Motion Correction and Signal Enhancement in Optical Coherence Tomography

Bewegungskorrektur und Signalverbesserung in der Optischen Kohärenztomographie

Der Technischen Fakultät
der Friedrich-Alexander-Universität
Erlangen-Nürnberg

zur

Erlangung des Doktorgrades Dr.-Ing.

vorgelegt von

Martin Kraus

aus

Forchheim, Deutschland

Als Dissertation genehmigt
von der Technischen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung:	19.12.2016
Vorsitzende des Promotionsorgans:	Prof. Dr.-Ing. Reinhard Lerch
Gutachter:	Prof. Dr.-Ing. J. Hornegger
	Prof. Dr. James G. Fujimoto, M.I.T.

Abstract

Optical Coherence Tomography (OCT) is a non-invasive optical imaging modality with micron scale resolution and the ability to generate 2D and 3D images of the human retina. OCT has found widespread use in ophthalmology. However, motion artifacts induced by the scanning nature restrict the ability to have reliable quantification of OCT images. Furthermore, OCT suffers from speckle noise and signal quality issues.

This work addresses these issues by treating the motion correction problem as a special image registration problem. Two or more 3D-OCT volumes with orthogonal scan patterns are acquired. A custom objective function is used to register the input volumes. As opposed to standard image registration, there is no reference volume as all volumes are assumed to be distorted by motion artifacts. To improve the robustness of the correction algorithm, multi-stage and multi-resolution optimization, illumination- and tilt-correction and custom similarity measures and regularization are employed. After registration, the corrected volumes are merged and a single volume with less noise is constructed by adaptively combining the registered data.

A large-scale quantitative evaluation was performed using data acquired from 73 healthy and glaucomatous eyes. Three independent orthogonal volume pairs for each location of both the optic nerve head and the macula region were acquired. The results of two motion correction algorithm profiles were compared with performing no motion correction. The evaluation measured registration performance, reproducibility performance and signal improvement using mutual information, error maps based on the difference of automatic segmentation of retinal features and a no-reference image quality assessment. In all three of these aspects, the proposed algorithm leads to major improvements, in accordance with visual inspection. For example, the mean blood vessel map reproducibility error over all data is reduced to 47% of the uncorrected error.

The algorithm has been deployed to multiple clinical sites so far. In addition, the technique has been commercialized. The main application is structural imaging for clinical practice and research. The removal of motion artifacts enables high quality en face visualization of features. The technique has also been applied to hand held OCT imaging and small animal imaging. Furthermore, applications in functional imaging in the form of intensity based angiography and Doppler OCT have been demonstrated.

Overall, the motion correction algorithm can improve both the visual appearance and the reliability of quantitative measurements derived from 3D-OCT data substantially. This promises to improve diagnosis and tracking of retinal diseases using OCT data.

Kurzübersicht

Optische Kohärenztomographie (OCT) ist ein nichtinvasives optisches bildgebendes Verfahren mit mikrometergenauer Auflösung und der Möglichkeit 2D und 3D Bilder der menschlichen Retina zu erzeugen. OCT ist weit verbreitet in der Augenheilkunde, allerdings behindern Bewegungsartefakte die durch Scannen induziert werden eine zuverlässige Quantifizierung von OCT Bildern. Weiterhin leidet OCT an Speckleraushen und Problemen der Signalqualität.

In dieser Arbeit werden beide Probleme adressiert indem das Bewegungskorrekturproblem als ein spezielles Bildregistrierungsproblem behandelt wird. Zwei oder mehr 3D-OCT Volumen mit orthogonalen Scanmustern werden aufgenommen. Eine spezielle Zielfunktion wird zum Registrieren der Eingangsvolumen benutzt. Im Vergleich zu üblichen Registriermethoden gibt es kein Referenzvolumen da angenommen wird, dass alle Volumen durch Bewegungsartefakten verzerrt sind. Um die Robustheit des Korrekturalgorithmus zu verbessern werden Mehrstufen- und Mehrfachauflösungsoptimierung, Beleuchtungs- und Neigungskorrektur sowie eine spezielle Ähnlichkeitsmetrik und Regularisierung verwendet. Nach der Registrierung werden die korrigierten Volumen zu einem einzelnem Volumen mit reduziertem Rauschen verschmolzen indem die registrierten Daten adaptiv kombiniert werden.

Eine groß angelegte Evaluation mit Daten von 73 gesunden und glaukomatösen Augen wurde durchgeführt. Drei unabhängige orthogonale Volumenpaare von den Regionen des Sehnervs und der Makula wurden aufgenommen. Die Ergebnisse von zwei Bewegungskorrekturalgorithmusprofilen wurden mit keiner Korrektur verglichen. Die Evaluation maß Registrierungsleistung, Reproduzierbarkeitsleistung und Signalverbesserung mittels Mutual Information, Fehlermaps basierend auf der Differenz von automatischen Segmentierungen von Merkmalen der Retina sowie einer referenzlosen Bewertung der Bildqualität. In allen drei dieser Aspekte führt der vorgeschlagene Algorithmus in Übereinstimmung mit visueller Begutachtung zu großen Verbesserungen. Zum Beispiel reduzierte sich der mittlere Blutgefäßmapfehler der Reproduzierbarkeit über alle Daten auf 47 % des unkorrigierten Fehlers.

Der Algorithmus wird mittlerweile in mehreren Kliniken eingesetzt. Außerdem wurde der Algorithmus kommerzialisiert. Die Hauptanwendung ist strukturelle Bildgebung im klinischen Alltag und der Forschung. Die Entfernung von Bewegungsartefakten ermöglicht eine qualitativ hochwertige en face Visualisierung von Features. Der Algorithmus wurde auch für tragbares OCT und Kleintierbildgebung eingesetzt. Weiterhin wurden Anwendungen im Bereich der funktionalen Bildgebung in der Form von Intensitätsbasierter Angiographie und Doppler OCT demonstriert.

Zusammengefasst kann der Bewegungskorrekturalgorithmus sowohl die offensichtliche Bildqualität als auch die Zuverlässigkeit von quantitativen Messungen auf 3D-OCT Daten substanziell verbessern. Dies verspricht die Diagnose und die Verfolgung von Krankheiten der Retina mittels OCT zu verbessern.

Acknowledgments

I'd like to thank a number of people who have helped me one way or another to write this thesis. First of all, I'd like to thank Prof. Dr. Joachim Hornegger and Prof. James G. Fujimoto, PhD who gave me the opportunity, support and advice that was necessary to successfully work on this project. Also, I'd like to thank Prof. Dr. Bjoern Eskofier and Prof. Dr. Elmar Nöth for putting the trust in me that led me to pursue a doctorate in the first place. Furthermore, I'd like to acknowledge Markus Mayer and Rüdiger Bock who were my advisors on the diploma thesis which started this project. My colleagues from the Pattern Recognition Lab deserve thanks for the great working atmosphere, good conversations and great celebrations. Especially, I thank Lennart Husvogt and Thomas Köhler for proofreading and providing valuable input and my long time great room mate André Aichert.

This work would not have been possible without the excellent collaborators. Therefore, I'd like to thank Jay S. Duker, MD, Gadi Wollstein, MD, Joel S. Schuman, MD, David Huang, MD, PhD, Prof. Dr. Wolfang Drexler, Prof. Dr. Georg Michelson and their groups for the great collaboration and experience.

During my work I was fortunate enough to be part of the Erlangen School of Advanced Optical Technologies (SAOT) and the DFG Training Group 1773 "Heterogeneous Image Systems". These provided funding, education and the ability to exchange ideas with other graduate students. I'd also like to gratefully acknowledge funding from the German Research Foundation within the grant DFG-HO-1791/11-1.

During my travels I've had the pleasure of meeting and working with many great people that also became friends. For this I'd like to thank, in no particular order, Marieh Esmaelpour, PhD, Lauren Branchini, MD, Teresa Torzicky, PhD, Ben Potsaid, PhD, Jonathan Liu, PhD, Bernhard Baumann, PhD, Kenny Tao, PhD and Michael Giacomelli, PhD.

Finally, I'd like to thank my family and friends for the support. Especially, I'd like to thank my parents for letting me go my own way and providing me with the safety and support that was necessary. This work is dedicated to my grandfather Adolf Kronas.

Martin Kraus

Contents

CHAPTER 1

Introduction

1.1 Background

The human eye is one of the most important sensory organs. Good vision is important for the quality of life, be it for social interaction, information reception, mobility and other things. Therefore, the diagnosis and treatment of diseases of the eye presents a worthwhile goal. Both diagnosis and treatment are facilitated by the availability of reliable imaging modalities.

Optical Coherence Tomography (OCT) [Huan 91] is a modality that has become a clinical standard in opthalmologic care. Being an optical technique, it is uniquely suited to the transparent tissue found in the eye. The ability to do 2D and 3D imaging, micrometer scale resolution, high sensitivity and non-invasiveness are key features for its widespread use, among others [Drex 15]. However, the reliability of OCT data is negatively influenced by aspects of motion artifacts and speckle noise. Motion artifacts result from the fact that OCT data sets are typically acquired over multiple seconds combined with involuntary motion of the eye relative to the imaging instrument during acquisition. Due to the scanning nature of OCT this leads to distortions in the obtained data, which cause inaccuracies in quantitative measurements that are extracted from the data set. Moreover, Speckle noise is inherent to the detection method used in OCT. It leads to a grainy look of OCT images and effectively lowers image signal-to-noise ratio (SNR) and resolution. This too has a negative influence on data quality and the reliability of obtained measurements.

Several approaches to deal with these problems have been developed so far. Common approaches are to either try to improve the OCT hardware itself with regards to these effects or apply special software methods in post processing. Most of the existing methods either lead to a significant increase in cost and complexity of OCT systems or are limited in the ability to correct the data.

1.2 Scope of the Work

This work is primarily concerned with finding ways to eliminate motion artifacts in 3D-OCT volume data. For this purpose, a post processing based algorithm is developed that takes two or more 3D-OCT volumes as input. No additional hardware which would increase the complexity and cost of the OCT system should be required. However, the different input volumes may have been scanned with different scan patterns, especially using so-called *orthogonal raster scans*.

Special care has to be taken to fulfill multiple, potentially conflicting criteria as best as possible. These are:

- The ability to correct motion with high precision, while being able to deal with large motion.
- Ease of integration into existing OCT systems.
- Robustness of the algorithm with respect to low input data quality, data inconsistencies etc.
- Practicality of the algorithm in a clinical setting through aspects such as computation time, ease of use etc.

In addition to producing motion corrected volumes based on these constraints, the output volume also should have increased signal quality and reduced speckle noise levels.

1.3 Contribution to the Progress of Research

The main contribution of this work is the introduction, evaluation and application of a novel 3D-OCT motion correction algorithm. An algorithm was developed in collaboration with Prof. James G. Fujimoto's group at the Massachusetts Institute of Technology (MIT). It uses two or more 3D-OCT volumes with orthogonal raster scan patterns as input. Utilizing a novel registration framework that operates without a fixed reference, the volumes are registered to a common space and at the same time motion corrected. Special regularization based on the time structure of the OCT acquisition process for each volume is employed within a custom objective function. Registration is performed by optimizing said objective function using non-linear and multi-resolution optimization techniques. After registration, the intensity information of the registered volumes can be combined into a single merged volume using an adaptive weighted sum. This results in increased SNR. The merged output volume is motion corrected, has improved signal quality and reduced speckle noise. This initial contribution was based on the diploma thesis of the author [Krau 09] and has been published in an extended form in [Krau 12].

A joint patent application between the University of Erlangen and MIT based on this method was also filed [Krau 11] and was subsequently licensed exclusively to Optovue Inc., Fremont, CA, USA. Meanwhile, the patent issued in the United States [Krau 16].

As part of the collaboration with MIT, the author also helped in developing custom OCT acquisition software for a novel prototype swept source OCT system, in addition to working on motion correction. This collaboration led to multiple joint publications [Baum 11a, Tsai 11a, Jia 12, Ahse 13, Liu 13b, Nadl 13, Tsai 13a, Liu 13a, Tsai 13b, Seba 12, Tan 12, Tsai 11b, Baum 11b, Wang 13, Wang 14a, Jia 14a, Wang 14b].

In addition, the work led to a collaboration with the lab of Prof. Wolfgang Drexler at the Medical University of Vienna which also resulted in multiple joint publications [Kaji 13, Esma 14],

Ongoing work to improve the robustness and performance of the algorithm on real clinical OCT data led to the development of an advanced correction algorithm. This algorithm introduced several enhancements such as a two-stage, multi-resolution optimization process with tilt correction, robust intensity similarity measures and regularization and illumination correction, among others. In order to evaluate the algorithm within a clinical setting, a large-scale quantitative evaluation was performed. This work led to another first-author publication [Krau 14].

Finally, as part of ongoing collaboration and due to the ease of integration and quality of the resulting data several clinical and pre-clinical research studies and method have been augmented through use of the method [Adhi 14, Ferr 14, Jia 14b, Liu 14, Alas 15]. Also, a chapter on OCT motion correction was contributed to a standard OCT book [Drex 15].

The key contributions of this work are:

- Knowledge in advanced image processing was applied to the field of Optical Coherence Tomography.
- A specialized registration approach that can be considered novel in both fields was developed for solving the problem of motion artifacts and signal quality in OCT.
- A large scale quantitative evaluation was performed, showing the clear advantage of the method.
- The developed algorithm is fully automatic, easy to use and fast enough to be used in clinical practice.
- Based on these achievements, the technology could already be commercialized and integrated into an OCT product.

1.4 Structure of this Work

The structure of this work is as follows: The first part is concerned with the fundamentals of technical and medical OCT. First, OCT itself is explained. Subsequently, we focus on OCT in the context of ophthalmologic imaging, specifically of the retina. The final chapter of this first part describes the body of prior work that exists both in OCT motion correction and signal improvement.

The second part of this work is concerned with 3D-OCT motion correction using image registration methods and orthogonal scan patterns. First, the motion correction approach itself is described in detail. Next, the approach to evaluate the proposed method is described. Subsequently, results are presented and discussed. The final chapter of this part describes further applications of the algorithm.

The work closes with a part containing an outlook for future areas of research and challenges and finally summary and conclusion.

Part I

Fundamentals in Technical and Medical OCT

CHAPTER 2

Optical Coherence Tomography

In the following chapter an overview of the technical background of OCT technology is given. For the design of post-processing algorithms it is essential to understand the key features and limitations of the imaging modality at hand.

2.1 Basic Principle and Time Domain OCT

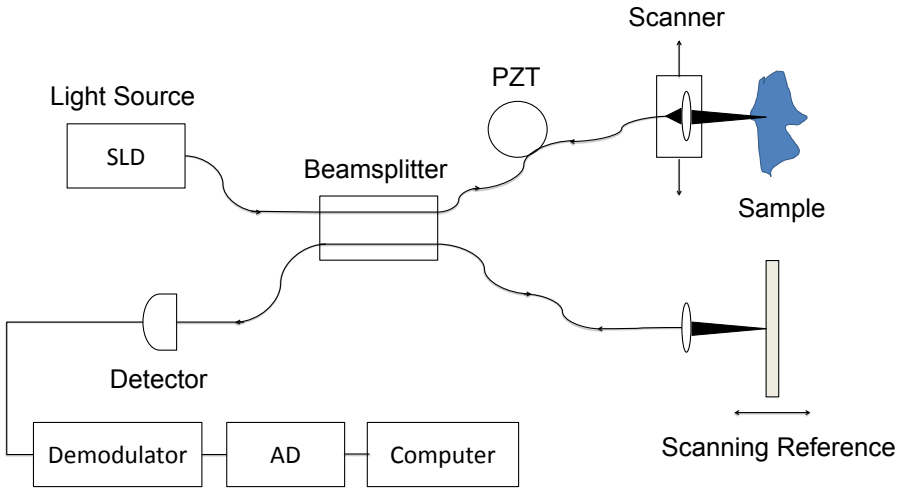

Figure 2.1: Early OCT system. Drawing based on [Huan 91].

OCT is closely related to low coherence interferometry [Ferc 86, Ferc 88] and femtosecond ranging [Fuji 86]. A schematic view of the first OCT system [Huan 91] is depicted in figure 2.1. In today's nomenclature, this system belongs to the class of Time Domain OCT (TD-OCT). The system is based on a Michelson type interferometer. A beam from a light source with broad spectral bandwidth, in this case

a Superluminescend Diode (SLD) is split into two parts in a beam splitter. Part of the light is directed into the so-called *sample arm* (lower right part of figure 2.1). In the sample arm, optics ensure that the light is focused onto the object. Also, a scanning mechanism enables the lateral scanning of the beam over the sample. The other part of the light is directed into the so-called *reference arm* (upper right part of figure 2.1). In the reference arm, the light is reflected back by a mirror. The optical path length of the reference arm is continuously varied by scanning this surface back and forth.

In the sample arm, part of the light is reflected back by the object, depending on its optical properties. The reference arm reflects the light after having traveled a defined distance. Light reflected from both arms moves back into the beam splitter and interferes there. The interference signal is then detected by a photo diode, and converted to a digital signal and fed into a computer for further processing and display. The system of Huang also used a piezoelectric transducer (PZT) and a demodulator to perform heterodyne detection, increasing SNR.

Light returning from the sample and reference arm will only show a clear interference signal if the difference in optical path length between the path traveled in the sample arm and the one traveled in the reference arm is within the *coherence length* of the light emitted by the light source. If the path lengths are not well matched up to the coherence length of the source, the interference signal rapidly vanishes [Huan 91]. OCT uses low coherence light sources with corresponding low coherence length of typically only a few microns. The amplitude of the interference signal is also dependent on the amount of light that is reflected by the sample and reference arm. The reflectivity of the reference arm can be assumed to be constant. Therefore, for a particular reference arm delay, the amplitude of the interference signal will depend mainly on the amount of light that is back reflected from the sample and that is path length matched up to the coherence length. This allows OCT to effectively detect only back-reflections corresponding to a certain axial depth in the object. The axial direction is the direction along the propagation direction of the beam into the object. OCT scans the optical path length of the reference arm and measures the amplitude of the interference signal. This corresponds to measuring the amount of back-reflection from the object with respect to different axial depths along the light beam going into the object. This 1D profile of back reflected light in relation to axial depth is called an axial scan or *A-scan*. By scanning the sample arm beam in *lateral* direction while acquiring 1D A-scans through scanning of the reference arm, 2D and 3D images of the object can be acquired.

Key parameters in OCT operation are:

- Imaging speed: Measured in A-scans per second. The faster the system, the less time is spent acquiring an image with the same number of transverse samples.
- Sensitivity: The smallest fraction of sample arm light that is back-reflected and that can still be detected. Sensitivity is usually inversely related to imaging speed.
- Imaging resolution: Full-width-half-maximum (FWHM) of the point spread function (PSF) in axial and transverse direction

- Imaging range: The maximum optical path length difference that is still contained within the A-scan. In TD-OCT this is only limited by the range of scanning of the reference mirror.

- Operating wavelength: The wavelength of the light used for imaging. This determines the penetration depth into different materials and tissues and also absorption, scattering, dispersion, etc.

- Noise: Both electronic noise and speckle noise which is caused by coherent detection.

An interesting aspect of OCT is that resolution in the axial and transverse directions is decoupled. For example in confocal microscopy [Webb 90, Mast 98], the axial resolution is determined by the numerical aperture (NA) of the imaging beam. In OCT, however, the axial resolution is determined by the coherence length of the light source while the transverse resolution is by the NA. This allows OCT to also image with high axial resolution when NA is limited [Huan 91].

2.2 Fourier Domain OCT

In Fourier-Domain OCT (FD-OCT) the interference signal is spectrally resolved [Ferc 95, Haus 98]. It is based on the observation that the inverse Fourier transform of the spectral components of the interference signal gives rise to the A-scan information. As opposed to TD-OCT, the reference arm mirror does not need to be scanned. In addition, it was discovered that the detection of multiple spectral components has an inherent sensitivity advantage [Chom 03, Leit 03a, Boer 03]. Together, this enables much faster imaging speeds than TD-OCT. Hence, most commercial OCT systems nowadays use Fourier domain detection.

2.2.1 Spectral Domain OCT

One way to acquire the spectral channels of the interference signal is to use a spectrometer instead of a single photo detector. Figure 2.2 depicts a schematic of a spectrometer based OCT system. As is common in FD-OCT, the reference mirror does not need to be scanned anymore and the spectrometer replaces the photo detector. The different channels from the spectrometer are usually acquired by a line scan camera and then sent to the computer for further processing.

While Spectral Domain OCT systems can be much faster than TD-OCT systems, the imaging range in a spectrometer based system is dependent on the ability of the spectrometer to separate the different spectral components and the number of spectral channels that are acquired in total [Ferc 03]. Hence, the imaging range tends to be more limited which is why TD-OCT systems still have a niche in applications that require a large imaging range with good axial resolution.

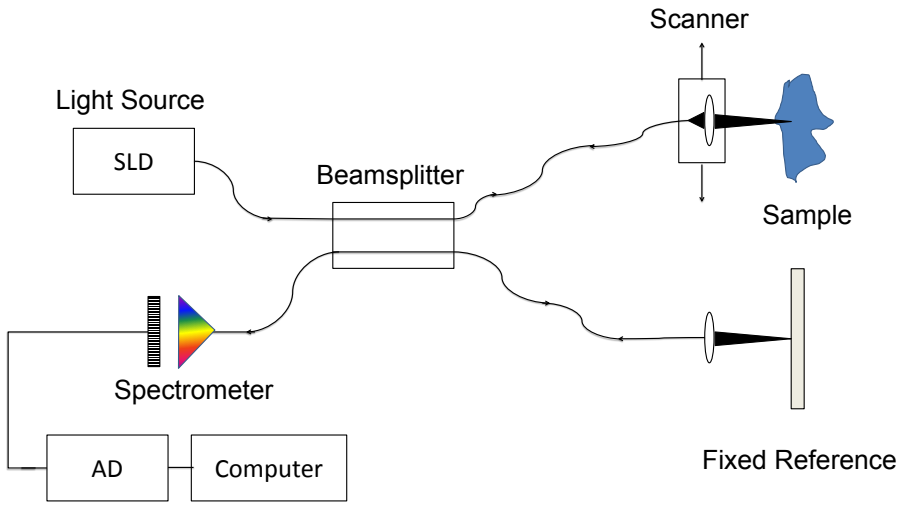

Figure 2.2: Camera based Spectral Domain OCT system schematic.

2.2.2 Swept Source OCT

Another principal way to perform FD-OCT is to employ a swept light source or swept source [Chin 97, Habe 97]. Such a light source instantaneously emits near monochromatic light but the wavelength of the light is swept over time. Instead of a spectrometer, a single photo detector can be used again. This is because the spectral components of the interference signal are encoded in time. The imaging range in this case is determined by the ability of the light source to emit very narrow instantaneous spectra and the detection bandwidth of the system [Ferc 03]. Figure 2.3 depicts a schematic of a swept source OCT system.

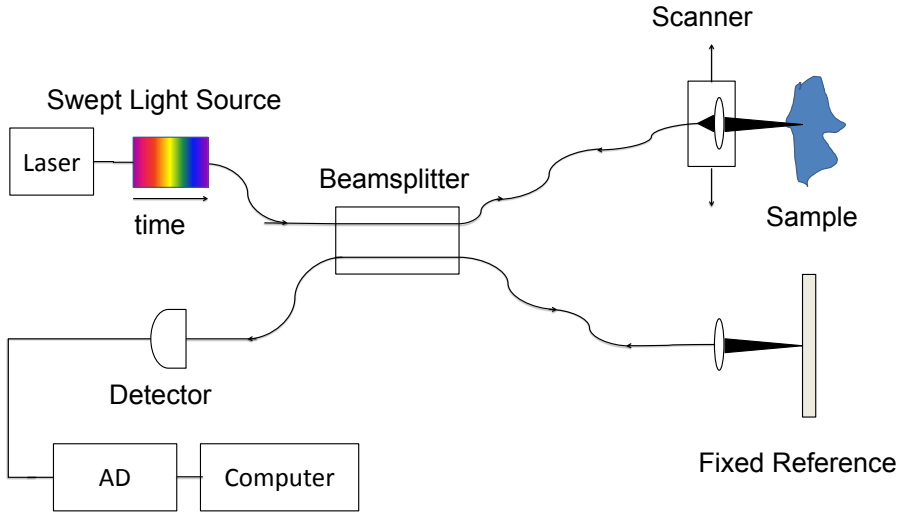

Figure 2.3: Swept Source OCT system schematic.

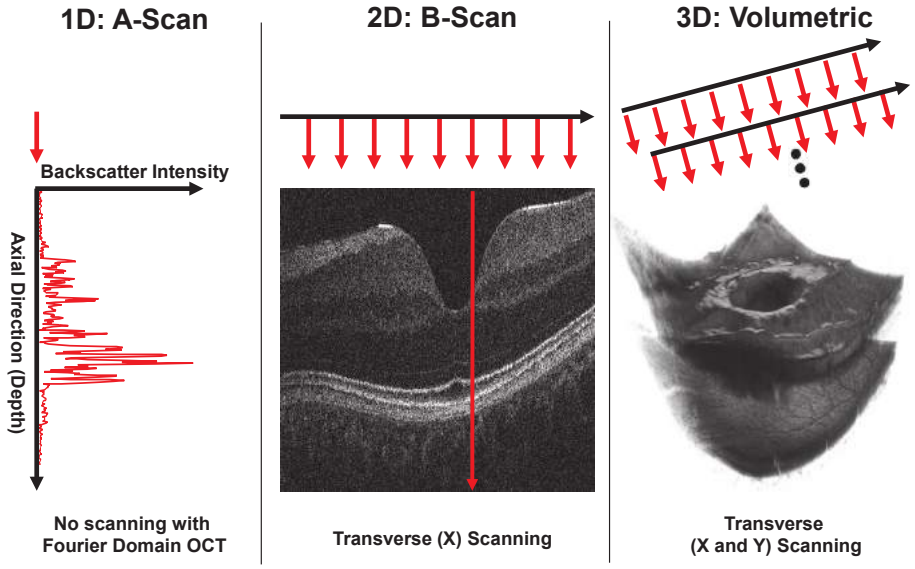

Figure 2.4: OCT scanning schematic.

2.3 OCT Scanning

Figure 2.4 shows how scanning can be performed in OCT in order to create multi-dimensional images. OCT systems typically have two galvanometer mirrors that allow the beam to be moved in the transverse (X,Y) plane in a programmable fashion. This motion is controlled by a *scan pattern* which specifies the trajectory of scanning. Assuming FD-OCT, no scanning is necessary to obtain 1D A-scans. The scanner coordinate system is spanned by the axial direction (z-axis) and two orthogonal directions (x- and y-axis) which correspond to the degrees of freedom of the two galvanometer mirrors which position the imaging beam. In order to acquire a 2D image or *B-scan*, the beam is scanned laterally while acquiring A-scans. One example way to scan would be a linear scan, i.e. the beam trajectory is a line in the transverse dimensions of the scanner coordinate system. Other types of 2D images are possible, though. For certain applications, performing circular 2D scans is useful [Wang 09].

Given a sufficiently fast OCT system, 3D images can be generated by scanning such that a two-dimensional grid of transverse locations is traversed while acquiring A-scans. The simplest way to perform 3D scanning is to use a so-called *raster scan*. Figure 2.5 shows a schematic of a raster scan. A raster scan consists of a series of linear scans and is determined by two directions. The first direction is the so-called *fast direction* along which the linear B-scans are performed. After each linear scan,the beam is moved one step in the *slow direction*. This process is repeated until the whole regular grid of A-scan locations is traversed.

Due to the limited acceleration and frequency response of galvanometer mirrors used in OCT systems, not 100 percent of the time can be spent imaging. In order to connect segments, the scanner has to spend time changing position and velocity. During this time no A-scans are acquired. In raster scans the beam has to move back from the end position along the fast direction after each B-scan to the

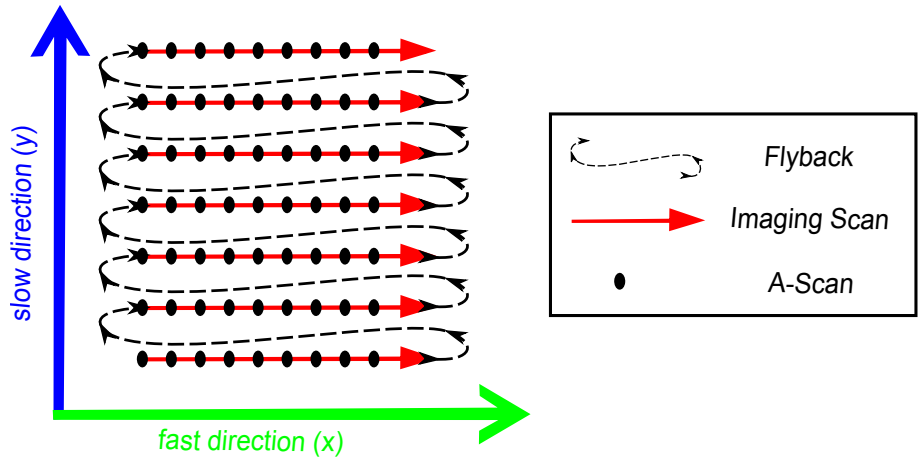

Figure 2.5: OCT raster scanning schematic.

start position for the next B-scan. The time spent for this repositioning is called *flyback time*.

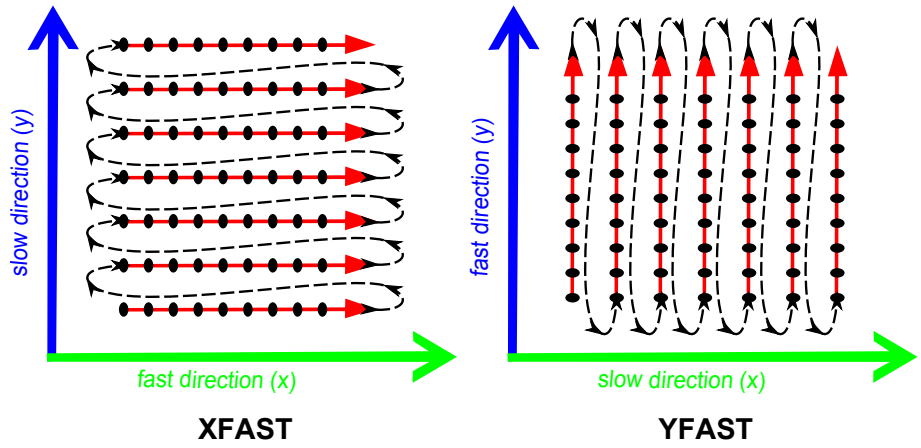

Figure 2.6: Orthogonal raster scanning.

A specific kind of scan patterns that are highly relevant for this work are so-called *orthogonal raster scans*. Specifically, in an XFAST scan pattern the x-axis is the fast scan direction while the y-axis is the slow direction. The YFAST type scan pattern switches these directions, making the y-axis the fast direction. The scan patterns are orthogonal because the 2D vectors that specify the fast scan direction in each pattern are orthogonal to each other. Using a pair of orthogonal scan patterns the same grid of scanner coordinate A-scan locations can be sampled. The difference being the *order* of traversal.

2.4 Summary

In this chapter, a technical background on OCT technology was given. The basic operation principle of OCT was explained using the first time domain OCT system as an example. Key imaging parameters were identified. Subsequently, the more

recent technology of FD-OCT in its incarnations of Spectral and Swept Source OCT were described. Finally, the formation of two- and three dimensional OCT images by the use of lateral scanning while acquiring A-scans was introduced. Last but not least, raster scanning for 3D imaging and orthogonal raster scans in particular were described.

CHAPTER 3

Ophthalmologic Imaging using OCT

In the following chapter, we will take a deeper look at the use of OCT in the context of ophthalmologic practice. OCT is widely used as a tool for diagnosing eye diseases and monitoring their progression and the response to treatment. In this context it is important to have a basic understanding of the anatomy of the human eye. The use of OCT in diagnostic imaging will be discussed. OCT images exhibit characteristic image features which are relevant clinically and when performing image processing. Eye motion plays a key role in visual perception. Yet, motion artifacts that are caused by them are a key problem that limit the reliability of quantitative measures derived from OCT data.

3.1 Basic Anatomy of the Eye

A simplified view of the anatomy of the eye is shown in figure 3.1 [Bomm 06]. The outer surface of the eye is made up by the white tissue of the sclera. In the front of the eye the outer surface is not made up from sclera but from transparent tissue called the cornea. Here, light enters the eye. Behind the cornea there is the anterior chamber. The iris contains the pigment which determines eye color. It acts as a variable shutter that can allow a varying amount of light to enter the eye. Light that is not obscured by the iris passes through the lens. Shape and refractive properties of the lens and the cornea cause the light to be focused onto the back of the eye where the retina is located. Before reaching the retina, the light passes through the transparent matter of the vitreous body. The retina contains photoreceptors that sense incoming light and generate biological signals. These signals are transported towards the optic nerve head (ONH) via nerve fibers. The nerve fibers converge into the optic nerve where they exit the eye towards the brain. The ONH is also the location where retinal vessels that supply the retina

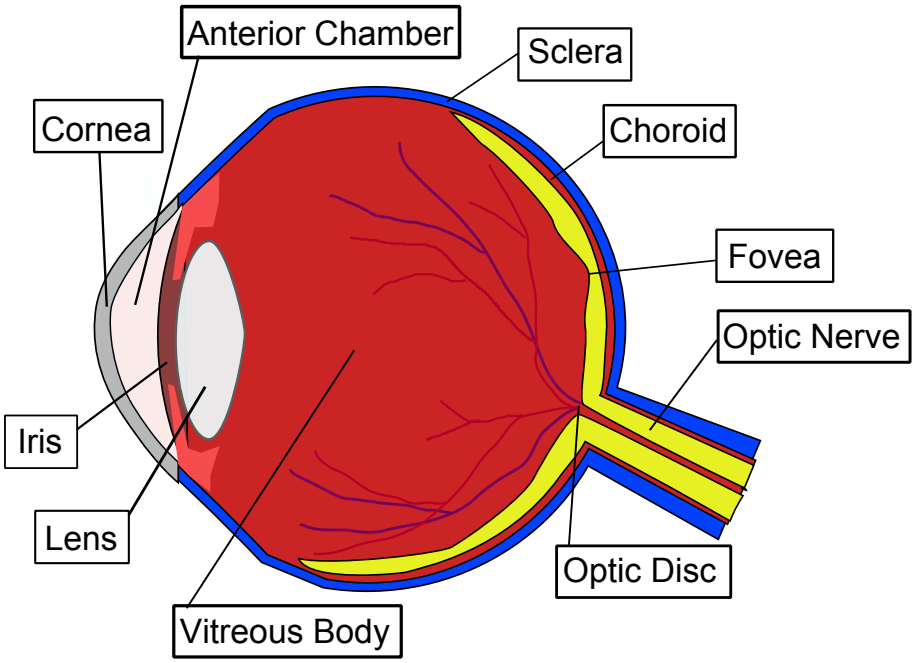

Figure 3.1: Simplified anatomy of the human eye, based on [Bomm 06].

enter and exit the eye. Additional blood vessels are located behind the retina in the so-called choroid. Last but not least, the retina has an area of best vision called the fovea. Here the density of photoreceptors is maximal. The fovea itself is located approximately on the optical axis of the lens.

Figure 3.2 shows a so-called fundus photograph of a human retina as seen through the front of the eye by a fundus camera. The ONH is visible as a yellow circular area from which blood vessels emerge and spread across the retina. The fovea can be seen as a slightly darker region that is also free of blood vessels, the so-called foveal avascular zone.

The retina itself is a layered structure of different functional tissues and is between 0.1 mm and 0.56 mm thick [Rior 08].

> "The layers of the retina, starting from its inner aspect, are as follows: (1) internal limiting membrane; (2) nerve fiber layer, containing the ganglion cell axons passing to the optic nerve; (3) ganglion cell layer; (4) inner plexiform layer, containing the connections of the ganglion cells with the amacrine and bipolar cells; (5) inner nuclear layer of bipolar, amacrine, and horizontal cell bodies; (6) outer plexiform layer, containing the connections of the bipolar and horizontal cells with the photoreceptors; (7) outer nuclear layer of photoreceptor cell nuclei; (8) external limiting membrane; (9) photoreceptor layer of rod and cone inner and outer segments; and (10) retinal pigment epithelium" [Rior 08, p. 13]

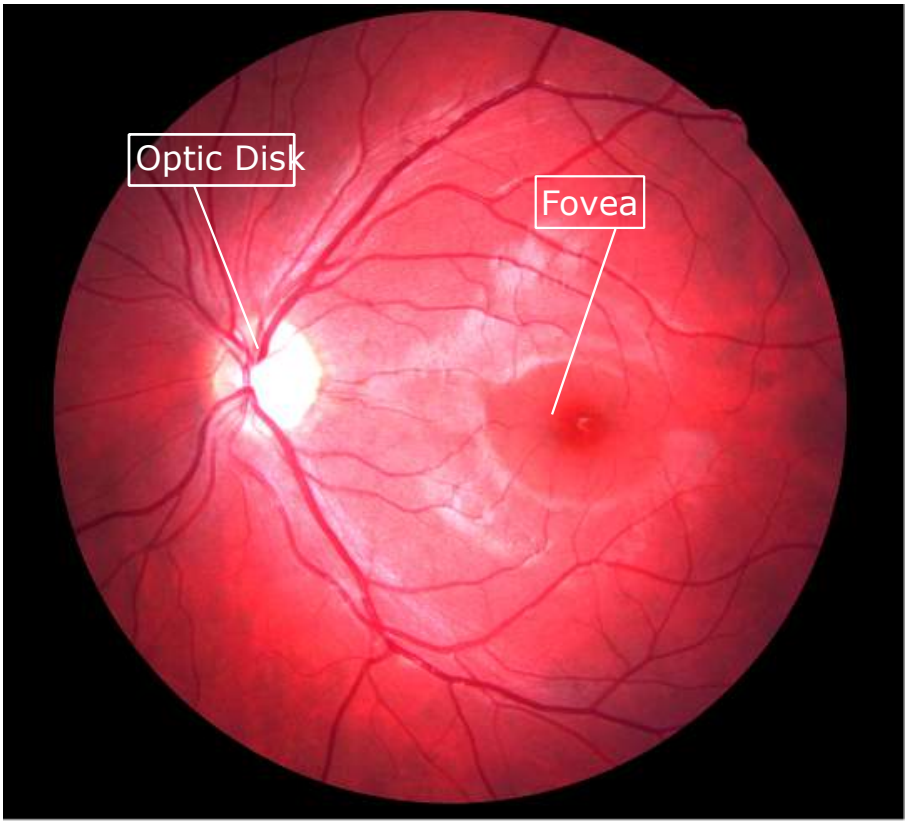

Figure 3.2: Fundus photograph of a human eye.

3.2 Eye Motion

The eye is located inside the eye socket and there are several muscles that allow the eye to rotate within the socket. Through this mechanism, the optical axis of the eye and with it the foveal zone of best vision can be moved to point at different objects in the world. This is called fixation. A fast involuntary change in fixation is called a *saccade*.

In addition to macroscopic and voluntary motion such as fixating on an object there is also smaller, involuntary motion of the eye. Eye motion plays a key role in being able to see at all. Due to a process called *neural adaptation*, a static stimulus on the retina will cause neural activity to fade within a short time. The neural system *adapts* to the stimulus, there is no more excitation [Mart 04].

To counteract neural adaptation, involuntary eye motion causes the fixation to change over time. This moves the image that is projected onto the retina. The image which is presented to the photoreceptors and subsequent neurons is not static anymore, hence avoiding adaptation [Mart 04].

There are three main types of involuntary eye motion, namely tremor, drifts and microsaccades. Tremor is an aperiodic, wave-like motion with a frequency of about 90Hz and very small amplitude [Mart 04]. Drifts occur together with tremor and in between microsaccades [Mart 04]. Finally, microsaccades are “small, fast, jerk-like movements“ [Mart 04]. Their frequency ranges from 0.5 to Hz, with a typical duration of about 25 ms [Mart 04]. The amplitude of microssades ranges

from 5 min up to 2 degrees in fixation angle. Other sources report that saccadic motion can lead to a change in fixation angle of up to 4 degrees [Pova 09, Engb 03].

3.3 Retinal OCT Scanning

Imaging of the retina is one of the main uses of OCT in ophthalmology. OCT is also used to image the anterior part of the eye, but this use case is less important currently. As mentioned before, when scanning the retina with OCT, the eye itself is part of the optics of the system [Swan 93]. Figure 3.3 shows a schematic of how scanning the retina is performed. The OCT imaging beam is shown in red. Before reaching the retina, the OCT beam has to pass through the cornea and lens, not be blocked by the iris and pass through the vitreous body. The OCT system sends a collimated beam onto the eye in an angle β with respect to the optical axis of the eye [Swan 93]. The cornea and the lens together refract and focus the beam onto the retina. The beam then reaches the retina under an angle γ relative to the optical axis. Using a first-order approximation, the relationship between these two angles is linear. Therefore by varying the incident angle β, the beam can be scanned over the retina. When the OCT system and the eye are in good alignment with respect to each other, the beam will pivot around a fixed point (shown as a black dot in each view) independent of the incident angle. Therefore, a linear scan in the scanner coordinate system (see section 2.3) will result in a fan-like geometry of the resulting imaging beams. The imaged area of neighboring A-scans will therefore not be exactly parallel. However, since the deviation from parallelism is very small and the retina itself is a curved surface this aspect is often omitted. Also, in reality there are two degrees of freedom instead of one, namely the x- and y- coordinates of the scanner coordinate system.

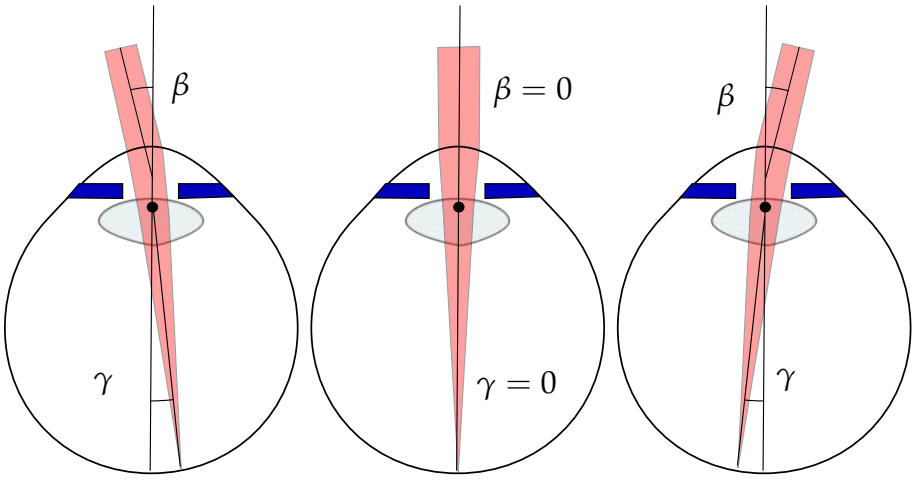

Figure 3.3: Ideal alignment for scanning an eye in OCT.

3.3.1 2D and 3D Imaging

Figure 3.4 shows a sample linear B-scan of a healthy human foveal area. The image consists of 1000 A-scans and was acquired with a prototype 850 *nm* based high

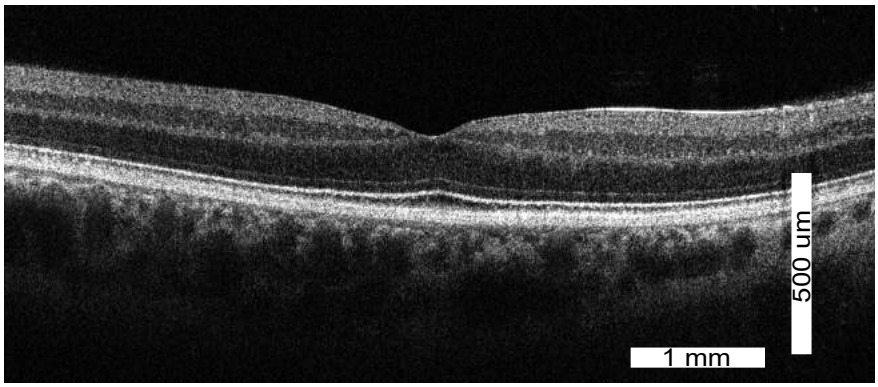

Figure 3.4: Sample linear B-scan of a human fovea.

speed ultra high resolution (UHR) OCT system [Pots 08]. Due to the high dynamic range of typical OCT images, the image is shown in log-scale, as is common for OCT. The region at the top is the vitreous body which does not reflect light, therefore it shows up as a dark region. Below, the retina is visible as a layered structure. In the horizontal center of the image the fovea can be seen as a pit like structure. The use of 2D images is very common for obtaining a qualitative view of a certain region of the retina. However, it is inherently difficult to align the system such that a single 2D slice captures focal pathology for example.

In addition to 2D diagnostic imaging the use of 3D-OCT is becoming more and more common in clinical practice. This is enabled by increased system speed. A key advantage of 3D-OCT is the more comprehensive data acquisition which makes it less susceptible to miss focal pathology. In addition, 3D-OCT enables the clinician to register fundus features with the volume through use of OCT *fundus projections* [Hitz 03]. Figure 3.5 shows different views of a 3D-OCT data set. Typical 3D-OCT volumes are acquired as a *raster scan* (see figure 2.5). Therefore, the volume data consists of a regular grid of A-scans in the *scanner coordinate system*. A *fundus projection* of a volume is a obtained by integrating the volume intensity data over the z-direction. This leads to a two dimensional image that corresponds to a fundus photograph (see figure 3.2). Figure 3.5 (a) depicts an OCT fundus projection. An alternative way of displaying a 3D-OCT volume is by using volume rendering techniques and is shown in figure 3.5 (b). Last but not least, 3D data allows for the extraction of arbitrary 2D slices out of the data. Figure 3.5 (c) and (d) show two central slices in the x-z and y-z planes, respectively.

3.3.2 Diagnostic Imaging

The first *in-vivo* imaging of the retina was performed using a time domain OCT system by Swanson et al. [Swan 93]. The development of FD-OCT systems led to a significant improvement of the practically obtainable axial resolution and imaging speed. Current commercial FD-OCT that are in clinical use provide scan speeds of around 25000 A-scans per second and axial resolutions of around 5 μm in tissue [Drex 08, Ho 09]. Also, OCT is a non-invasive technique as it is using safe levels of exposure to near infrared light, requiring no contact.

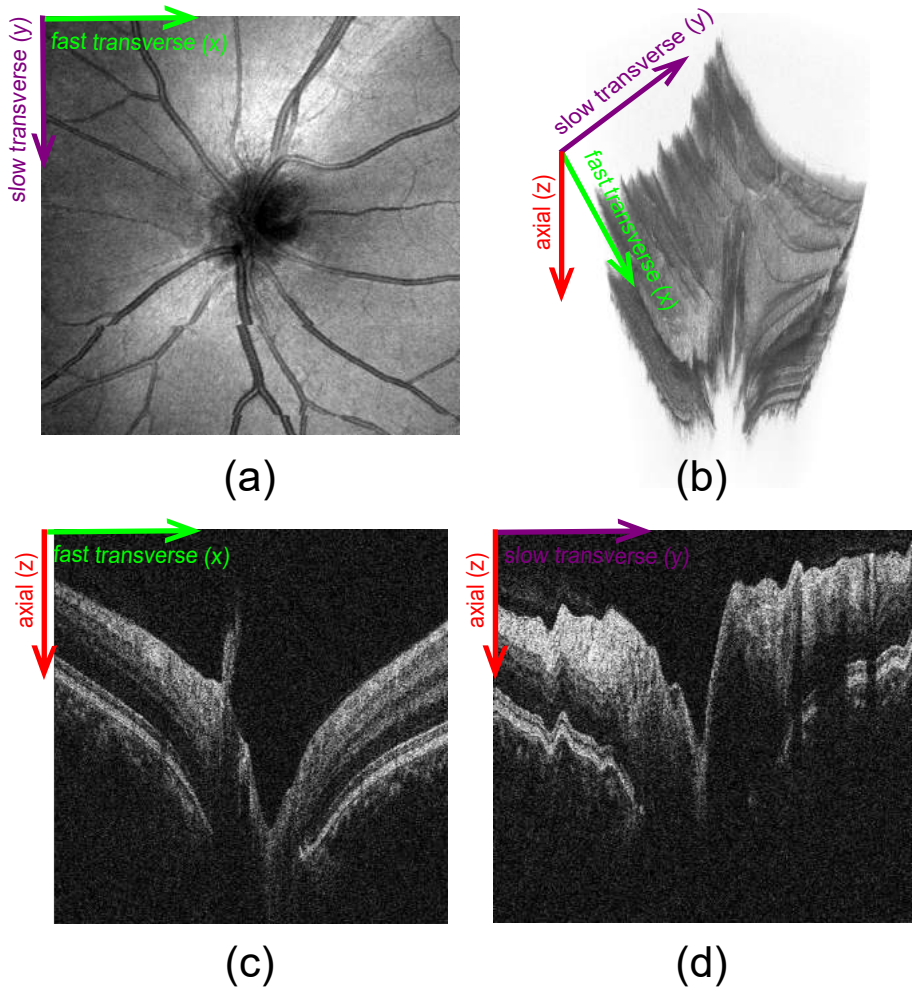

Figure 3.5: Different views of 3D OCT Volume: (a) Fundus projection (b) 3D rendering (c) Central slice along fast scan direction (d) Central slice along slow direction.

Many diseases of the retina such as glaucoma, age-related macular degeneration (AMD) and diabetic retinopathy manifest through changes in the layer structure of the retina. In some forms of glaucoma for example, there is a progressive thinning of the nerve fiber layer (NFL) around the ONH [Schu 95]. Another example is the formation of lipid accumulations, so-called drusen, that lift the retinal pigment epithelium (RPE) [Ryan 13, Grou 04]. These are indicative of some forms of AMD.

Therefore, being able to image the retina with high data quality and resolution plays a key role in diagnosis and management of diseases. Due to the transparent nature of most tissues of the eye, optical technologies are well suited for this task. When imaging the eye, the refractive part of the eye (cornea and lens) itself becomes part of the optical system, limiting the NA of the imaging system. This limits the axial resolution of conventional optical imaging methods such as confocal microscopes [Webb 90, Mast 98]. In OCT however, the axial resolution is dependent on the coherence length of the light source and therefore independent of NA. Therefore, the axial resolution can be much better than what would con-

ventionally be possible. This is important for discriminating the different layers of the retina and seeing the changes that occur in disease early.

In general, it is also important to be able to do this reliably and in a quantitative way. For example, one has to measure the thickness of a certain layer. In addition, the measurement has to be reliably associated with a location on the retina. An example scenario would be measuring NFL thickness on a 3.46 *mm* diameter circle around the ONH for glaucoma diagnosis [Schu 95]. If the thickness measurement is correct but it is not clear or uncertain where on the retina it was measured the use is limited.

3.4 Factors Influencing Image Quality

Good image quality and data that allow for reliable measurement of relevant parameters of the the subject's retina are key to the diagnostic ability of OCT. The OCT system itself, the subject's eye and the static and dynamic alignment due to motion give rise to several effects that influence the quality of OCT images and the reliability quantitative measurements. More specifically, the main image parameters that are influenced are (retinal) signal level, SNR and image distortion.

3.4.1 Speckle Noise

Similar to other imaging modalities such as ultrasound, OCT is affected by speckle noise [Good 76, Schm 99b]. "Speckle noise reduces contrast and makes boundaries between highly scattering structures in tissue difficult to resolve" [Schm 99b, p. 95]. Figure 3.6 shows a magnified area of figure 3.4. While layer boundaries are visible, the layers themselves are not homogeneous in intensity but have a grainy look to them. At the top of the image, where there is no retinal tissue, speckle noise still leads to individual pixels with relatively high intensities.

The speckle pattern in an OCT image can be seen as noise but is actually related to the imaged object and its micro-structure [Schm 99b]. In linear scale, the nature of speckle noise can be considered multiplicative [Wong 10]. In log-scale, in which OCT images are usually displayed, this reduces to an additive noise. However, while speckle noise in log-scale can be considered to have zero mean, it is not normally distributed [Kara 05, Bash 00]. Also, the speckle pattern is very sensitive to a number of factors such as incident angle of the beam, wavelength and polarization [Schm 99b]. As such, the speckle pattern changes when the same location is imaged twice with slightly different incident angle of the beam. Even the slightest motion between OCT system and subject will cause this.

3.4.2 Blinking

When a blink happens during image acquisition the eyelid will effectively totally block the beam path to the retina for a certain amount of time. Therefore, the A-scans acquired during this time will only show background and no retinal signal. Figure 3.7 shows views of an OCT volume with a blink happening during acquisition. As marked by the red bars, the effect of the blink can clearly be seen in the

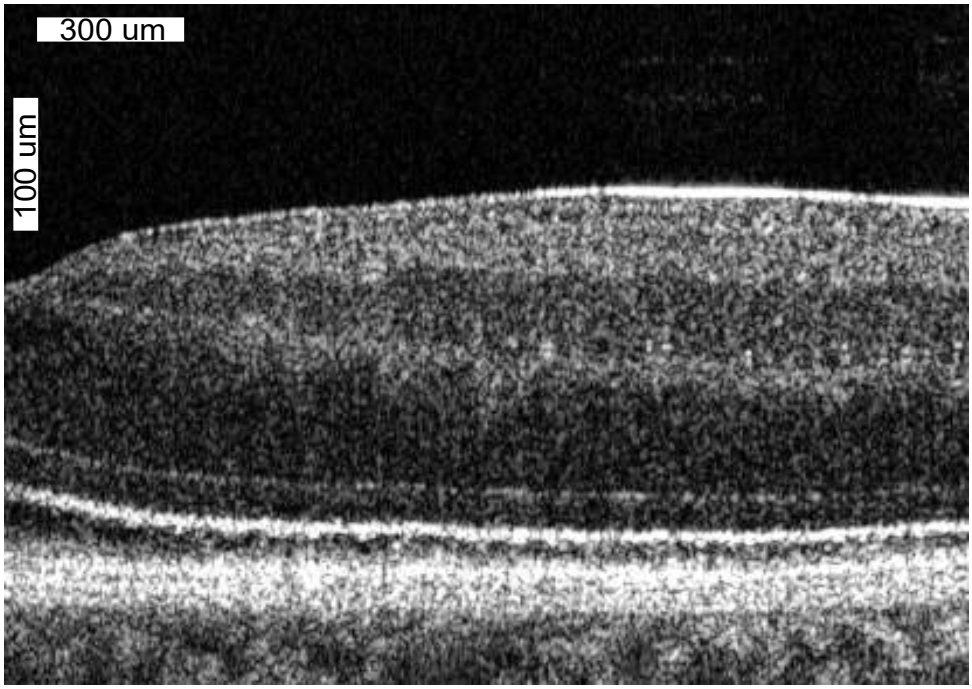

Figure 3.6: Zoomed excerpt of figure 3.4 showing speckle noise.

en face view (a) and in the slice along the slow direction (b), while a slice along the fast direction that was acquired while there was no blinking remains unaffected (c).

3.4.3 Illumination / Floaters

Depending on the concrete alignment between the OCT system and the eye, a certain fraction of the incoming light will reach the retina and be back-scattered by the tissue there. Of the back scattered light, again only a certain fraction will be collected back by the OCT system, interfere and be detected. Again, this is alignment dependent.

So-called floaters can block significant fractions of the incoming of outgoing light [Clin 80]. These are opacities that are located inside the vitreous body that can block the beam partially or fully. In addition, due to changes in alignment over time the shadow of a floater can move on the retina, blocking signal from different areas in two subsequent volume acquisitions. Usually, a single floater affects a relatively small transverse area. Another source of diminished signal are opacities in the lens such as in cataract [Velt 06].

Figure 3.8 shows en face views from two subsequently acquired volumes of the same subject. While the area of the retina that is covered is mostly the same, there are differences in illumination: For example the volume in (a) shows a relatively focal shadowed area (marked by the red arrows). The same area in the second volume (b) does not show this shadowing effect. The likely reason for this is that the beam path was blocked at these locations by a floater or opacity in the first volume. When the beam scans the same retinal area again for the second volume the

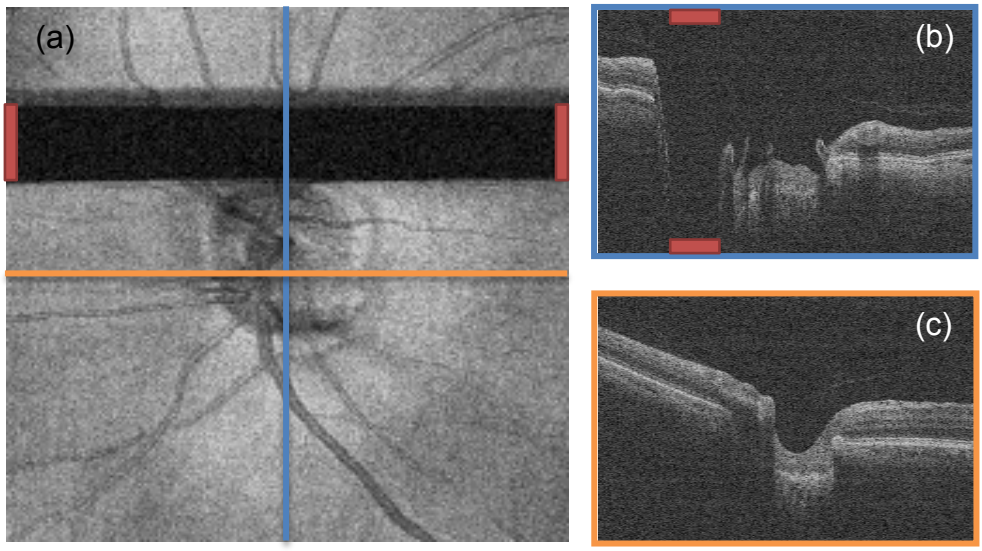

Figure 3.7: Views of an example OCT volume with blink during acquisition.

alignment of the OCT system with the eye has changed or the floater has moved in such a way that the area is no longer in shadow. Illumination differences for corresponding locations in subsequent volumes lead to inconsistencies in the observed intensities.

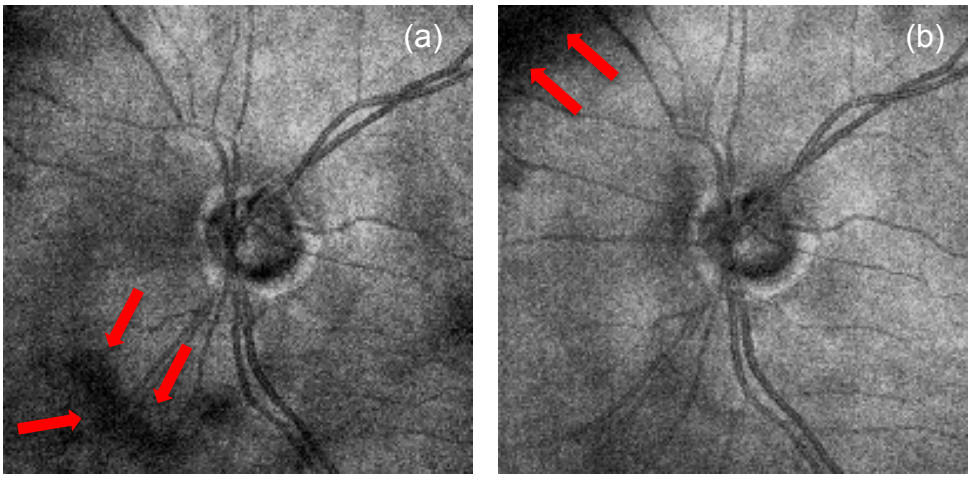

Figure 3.8: En face views of two subsequently acquired volumes exhibiting time dependent illumination effects.

3.4.4 Vignetting

The shadow in the marked area in figure 3.8 (b) could be a result of vignetting of the beam. As the incident angle of the OCT beam with respect to the optical axis changes, the beam can be blocked partially or fully by the iris. This effect is called vignetting. The actual incident angle is dependent on motion and therefore, the same anatomical location can be significantly vignetted in one volume acquisition

and while it is unaffected in a subsequent one. In addition, if the transverse area scanned by OCT is larger, it is generally harder to avoid vignetting effects.

3.4.5 Tilt

Due to static or dynamic alignment, the OCT might not be pivoting around a point on the optical axis as it is scanning (see section 3.3). In this case the optical path length until the retina is reached (assuming a circular retinal surface) becomes dependent on the scan angles. Figure 3.9 shows a schematic of the phenomenon. First, let us consider a configuration where the OCT beam pivots at the center of the lens, i.e. on the optical axis (middle in figure 3.9). Here, the optical path length until reaching the retina is approximately the same, hence the retina appears aligned with the horizontal axis in the corresponding B-scan. On the other hand, if the beam is not pivoting in the center (left and right in figure 3.9) the path length becomes longer to the respective other side of the retina, while becoming shorter to the same side. This leads to the characteristic tilting that can be seen in the corresponding B-scans.

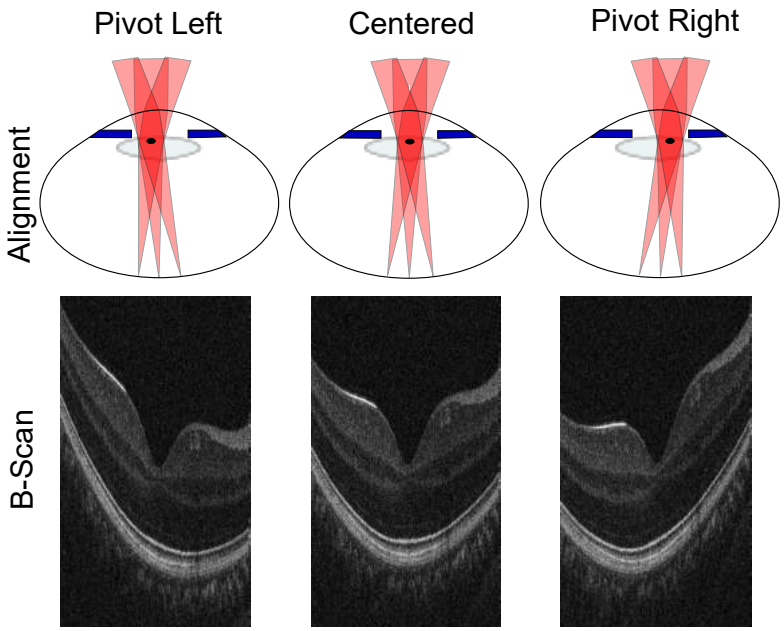

Figure 3.9: Schematic of alignment of the eye with respect to the OCT device and resulting tilting of the image (not up to scale).

To first order, an additional path length difference that is proportional to the incident angle(s) is the result. This results in a translation of the content of the corresponding A-scans in axial direction. In a linear B-scan, the retina will ”tilt” in accordance with a line of a certain slope. The appearance effect is amplified by the high axial resolution of OCT.

3.4.6 Motion Artifacts

Depending on the speed of the OCT system and the number of A-scans that need to be acquired, an OCT scan can take multiple seconds. For example, a 200×200 A-scan 3D raster scan will take approximately 2 seconds to acquire on a system with a speed of 20000 A-scans per second. If the eye moves relative to the OCT system during scanning, motion artifacts can occur. These result in spatial distortion of the acquired data. In addition, motion can change other parameters that influence image quality during a single acquisition and also between subsequent acquisitions. This can lead to inconsistent image parameters over time.

In general, relative motion between the OCT system and the subject's eye causes a time dependent change in the alignment of the beam with respect to the eye. This time dependent change of the optical configuration causes variation in the beam path and/or changes of the optical path length until the retina. Figure 3.10 shows views of an example volume exhibiting typical motion artifacts. We can distin-

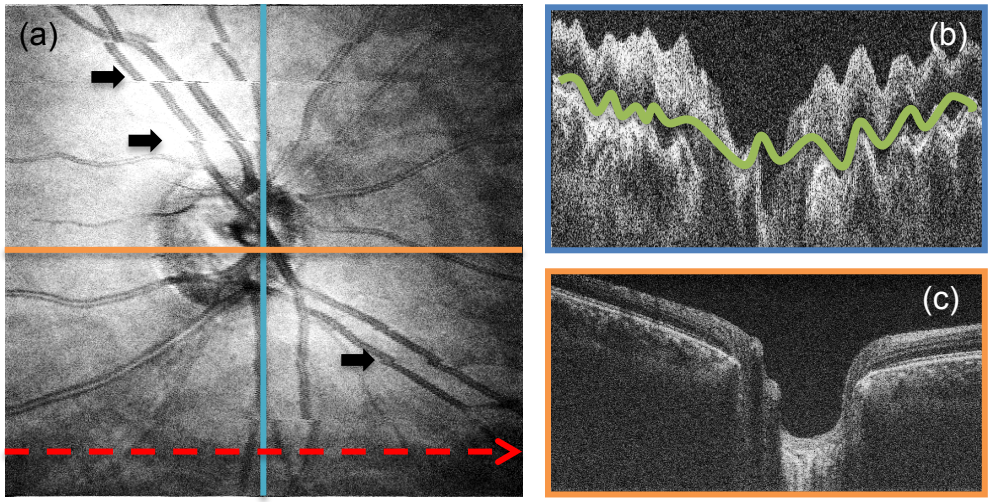

Figure 3.10: Example of motion artifacts in a 3D-OCT volume: (a) Fundus view (b) Central slice along the slow scan direction (c) Central slice along the fast scan direction. The red line in (a) shows the fast scan direction. The green line in (b) symbolizes distortion caused by axial motion.

guish between transverse and axial motion artifacts. These names relate to the direction in which the artifact manifests in the acquired data. Transverse motion artifacts are caused by change in fixation of the subject. These changes can be caused by voluntary and involuntary eye motion. A change in fixation leads to a change of the incident angle of the OCT beam onto the eye (see figure 3.3). This change of incident angle causes the beam to be laterally displaced on the retina. Consequently, A-scans will be recorded from this displaced position. Note that such a changed incident angle could also be achieved if there was no motion. Instead, the delta in angle could be reached by a corresponding adjustment of the transverse scan coordinates.

The effect is that the actual scan pattern of the OCT beam on the retina is not a regular sampling. Depending on the concrete motion profile, the regular scan

pattern in the scanner coordinate system is mapped onto a set of locations on the retina which are not regularly spaced. Instead, certain areas might be sampled twice while others are not sampled at all. In the fundus projection in figure 3.10 (a) this can be seen in the breaks in the vessels that are shown. These breaks are not anatomical but caused by transverse motion artifacts.

The second fundamental type of artifact is caused by motion in the axial direction. Figure 3.10 (b) shows a slice of an OCT volume along the slow scan direction which shows axial motion artifacts. They are caused by the eye and/or retina moving toward or away from the OCT system, i.e. along the axial direction. Axial motion can be caused by changes in blood pressure caused by the heartbeat and by respiration. Compared to saccadic transverse motion, axial motion is slow and low frequency. The effect of axial motion is a translation of the content of the acquired A-scans along the axial dimension due to a change in optical path length until the retina is reached. Therefore, the retina will move up and down in a B-scan image. The green profile line in figure 3.10 (b) symbolizes this time dependent effect.

While the views from figure 3.10 (a) and (b) show clear signs of motion artifacts, the slice view from figure 3.10 (c) does not show obvious motion artifacts. This is because the effect of motion on the data set depends on the scanning process. The slice in question is along the *fast scan direction*. The A-scans along this direction are rapidly acquired in sequence. Motion dependent effects are effectively "frozen out" due to the short acquisition time of a single B-scan compared to the time it takes to acquire the whole volume and the speed of the motion. Therefore, in raster scan type patterns, distortions and/or breaks due to motion will predominantly be visible when looking along the slow scan direction where the time difference between neighboring A-scans is much larger.

Figure 3.11 shows the relationship between object and scanner coordinate system when affected by motion in the transverse plane. Shown on the left is an en face view in the scanner coordinate system. Dotted arrows indicate B-scans, dots indicate individual A-scans. In each B-scan, individual dot pairs have been made black to show the correspondence in the two spaces. The background shows an en face fundus projection as it would be acquired given motion. The volume data is defined in the scanner coordinate system, therefore this view corresponds to what a projection of a recorded volume would look like. The two red arrows indicate discontinuities from motion. The right side shows an en face view in the corresponding object coordinate system. Arrows colored with the same color as left indicate where B-scans from the scanner coordinate system are located in the object coordinate system. The background shows an en face view of the object in the object coordinate system. Individual black dots on each B-scan indicate corresponding A-scans in the two coordinate systems. Note that in the object coordinate view, there are no motion artifacts. However, due to motion, the regular raster pattern from the scanner coordinate system is mapped to an irregular pattern in object coordinates. Therefore, the scanner coordinate representation is not a true representation of the object.

Another key concept when describing motion artifacts is the notion of in-plane versus out-of-plane motion. The plane here references the plane of which a B-scan acquires image data. So in a scan pattern with fast scanning in x direction

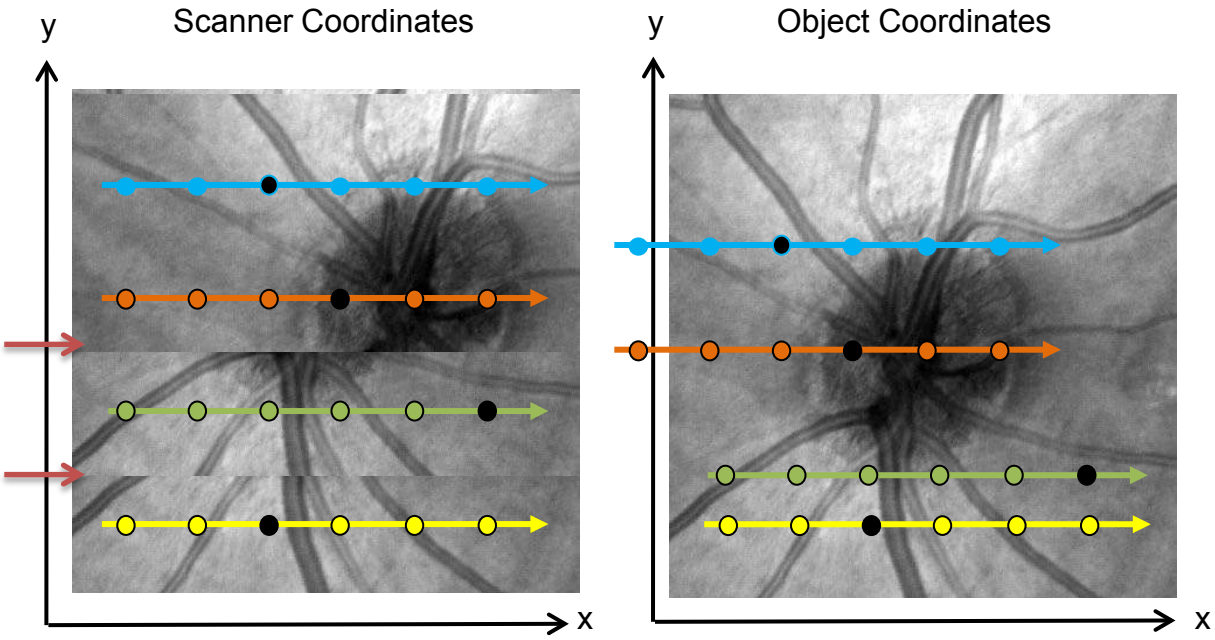

Figure 3.11: Relation between scanner and object coordinates under the presence of motion.

the B-scan plane would be the x/z plane. Correspondingly the y/z plane for a pattern with fast scanning in y direction. In-plane motion artifacts lead to shifts of the image content that pertain to this plane. Therefore, an axial shift, caused by axial motion is always considered in-plane as it leads to a shift along the z axis. For an XFAST type scan pattern transverse motion that leads to a shift along the x axis would also be considered in-plane, and analogously for a YFAST pattern. In-plane motion has the convenient effect that the data can be compensated while only considering the current B-scan. On the other hand, if there is out-of-plane motion, the correct data for a certain location might be in another B-scan or might not have been scanned at all. Some correction approaches simplify the modeling of the problem by only considering in-plane motion (see chapter 4).

As an important consequence of these effects, motion artifacts cause problems in the reliability of 3D-OCT data for clinical purposes. They cause an uncertainty in which object location was imaged by a certain A-scan. Also, distortions caused by motion artifacts do not preserve distances and angles in the volume. Measuring the volume of a lesion in the retina or the average thickness of a retinal layer at predefined positions are two examples of measurements of which the reliability can be severely impacted by motion artifacts.

3.5 Summary

In this chapter, several aspects of OCT imaging in ophthalmologic practice were described. First, the basic anatomy of the eye was described. The eye is transparent and also has optics to focus light onto the retina which has a layered structure. In order to prevent a phenomenon called neural adaptation, the eye also has to keep

moving and change the fixation in order to see. When scanning the retina of the eye with OCT, a collimated beam is focused by the eye optics itself onto the retina. By varying the incident angle of this beam to the eye, different regions of the retina can be imaged. This way both 2D and 3D imaging can be performed. For 3D imaging specifically, raster scanning is most common. Many eye diseases manifest in changes in the layered structure of the retina. As OCT can image the layered structure non invasively and with high quality, OCT is well suited for diagnostics and disease progression tracking provided that quantification is reliable.

Several factors influence this ability. These include speckle noise, blinking, illumination effects and floaters, tilt and finally motion artifacts. Motion artifacts result from relative motion between the subject and the OCT device during acquisition. They cause a distortion of the acquired data as the areas that are imaged are deviating from where the expected location would be. Furthermore it can be distinguished between transverse and axial motion artifacts. How motion artifacts manifest in an OCT volume depends on the scan pattern that is used. In a raster scan for example, the fast scan direction will be relatively undistored, as opposed to the slow one. Motion causes the relationship between object and scanner coordinate to change over time. Therefore, a regular grid in the scanner coordinate system might be mapped to a irregular one on the object. This causes an uncertainty in where on the retina a certain A-scan was taken and presents a problem for accurate quantitative measurements using OCT.

CHAPTER 4

Prior Work

The following chapter is concerned with prior work on motion artifact correction and avoidance in OCT and methods to improve the signal quality in OCT images. Parts of this chapter have already appeared as part of a review of the state of the art as part of a book chapter on the topic by the author [Drex 15].

4.1 Motion Correction

Since motion artifacts constitute a serious issue especially for retinal OCT imaging, considerable work has been performed by different groups to overcome the problem. In the following sections, we give an overview of the state of the art in OCT motion correction techniques. One basic feature of a particular motion correction technique is whether it needs additional hardware support, i.e. the OCT system needs to be built with the motion correction technique in mind or additional imaging modalities need to be available. There are two basic ways to address the problem. Hardware-based methods try to avoid motion artifacts during the acquisition itself though a specific system design:

- Freeze out motion by improving the encoding of spatial dimensions in time, i.e. acquire the data set in a shorter time [Leit 03a, Hube 06, Klei 11, Klei 12, Wies 10, Boni 10].
- Measure the deviations that originate from changes in relative position and actively apply corrections to the galvanometer mirror positions during acquisition: tracking OCT [Pirc 07, Ferg 04, Hamm 05, Magu 07],

Software-based methods on the other hand try to correct motion artifacts retrospectively using image processing, specifically image registration:

- Use images from another modality that does not suffer from motion artifacts as OCT does, as a reference to correct the OCT data [Capp 11, Ricc 09].
- Correlate consecutively acquired data to filter out the effects of motion [Swan 93, Zawa 07, Anto 11].

- Correct motion artifacts using additional OCT data with orthogonal fast scan axis [Zawa 07, Pots 08, Hend 13].

In principle, these basic approaches can also be combined in a concrete method. In the following sections, we review selected state of the art methods for each approach.

4.1.1 Acquisition Speed

One fundamental way to alleviate the motion artifact problem in OCT is to increase the imaging speed of the OCT system, which continues to be important. Higher speed means that a higher number of A-scans can be sampled per unit of time and that therefore a certain scan pattern can be sampled in less time. Since motion requires time to pass, short enough acquisition times can effectively be used to ”freeze out” motion in parts or in even in the whole OCT acquisition and minimize motion induced spatial distortion.

Historically, the move from TD-OCT to FD-OCT enabled an order of magnitude increase in acquisition speeds. This was enabled by the inherent sensitivity advantage of FD-OCT [Leit 03a]. In addition, the reference arm did not need to be scanned anymore during the acquisition of a single A-scan. It is a reasonable assumption in time domain OCT, that there is effectively no motion within an A-scan. The higher speed of FD-OCT systems allows current commercial systems to effectively disregard motion within a single 2D B-scan, simply because the acquisition time is short enough compared to the speed and frequency of eye motion.

Within the realm of existing FD-OCT technology it has been shown that system speed can be improved tremendously with respect to standard commercial systems which operate at around 25 kHz A-scan rate. Using Fourier domain mode locked (FDML) swept source lasers [Hube 06] retinal OCT operating at up to 6.7 MHz has been shown [Klei 11, Klei 12, Wies 10], albeit with reduced sensitivity and resolution compared to commercially available systems.

There is an inherent sensitivity loss associated with running faster due to the maximally allowed light exposure on the eye, which is limited by safety standards [ANSI 07]. This puts an upper bound to the number of photons that can be collected per unit of time. This means that if one runs twice as fast there are only half as many photons available to be collected per A-scan. All other things being equal, this means that one pays an increase in speed with a loss in sensitivity. Especially for clinical applications, where subjects might have bad eye optics, opacities and floaters, a system with sufficient sensitivity headroom is necessary for imaging.

Another issue is that one might want to use the high speed of a system not just to lower the overall acquisition time and motion artifacts. Instead one might choose to acquire more A-scans in total, e.g. to sample more densely and/or to sample a larger area. This trade-off depends on the concrete data that one wants to collect.

Pending significant improvements in sensitivity, speed alone is unlikely to be the only solution for avoiding motion artifacts in OCT, at least as long as dense sampling of a clinically relevant area with good sensitivity and resolution is required. Such improvements might come from entirely alternative forms of OCT

such as full field OCT which has already been demonstrated for retinal imaging [Boni 10]. This technique illuminates and collects data from the full field at once and does not require the scanning of the OCT beam. This helps in achieving high speeds and allows for a higher light exposure. However, as of now, low sensitivity and axial resolution as well as issues with cross talk and uniform image quality limit the practicality of the technique.

4.1.2 Tracking OCT

The approach of tracking OCT is to continuously measure the motion induced deviation from a reference position and to apply a corresponding offset to the galvanometer mirrors and/or the reference arm mirror to cancel this deviation. This effectively compensates for the deviation in scan position that is caused by object motion and therefore removes motion artifacts. Key factors in tracking OCT are the accuracy of measurement of the deviation and its correction and the update rate of the system, i.e. how fast the system can react to a motion induced change in relative position.

Pircher et al. [Pirc 07] employed axial tracking of the eye motion in the context of retinal imaging using time domain en face OCT. This modality acquires one en face plane of information at a time using rapid scanning. Therefore, it is very sensitive to axial motion even in the order of the axial resolution of the system. Axial deviation due to motion was measured by using a second Fourier domain channel at 1300 nm that was used to continuously track the position of the cornea. The measured deviations in cornea position were used to generate a correction signal for a voice coil in the reference arm to rapidly change the reference arm length. This system achieved an update rate of 200 Hz.

Although there is some work on axial tracking, more commonly tracking in OCT is used to correct for transverse motion. Among others, Ferguson et al. used a secondary sensing beam that rapidly scans a circular area on the fundus, for example around the optic nerve head [Ferg 04, Hamm 05, Magu 07]. The system extracts correction information from this secondary channel in a closed loop running at 1 kHz and applies the correction to the galvanometer mirrors. The reported accuracy of the technique is less than one spot diameter. Transverse tracking is also employed in commercial OCT devices such as the Heidelberg Engineering Spectralis (HRA+OCT,Heidelberg Engineering, Heidelberg, Germany). A Scanning Laser Ophthalmoscope (SLO) [Webb 87] is used to rapidly acquire 2D images of the fundus. SLO is a scanning imaging modality that is similar to OCT in this respect. However, SLO typically scans much faster than OCT and can more effectively freeze out object motion. These images are then registered to a reference SLO view. The shift between the two images corresponds to the deviation in scan angle. A correction signal is then applied to the galvanometer mirrors to compensate for this deviation. The use of this technique allows the system to acquire multiple 2D B-scans at roughly the same location and average them in order to remove speckle noise and increase SNR.

4.1.3 Use of Additional Modalities

This class of motion correction employs reference image data from a different modality that does not suffer from motion artifacts as OCT does. By registering the OCT data to the reference modality image, one can find corresponding locations between the two images. The OCT data can then be mapped onto the reference image. Since the reference image contains virtually no motion, the OCT data can be motion corrected if the mapping between the two images is accurate. Two modalities that are used for retinal imaging and which are not suffering from OCT like motion artifacts are fundus camera photography and SLO imaging. Motion in fundus photography will lead to a blurring effect. However, typical exposures are short enough to prevent this problem. Capps et al. [Capp 11] used an adaptive optics SLO that is running simultaneously with the OCT acquisition to estimate and correct for lateral motion. After imaging, the OCT data is registered to the SLO data in order to calculate the displacement caused by motion per A-scan. Subsequently, the OCT data is re-sampled onto a regular grid. Ricco et al. [Ricc 09] registered the OCT fundus view to an SLO reference image in order to correct for motion. The algorithm uses the vessel pattern visible in both modalities as features for the registration. After vessel detection, registration is performed in a two-step process: First, drift and tremor is corrected by using an elastic registration technique that is based on patch-wise affine transformations between the two images. Over all pixels (x, y), the sum of the terms

$$(m_7 I^{SLO}(x, y) + m_8 - I^{OCT}(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6))^2 \tag{4.1}$$

is minimized, where $I^{SLO}(x, y)$ is the SLO image and $I^{OCT}(x, y)$ is the OCT fundus image and $\mathbf{m} = (m_1, \cdots, m_8)$ is the parameter vector. m_1 to m_6 model a deformation of the OCT fundus image while m_7 and m_8 model a linear relationship between the intensities of both modalities. The same parameters $\mathbf{m}$ are shared over one patch. The second step attempts to correct discontinuities caused by microsaccades along the fast scan direction. The OCT en face pixels are treated as a time domain signal and the best alignment with the reference image is found using dynamic time warping. Inherent to this kind of technique is the need for having images from two modalities which poses a logistical problem.

4.1.4 Consecutive Data Correlation

The underlying idea of consecutive data correlation algorithms is to assume that the imaged object is inherently smooth and densely sampled by the OCT scan pattern. Therefore, high frequency spatial patterns in the data, i.e. jumps between consecutively acquired A-scans or B-scans are induced by motion only. By correlating consecutive data and shifting it such that the smoothness of the result is maximized, the high frequency artifacts that are induced by motion can be removed. Swanson et al. used 1D cross correlation between neighboring time domain OCT A-scans within a 2D intensity B-scan $I(x, z)$ to remove motion artifacts

[Swan 93]. The axial shift $\delta z(x)$ between consecutive A-scans that maximizes the cross correlation

$$\sum_z I(x,z)I(x+1,z+\delta z(x)) \tag{4.2}$$

is calculated for every A-scan. An absolute motion profile is then calculated by accumulating the relative shifts.

This 1D motion profile is subsequently filtered based on prior knowledge on the frequency distribution of axial motion. Specifically, the motion profile needs to be high-pass filtered; otherwise, the low frequency curvature of the retina is removed. Finally, the filtered motion profile was applied to the each A-scan to remove axial motion.

In 3D volumetric imaging using raster scans, correlation or registration can similarly be used to estimate the motion induced shift between consecutively scanned neighboring B-scans within the volume. Instead of 1D A-scans, 2D B-scans need to be correlated/registered with each other here. Once consecutive B-scans have been registered, the shifts can again be filtered in order to preserve low frequency curvature of the scanned object. The underlying motion model assumes that motion only occurs in between B-scans, i.e. that B-scans themselves are rigid. Furthermore, correlation of consecutive B-scans can only correct for in-plane motion, which is motion that causes a shift of the image content in axial and/or in the direction of the fast scan direction of the raster scan. In reality however, transverse motion such as that caused by saccades can also take place in direction of the slow raster scan direction. In this case, techniques that are based on subsequent B-scan correlation produce inadequate results. For example, Zawadzki et al. used consecutive B-scan registration [Zawa 07]. Antony et al. corrected for axial motion artifacts in 3D raster scans using an approach based on layer segmentation and fitting of a thin plate spline surface to said segmentation followed by multiple steps of smoothing [Anto 11].

4.1.5 Orthogonal Scanning Based

The final class of motion correction algorithms applies to 3D volumetric imaging and employs orthogonally scanned data. This means that one or more B-scans are acquired with a fast scan axis that is orthogonal to the B-scan direction of the 3D raster scan, which is to be corrected. In the extreme case, two or more full raster scans with orthogonal fast scan axis are acquired and all of them are corrected.

One idea of using orthogonal scans is to acquire a few orthogonal B-scans in addition to a raster scanned volume and use the orthogonal B-scans as ”guideposts” to which the raster scanned volume is registered. It is assumed that no motion takes place during the acquisition of the guidepost scans. Within the context of these algorithms, they function as a motion free reference. When the raster scanned volume can be accurately registered to the guidepost scans it can be roughly motion corrected.

After consecutive B-scan registration to remove axial and in plane transverse motion Zawadzki et al. used a single orthogonal guidepost B-scan in the center of the volume (x coordinate) to remove the flattening artifact that results from unfiltered correlation of B-scans [Zawa 07]. For each B-scan along the slow scan axis

with coordinate an A-scan of the guidepost scan is associated with one A-scan of the raster-scanned volume. These matched A-scans are again aligned by maximizing the correlation in dependence of an axial shift. The found shift is then applied to the corresponding B-scan. This way reference information from the guidepost scan is used to correct the flattening artifact and residual axial motion within the raster-scanned volume. Potsaid et al. extended this concept and used three instead of one guidepost scans [Pots 08]. This allows for increased robustness as instead of one pair of A-scans three pairs are correlated to find the axial motion profile. However, these methods have the limitation of a non realistic motion model because motion is usually corrected on a per B-Scan level.

A consequent extension of the guidepost concept is to acquire an additional orthogonal whole raster scanned volume. The concept is that the fast direction of one scan can be used to correct the slow direction of the other and vice versa. The dense data that is available allows for unique opportunities to correct motion in all three directions, including out of plane motion.

Tolliver et al. used two orthogonal raster scans and an approach based on matching A-scans from the two volumes to each other to estimate and recover motion in all three dimensions [Toll 09]. In a first step, each A-scan is transformed into a feature vector using a shift invariant 1D Haar wavelet transform. Then a classifier with the goal of assessing the probability whether two A-scans are similar, i.e. they were sampled from close locations on the retina is trained. Here, A-scans that are on the same B-scan and spatially close to a certain A-scan are assumed to be similar for the purpose of training the classifier. The classifier is used to compute "pseudo-match probabilities" between A-scans from both volumes. In a subsequent step, Bayesian smoothing is used to incorporate the prior knowledge of piece-wise smooth eye motion. Instead of choosing the most likely matches given the classifier output, a less likely but piece-wise smooth set of matches is favored. In addition, axial motion correction is performed. Unfortunately, this approach was never formally published which leaves several details and the results of this approach unknown.

Hendargo et al. developed an OCT motion correction and volume stitching algorithm for the specific context of speckle variance angiography [Hend 13]. They utilized OCT layer segmentation and orthogonal scan patterns. In speckle variance OCT, motion artifacts can be detected by the highly increased variance of the corresponding set of B-scans. Hendargo et al. used this fact to subdivide segmentation based 2D projections of the angiography volume into several strips. The angiography information was subsequently enhanced using Gabor filtering. The strips were then globally aligned with one 2D translation per strip by maximizing correlation between strips. Subsequently, a spline based registration step was performed to compensate for additional deformation between the globally aligned patches. Finally, composite images were generated. In addition, the technique is able to mosaic 2D vasculature images from multiple locations on the retina to create a wide field 2D angiography image. One limitation of this approach is that it requires a specific OCT scan pattern which allows for angiographic information to be acquired. This limits the generality of the method.

4.2 Signal Improvement Methods

Signal improvement methods in OCT are mainly concerned with increasing SNR and removing the effects of speckle noise (see section 3.4.1) from the images. Both goals help to ease subsequent processing steps, such as automatic segmentation, which are necessary for quantitative analysis of OCT data.

4.2.1 Physics-based Methods

One physical method for OCT system design to reduce speckle noise is to employ polarization diversity, that is using unpolarized light in both sample and reference arm [Schm 99b]. However, the technique is limited as it increases SNR by a factor of $\sqrt{2}$ at best [Schm 99b].

Spatial compounding and more specifically angular compounding is a method of simultaneously acquiring multiple A-scans that illuminate a certain location with slightly different incident angles [Schm 99b, Schm 97, Bash 00]. The difference in incident angle is sufficient to de-correlate the speckle noise pattern that is visible. The actual signal however, will remain correlated. The A-scans can then be combined by averaging intensity in log-space. For N A-scans, the SNR gain is at most $\sqrt{N}$ using this method [Schm 99b]. Another method of spatial compounding involves acquiring multiple A-scans from slightly laterally displaced locations on the sample and combining these [Szku 12].

The speckle noise pattern that is observed also depends on the frequency and bandwidth of the light source [Schm 99b]. This is exploited in frequency compounding by splitting the spectrum of the light source into N sub bands [Schm 99b]. From each band an A-scan can be computed. Similarly to angular compounding, these A-scans can then be averaged to a single one, yielding an SNR gain of at most $\sqrt{N}$. However, due to the reduced bandwidth available for each band, the axial resolution degrades by a factor of N [Schm 99b].

4.2.2 Post-Processing

Due to the effects of motion on the alignment of the OCT system with the eye, the incident angle on the retina changes automatically over time and with that the speckle noise pattern. This can be exploited by scanning the same area multiple times and combining the acquired images. However, due to motion, not only the incident angle changes but also the position on the retina from which the A-scans are acquired might change. Therefore, tracking (see section 4.1.2) or registration of the images and removal of scans from the wrong location are needed prior to combining the scans. In practice, a linear B-scan is repeatedly scanned, yielding a single high quality 2D image as a result [Papp 12]. Basically all commercial OCT systems support this mode. The improved visualization of features and data quality have been found to be clinically useful [Saka 08].

More advanced signal processing based approaches can be used to removing unwanted speckle noise. These methods can be distinguished based on whether

they operate on a single or on multiple B-scans at the same time. For example, Wavelet analysis and thresholding methods have been used both for single [Adle 04, Xian 98] as well as multi-frame denoising [Maye 12, Chit 12].

4.3 Summary

In this chapter the state of the art in OCT motion correction and signal improvement has been considered. Several approaches exist for motion correction, one of which is to increase the acquisition speed of the system. This can help to effectively ”freeze” out the motion, but since the maximum allowed light exposure is limited by safety standards, it comes at a sensitivity cost. A second major direction is the use of tracking hardware to monitor and correct where the beam is pointing on the retina. Unfortunately, this leads to an increase in cost and complexity of the OCT system. The acquired OCT images can also be related to an image from another modality which does not suffer from motion artifacts. Consecutive data correlation operates under the assumption that high spatial frequencies for example between neighboring B-scans are caused by motion. By filtering out such frequencies, the motion artifacts can be removed. This technique is very simple but often the underlying assumption is violated which can lead to inaccurate results. The final class of motion correction algorithm is based on using orthogonally scanned data. This ranges from a single or a few orthogonal ”guidepost” B-scans to acquiring whole volumes that are acquired using an orthogonal raster scans.

The prior work in signal improvement methods can be divided into methods that are based on a modified physical setup or post-processing based methods. Physical methods use the fact that speckle noise in the signal is not correlated over different incident angles, polarization or frequency. The system acquires multiple samples of the same location such that speckle is uncorrelated. Then speckle can be reduced by combination of the signals. Post-processing approaches use the fact that speckle decorrelates easily when imaging the same area multiple times. Then the multiple samples can also be combined and speckle noise can be reduced. Another principal direction is to employ special denoising methods based on digital signal and image processing. These methods either operate on single images or multiple images from the same location are combined.

Part II

3D-OCT Motion Correction using Image Registration and Orthogonal Scanning

CHAPTER 5

Motion Correction Approach

The following chapter constitutes the main method part of this work. The 3D-OCT motion correction and signal enhancement algorithm utilizing image registration and orthogonal scan patterns will be deduced and described. Parts of this chapter have been published in prior publications of the author [Krau 12, Krau 14].

5.1 Definitions

As input for the algorithm $N_v \geq 2$ OCT volumes are available which are scanned with two orthogonal scan patterns denoted as **XFAST** and **YFAST** in such a way that there is at least one volume of each type available. Furthermore, the volumes are supposed to be acquired using the same OCT system and contain data (excluding motion effects) of approximately the same area on the retina. The two scan patterns traverse the same grid of A-scan sampling locations in the scanner coordinate system.

To simplify the explanation, the following sections will consider the case of $N_v = 2$, i.e. that there are only two input volumes. This restriction will be lifted in section 5.5.11. The two volumes are defined on a common, regular 3D grid in the scanner coordinate system. They are denoted $\mathbf{X}_{i,j,k}$ and $\mathbf{Y}_{i,j,k}$ where $i = 1, ..., w, \quad j = 1, ..., h$ and $k = 1, ..., d$ are the indices in the two transverse (x,y) and axial directions (z), respectively. Each volume therefore consists of $w \times h$ A-scans with d axial pixels each.

For each volume, $\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$ associate a time t with each A-scan grid point (i,j) of the respective volume at which the A-scan was recorded. We can describe the content of the volumes using a function

$$A(x,y,z,t) : \mathbb{R}^4 \mapsto \mathbb{R}^d, \tag{5.1}$$

where x ,y ,z are three coordinates in the scanner coordinate system and t is the time of acquisition. A returns a d-dimensional vector of (log-scaled) intensities that make up the content of the A-scan. A is dependent on x and y scanner coordinates, an offset in z direction that can be used to translate the content of a given A-scan and a time t. $A(x,y,z,t)$ is time dependent because the A-scan content depends on the alignment between the OCT system and the eye. Due to relative motion, this alignment can change over time.

Based on these definitions, the volumes $\mathbf{X}$ and $\mathbf{Y}$ are defined as

$$\mathbf{X}_{i,j,k} = A(x_i, y_j, 0, \mathbf{T}^{\mathbf{X}}_{i,j})^T \cdot \mathbf{u}_k \quad \forall i,j,k \tag{5.2}$$

and

$$\mathbf{Y}_{i,j,k} = A(x_i, y_j, 0, \mathbf{T}^{\mathbf{Y}}_{i,j})^T \cdot \mathbf{u}_k \quad \forall i,j,k, \tag{5.3}$$

where x_i and y_j are the scanner coordinate system coordinates in x and y direction, respectively, that are associated with the i-th and j-th grid point. $\mathbf{u}_k$ is a d-dimensional unit vector that is all zero, except for 1 at the k-th component. It is used here in the dot-product to select the k-th component of the vector that is output from the A function.

5.2 Deduction

The first key step in deducing the correction algorithm is to associate the output of the A function at an arbitrary time t to a fixed point in time, here arbitrarily chosen as $t = 0$. The A-scan contents at one time can be related to another time via a time dependent three dimensional offset such that

$$A(x,y,z,t) = A(x - D_x(t), y - D_y(t), z - D_z(t), 0) + \mathbf{e}, \tag{5.4}$$

where $D_x(t) : \mathbb{R} \mapsto \mathbb{R}$, $D_y(t) : \mathbb{R} \mapsto \mathbb{R}$ and $D_z(t) : \mathbb{R} \mapsto \mathbb{R}$ are time dependent coordinate offset functions. $\mathbf{e}$ is a d-dimensional vector that models noise etc. (see section 3.4.1). This models the basic effects of motion artifacts that are induced by a changing alignment between the system and the eye (see section 3.4.6). Other effects such as those caused by blinking, illumination, tilt and vignetting (see section 3.4.2,section 3.4.3, section 3.4.5 and section 3.4.4) are not modeled here.

As a second step, we introduce the concept of the volume data not only being defined at grid point locations. For this purpose we define an *interpolation function* $I(\mathbf{V}, x, y, z) : \mathbb{R}^{w \cdot h \cdot d + 3} \mapsto \mathbb{R}$ such that

$$I(\mathbf{V}, x_i, y_j, z_k) = \mathbf{V}_{i,j,k}, \tag{5.5}$$

where z_k is an axial coordinate corresponding to the k-th axial grid position. The function I interpolates the volume data $\mathbf{V}$ at scanner coordinates corresponding

to grid points. The underlying grid is a *fixed regular* 3D grid in the scanner coordinate system corresponding to the grid of A-scans that was scanned in the respective raster scan. In between grid point coordinates the function is supposed to approximate an A-scan that might have been acquired at corresponding scanner coordinates. Given sufficiently dense sampling and ignoring severe motion artifacts for now this seems reasonable.

Using the equations introduced so far we can write

$$I(\mathbf{X}, x_i, y_j, z_k) = A(x_i, y_j, 0, \mathbf{T}^{\mathbf{X}}_{i,j})^T \cdot \mathbf{u}_k \tag{5.6}$$

and

$$I(\mathbf{Y}, x_i, y_j, z_k) = A(x_i, y_j, 0, \mathbf{T}^{\mathbf{Y}}_{i,j})^T \cdot \mathbf{u}_k. \tag{5.7}$$

In a second step the reduction to $t = 0$ is performed such that

$$I(\mathbf{X}, x_i, y_j, z_k) = \left(A(x_i - D_x(\mathbf{T}^{\mathbf{X}}_{i,j}), y_j - D_y(\mathbf{T}^{\mathbf{X}}_{i,j}), -D_z(\mathbf{T}^{\mathbf{X}}_{i,j}), 0) + \mathbf{e}_1\right)^T \cdot \mathbf{u}_k \tag{5.8}$$

and

$$I(\mathbf{Y}, x_i, y_j, z_k) = \left(A(x_i - D_x(\mathbf{T}^{\mathbf{Y}}_{i,j}), y_j - D_y(\mathbf{T}^{\mathbf{Y}}_{i,j}), -D_z(\mathbf{T}^{\mathbf{Y}}_{i,j}), 0) + \mathbf{e}_2\right)^T \cdot \mathbf{u}_k, \tag{5.9}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are two error vectors.

Finally, the displacements can be moved to the other sides of the equations such that

$$I\left(\mathbf{X}, x_i + D_x(\mathbf{T}^{\mathbf{X}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{X}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{X}}_{i,j})\right) = (A(x_i, y_j, 0, 0) + \mathbf{e}_1)^T \cdot \mathbf{u}_k \tag{5.10}$$

and

$$I\left(\mathbf{Y}, x_i + D_x(\mathbf{T}^{\mathbf{Y}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{Y}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{Y}}_{i,j})\right) = (A(x_i, y_j, 0, 0) + \mathbf{e}_2)^T \cdot \mathbf{u}_k. \tag{5.11}$$

Note that except for the error terms, the right sides of these two equations are identical. The right sides also have A function invocations with a constant time argument of $t = 0$. In addition, the transverse coordinates form a regular grid consistent with the initial scanner coordinate system grid. At a single point in time the alignment between eye and OCT system is static. Therefore, the regular scanner coordinate grid maps to a regular grid of sampling locations on the retina. This is exactly what is desired.

Unfortunately, the values of the offsets $D_x(t)$, $D_y(t)$ and $D_z(t)$ that fulfill these equations are *unknown*. However, the equations show that given volumes where every required retinal location was sampled at least once and given the right offsets for each grid point and volume, two volumes that are the same regular, motion-free sampling of the retina can be constructed by offset interpolation of the original volume data.

In order to estimate the unknown displacements, we use the fact that given the right offsets the two re-sampled volumes need to be *similar*. For the two volumes

$\mathbf{X}$ and $\mathbf{Y}$, and a certain set of offsets $D_x(t)$, $D_y(t)$ and $D_z(t)$, we define the *residual volume* $\mathbf{R}(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z)$ at every grid point as

$$\begin{aligned}\mathbf{R}_{i,j,k}(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z) =& I\left(\mathbf{X}, x_i + D_x(\mathbf{T}^{\mathbf{X}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{X}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{X}}_{i,j})\right) - \\ & I\left(\mathbf{Y}, x_i + D_x(\mathbf{T}^{\mathbf{Y}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{Y}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{Y}}_{i,j})\right), \\ & \forall i, j, k.\end{aligned} \tag{5.12}$$

Given the right $D_x(t)$, $D_y(t)$ and $D_z(t)$ at the set of times where A-scans were sampled in each volume ($\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$) the magnitude of this residual should be minimal. By applying a loss function $L(r) : \mathbb{R} \mapsto \mathbb{R}^+_0$ to the corresponding residual and accumulating these values over all i,j and k we define a *similarity measure* between the volumes after they have been transformed by applying the displacement values as

$$S(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z) = \sum_i \sum_j \sum_k L\left(\mathbf{R}_{i,j,k}(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z)\right). \tag{5.13}$$

The value of this measure is low if the interpolated volume data is similar and high if it is not. Based on maximizing the similarity the objective becomes minimizing the accumulated loss

$$D_x, D_y, D_z = \underset{D_x, D_y, D_z}{\operatorname{argmin}} S(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z). \tag{5.14}$$

This can be interpreted as a special kind of *image registration problem* [Zito 03]. For each A-scan of every volume, a 3D-displacement vector is needed as an offset to sample the original volumes. The displacements that are associated with every A-scan position can be interpreted as one 2D *displacement field* per volume. At each A-scan position the displacement field contains a 3D displacement vector. These are to be found for each volume in such a way as to maximize the similarity of the two volumes in the transformed state. As opposed to a normal registration problem, there is no reference volume. This is because both volumes are affected by motion. Therefore, both volumes have to be transformed.

As mentioned before, the sought for displacement functions can also be seen as displacement fields. For a volume $\mathbf{V}$, the displacement field is denoted by $\mathbf{D}^{\mathbf{V}}_{o,i,j}$, where $o = 1, 2, 3$ denotes the dimension of the displacement expressed. i and j mark transverse indices as before. The value of the displacement field entries is defined as

$$\begin{aligned}\mathbf{D}^{\mathbf{V}}_{1,i,j} =& D_x(\mathbf{T}^{\mathbf{V}}_{i,j}) \\ \mathbf{D}^{\mathbf{V}}_{2,i,j} =& D_y(\mathbf{T}^{\mathbf{V}}_{i,j}) \\ \mathbf{D}^{\mathbf{V}}_{3,i,j} =& D_z(\mathbf{T}^{\mathbf{V}}_{i,j}).\end{aligned} \tag{5.15}$$

Therefore, for each volume the displacement field to be found can be seen as a 2D image with three channels.

Equation (5.14) does not specify that displacements at subsequent points in time are highly dependent on each other. This dependence stems from the fact

that the change in displacement over time is related to the change in alignment between retina and OCT system which is caused by motion. It can be assumed that the change in alignment is somewhat proportional to the motion that occurs. The amount of motion itself can also be considered proportional to time, i.e. in a very small amount of time there will likely be a smaller amount of motion.

These problem specific assumptions can be incorporated into the approach by use of a *regularization term*. This energy term E is defined as

$$E(D_x, D_y, D_z) = \sum_{t_l} L\left(\frac{\delta D_x(t_l)}{\delta t_l}\right) + L\left(\frac{\delta D_y(t_l)}{\delta t_l}\right) + L\left(\frac{\delta D_z(t_l)}{\delta t_l}\right), \tag{5.16}$$

where $L(x)$ is a loss function as before. This means taking the derivative of the displacement functions with respect to time, applying a loss function and accumulating over all times. The ordered set of all points in time at which A-scans were sampled in the input volumes t_l, where $l = 1, ..., N_v \times w \times h$ serves as a convenience for all the times contained in $\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$. The term expresses the notion that a solution is better if it models less time dependent change in the displacements.

Incorporating equation (5.16) into equation (5.14) yields

$$D_x, D_y, D_z = \underset{D_x, D_y, D_z}{\operatorname{argmin}}(S(D_x, D_y, D_z) + \alpha E(D_x, D_y, D_z)), \tag{5.17}$$

where α is a weighting factor that specifies the relative importance of the data term and the regularizer.

This concludes the deduction of the basic optimization problem that is used for motion correction in this work. In the following sections, the components necessary to make the method usable in practice will be described.

5.3 Processing-Pipeline Overview

Figure 5.1 shows an overview of the processing pipeline for motion correction and merging which will subsequently be described in detail. Starting from with a set of input volumes, several pre-processing steps are performed in order to alleviate inconsistencies in the data, reduce noise and speed up execution of the pipeline (section 5.4). Subsequently, combined registration and motion correction is performed (section 5.5). The registration, i.e. the optimization of the objective function, produces one displacement field for each input volume. By applying said displacement fields to the input volumes registered and motion corrected volumes are constructed (section 5.6). As these registered volumes exist in the same space, the data can be combined (section 5.7) in order to increase SNR and minimize holes in the data. The result of this process is a single merged and motion corrected volume. Optionally, additional functional channels from the input volumes beyond intensity (such as Doppler-OCT, polarization sensitive OCT (PS-OCT) or angiography information) can be mapped to the common space and merged (section 5.8). Finally, in order to improve execution speed, key parts of the pipeline are accelerated using Graphics Processing Unit (GPU) programming (section 5.9).

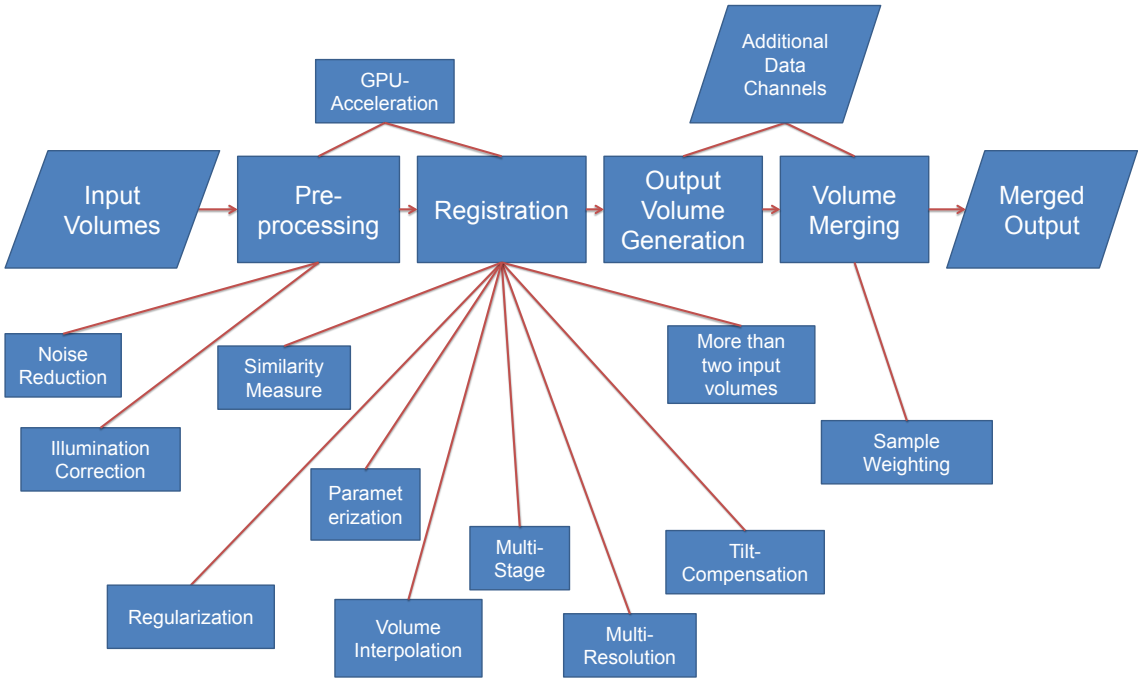

Figure 5.1: Processing pipeline schematic. Arrows indicate the flow of data between processing stages. Lines show associations between the stages and further topics.

5.4 Pre-Processing

Pre-processing consists of a series of optional processing steps that are applied to the input volumes and that seek to modify the data in a way such that subsequent pipeline steps can perform better. Key considerations are robustness of the algorithm and execution speed. In the following sections, the individual steps are described according to the order in which they are performed within the pipeline.

5.4.1 Noise Reduction

To reduce speckle noise (see section 3.4.1) in the input volumes, spatial median filtering is performed on the input data. Both 1D and 2D filtering can be used. 1D filtering operates along the axial direction only. The size of the filter is denoted $s_{med,1d}$. For 2D filtering, a filter of size $s_{med,2d} \times s_{med,2d}$ along the axial and fast scan direction is used. The concrete sizes have to be empirically chosen (see section 6.5) and can be dependent on:

- System SNR
- System axial resolution
- Image axial pixel spacing
- System transverse resolution
- Scan pattern transverse sampling

In order to speed up the process, both types of filters are available as a GPU implementation (see section 5.9).

5.4.2 Thresholding

Depending on the concrete OCT device and the settings used when computing A-scans from the acquired interference spectra, the overall brightness of the OCT images can vary substantially. In order to remove some of this variability, an automatic thresholding step can be performed on the input volume data. From the log-intensities of all the input volumes a histogram with 128 bins is computed. The intensity value associated with the maximum v_{mode} bin value is then used as a lower, the overall maximum intensity v_{max} as an upper threshold. For each volume $\mathbf{V}$ the data is then thresholded and scaled according to

$$\mathbf{V}_{i,j,k} = \begin{cases} \frac{\mathbf{V}_{i,j,k} - v_{mode}}{v_{max} - v_{mode}}, & \text{if } \mathbf{V}_{i,j,k} \geq v_{mode} \\ 0, & \text{otherwise} \end{cases} \tag{5.18}$$

leading to intensity values between zero and one as output of this step. This range is then quantized using 16 bit. In addition, the thresholding assures that the background level v_{bg} is close to zero. The choice for this approach is based on the following assumptions:

- Most of the voxels of a volume are background voxels.
- The background voxels are distributed such most common background intensity if close to the log-mean of all background intensities.
- There is no useful information at values below the mean intensity of the background.

Under these circumstances, v_{mode} will be close the mean of the background pixels. This approach will result in the same lower threshold independently of which background threshold was used in OCT preprocessing, as long as this threshold was below or at the mean of the background. The approach also has the nice property that multiple application of the thresholding will not progressively shrink the value range.

5.4.3 Illumination Correction

Illumination of certain locations as seen by OCT depends on the alignment between OCT system and subject (see section 3.4.3). Between successive volume acquisitions and also within a single acquisition this alignment can change, leading to different illumination of the same anatomical locations in different volumes. In addition, due to the optics of the eye and opacities and floaters different transverse locations are subject to different illumination.

The method is based on comparing interpolated intensities from subsequently acquired volumes within the residual computation (see equation (5.12)) and trying to minimize the difference by finding correct displacement vectors. This relies on

A-scan intensities from a certain anatomical location to be *consistent* over time. Changing illumination when acquiring A-scans of the same anatomical location multiple times can violate the assumption of a (log) additive noise component that is responsible for differences between these A-scans.

For compensating illumination effects, the *en face fundus projection* is used, which for a volume $\mathbf{V}$ (with log-scale intensities) is defined as

$$\mathbf{F}^{\mathbf{V}}_{i,j} = \frac{1}{d} \sum_{k=1}^{d} \mathbf{V}_{i,j,k}. \tag{5.19}$$

Furthermore, it is assumed that illumination effects are a relatively low-frequency, additive effect on said fundus projection.

The illumination component is estimated based on low-pass filtering the *en face* projection with a 2D Gaussian kernel with standard deviation σ_{Illum} which was empirically chosen to be 7.5 transverse pixels.

The resulting 2D image is denoted $\hat{\mathbf{F}}^{\mathbf{V}}$. A *bias field* [Hou 06] $\mathbf{B}_{i,j}$ is computed relative to a reference value v_{ref} such that

$$\mathbf{B}_{i,j} = v_{ref} - \hat{\mathbf{F}}^{\mathbf{V}}_{i,j}. \tag{5.20}$$

As a reference value, the 75th percentile of the set of values in $\hat{\mathbf{F}}^{\mathbf{V}}_{i,j}$ is used. This has the effect of 75 percent of the bias values being positive.

The bias field which has been calculated based on the *en face* fundus projection then has to be applied to the voxels of the corresponding volume $\mathbf{V}$. Uniformly applying a correction value to every pixel of an A-scan would lead to the effect of not only correcting the pixels which show the retina, but also to brighten or darken the pixels belonging to the background. Background pixels however, are not affected by illumination effects.

In order to solve this issue, the voxels of the volume are classified as belonging either to the retina or background. Correction is then only applied to retina voxels. We define a mask volume $\mathbf{M}^{\mathbf{V}}$ for the volume $\mathbf{V}$ as

$$\mathbf{M}^{\mathbf{V}}_{i,j,k} = \begin{cases} 1, & \text{if } \mathbf{V}_{i,j,k} \geq v_{ret} \\ 0, & \text{otherwise,} \end{cases} \tag{5.21}$$

where v_{ret} is an empirically chosen threshold value above which pixels are classified as belonging to the retina.

The bias field is then applied to the volume such that

$$\mathbf{V}_{i,j,k} \mapsto \mathbf{V}_{i,j,k} + \frac{\mathbf{M}^{\mathbf{V}}_{i,j,k} \cdot \mathbf{B}_{i,j}}{\sum_{k=1}^{d} \mathbf{M}^{\mathbf{V}}_{i,j,k}}. \tag{5.22}$$

If the resulting value exceeds the maximum intensity value for the given quantization, clamping is performed.

Figure 5.2 shows a schematic view of illumination correction performed on a 3D-OCT raster scan volume. The input fundus view shows variation in overall brightness which are caused by illumination effects. In the middle, the estimated bias field $\mathbf{B}_{i,j}$ is shown. The right top image shows the fundus projection after correction of the underlying volume data. In the bottom row, cross-sectional views of a slice of the volume corresponding to the blue line in the top views are shown.

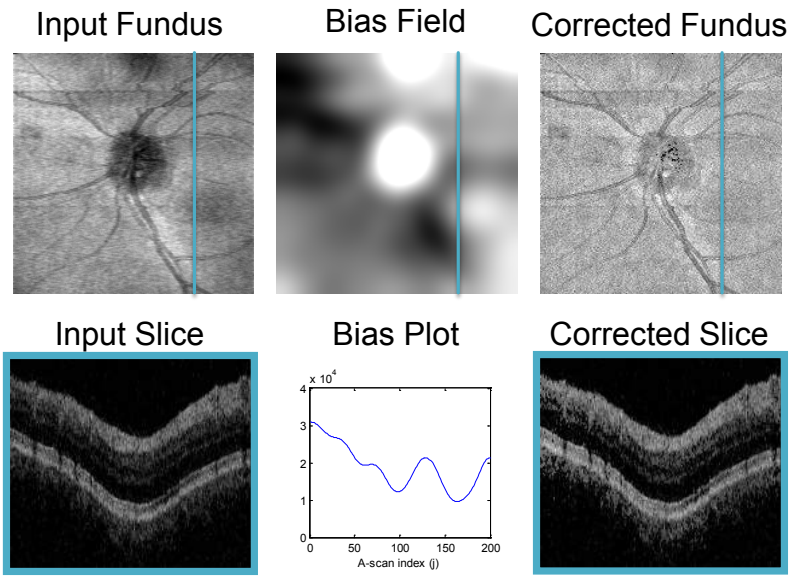

Figure 5.2: Illumination correction schematic

5.4.4 Data Down-sampling

Optionally, the volumes can be down-sampled in axial direction by a factor of two N_{down} times. Down-sampling is performed by first performing Gaussian filtering by applying an empirically chosen 1D (normalized) $(1 \quad 5 \quad 8 \quad 5 \quad 1)^T$ filter to every A-scan and then removing every second axial pixel. It has the effect of further reducing noise and reducing the data size for subsequent steps. However, the axial resolution of the volumes is also reduced by this step, which can limit the theoretically attainable precision of correction in axial direction.

5.4.5 Normalization

Subsequently, an additive bias and a scaling factor which are the same for all input volumes are applied to the intensities. The values are chosen such that the overall value distribution of all voxels of all volumes has zero mean and variance one. This value range normalization helps to standardize the intensities. This is important because parts of the objective function do not depend on the input volume intensities. Normalization tries to help to keep the relative weights of these terms within the objective function independent of the intensity range of the input volumes.

5.5 Registration

After pre-processing, the main registration step is used to estimate a displacement field for each volume. This is performed by optimizing a suitable objective function. The objective function itself consists of a weighted sum of a data similarity and a regularization term which is to be minimized. The specifics regarding the

design considerations and options and how these steps are performed will be described in the following sections.

5.5.1 Similarity Measure

For the similarity measure itself (see equation (5.13)), the main design choice is which concrete *loss function* L to use. One choice is to employ a square loss function $L_2(x) : \mathbb{R} \mapsto \mathbb{R}_0^+$ which is defined as

$$L_2(x) = x^2. \tag{5.23}$$

Together with the subtraction of interpolated intensities as part of the residual computation (see equation (5.12)) this results in the sum of squared differences (SSD) similarity measure. Assuming that intensities of corresponding anatomical locations in the volumes are the same except for additive Gaussian noise, this measure is statistically optimal [Sebe 00].

However, there are some problems with this model. First, speckle noise can not be considered to be of Gaussian distribution (see section 3.4.1). In particular, there is a salt-and-pepper noise component to speckle noise. Within the context of assuming a normal distribution such salt-and-pepper influences can be considered as outliers as they strongly violate the assumption. Secondly, effects such as illumination (see section 3.4.3) can lead to a systematic bias between intensities of corresponding locations. The error between intensities of corresponding locations might therefore not even be zero mean.

Illumination correction (section 5.4.3) and noise reduction (section 5.4.1) attempt to alleviate these issues. However, these effects can still play a role. In this case, the violation of the assumptions implied by using a square loss function can lead to systematic mis-matching of locations. This means that the solution that contains displacement values that minimize the overall squared error between interpolated intensities might not associate corresponding anatomical locations with each other. In addition, because of the quadratic loss, the error contribution from outliers dominates the objective function value. This can cause the optimizer to find parameters that alleviate primarily these outlier error contributions at the expense of overall performance.

In order to try to deal with outliers and inconsistencies in the data, an alternative loss function based on the pseudo Huber loss function [Hube 64] is proposed. The function is related to the L_1 norm which is defined as

$$L_1(x) = |x|. \tag{5.24}$$

Compared to the square loss function L_2, the L_1 loss function does not disproportionally penalize individual high values as they might be caused by high residual values due to data inconsistencies or noise. Therefore, the L_1 norm would lead to a higher robustness with respect with respect to these effects. However, L_1 is not continuously differentiable. Therefore, the whole objective function would not be continuously differentiable. This would conflict with the use of second order gradient based non-linear optimization methods that are employed in this work (see section 5.5.7).

As a compromise between having the robustness of L_1 and still keeping the objective function continuously differentiable, the pseudo Huber loss function [Hube 64] L_{H,ϵ_H} can be used. It is defined as

$$L_{H,\epsilon_H}(x) = \epsilon_H \cdot \left(\sqrt{1 + \left(\frac{x}{\epsilon_H} \right)^2} - 1 \right), \tag{5.25}$$

where ϵ_H is a small, positive constant. The closer ϵ_H is to zero, the closer the function approximates L_1. For absolute values of x much greater than ϵ_H, the slope of the function goes toward one. This implies that large absolute input values are not associated with very high loss values. Therefore, when using L_{H,ϵ_H}, individual high residual values due to inconsistencies in the data will have less influence on the overall objective function value than in the case of using L_2. This leads to potentially higher robustness. When using L_{H,ϵ_H} within the data similarity term, ϵ_H was empirically chosen to be 0.0001.

5.5.2 Regularization

The purpose of the similarity measure is to guide the optimizer towards a solution that allows for the matching of A-scans from corresponding anatomical locations. The regularization term on the other hand, is supposed to assure that the motion profile that is modeled by the displacement functions is realistic. In principle, this can be achieved by assigning a high (penalty) value to the regularization term for motion that is considered unrealistic. The key idea here is that the amount of distortion that can realistically happen within short time-spans (i.e. from one A-scan to the next) is limited. Therefore, the time-derivative of the displacement functions with respect to time is subjected to a loss function and thereby penalized.

For computing the regularization term, the derivative of the displacement functions with respect to time needs to be calculated at fixed time points corresponding to the sampling time for every acquired A-scan(see equation (5.16)). In practice, the displacement functions need only be defined at the finite number of sampling time points contained in $\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$. Therefore, the time-derivative of the displacement functions can be calculated using finite differences. Using D_x as an example this yields

$$\frac{\delta D_x(t_l)}{\delta t_l} = \frac{D_x(t_{l+1}) - D_x(t_l)}{t_{l+1} - t_l}, \tag{5.26}$$

where t_l is a certain point in time contained in $\mathbf{T}^{\mathbf{X}}_{i,j}$ or $\mathbf{T}^{\mathbf{X}}_{i,j}$ (see equation (5.16)) and t_{l+1} is the point in time corresponding to the following A-scan. Therefore, forward differences are used. The displacement functions associated with the other two dimensions are handled in the same manner.

In practice the loss function is applied at a slightly different position within the formula, equation (5.16) becomes

$$E(D_x, D_y, D_z) = \sum_{t_l} \left(\frac{L(D_x(t_{l+1}) - D_x(t_l))}{t_{l+1} - t_l} + \frac{L(D_y(t_{l+1}) - D_y(t_l))}{t_{l+1} - t_l} + \frac{L(D_z(t_{l+1}) - D_z(t_l))}{t_{l+1} - t_l} \right), \tag{5.27}$$

such that the loss function is applied to the difference in displacement *before* dividing by the time difference. This causes the regularizer output at a particular A-scan to be scaled linearly by the inverse time difference. This is relevant because the time difference is not constant everywhere. Typically the time difference will be the same from one A-scan to the next, but going from the last A-scan of a B-scan to the first A-scan of the next B-scan there is a greater time difference due to flyback. The choice of where to apply the loss function causes the regularizer output in these two cases to be scaled consistently.

Two different loss functions can be used within the regularizer term. The first one is the squared loss L_2. As with the application of it in the similarity measure, this loss function will penalize high changes in displacement from one A-scan to the next disproportionally. In effect, this leads to the modeling of smooth displacements over time. This is consistent with the notion that the expected change in alignment within a very short amount of time is also small.

On the other hand, typical motion profiles, especially in the transverse directions, are defined by longer periods of very little to no motion broken by individual, very fast motions of relatively high amplitude (saccades, see section 3.2). While the squared loss function is able to very well cope with the former periods, fast high-amplitude motions that need to be modeled are associated with very high penalization by the regularization term.

One way to better accommodate the modeling of saccadic motion within the optimization of the objective function would be to use a function such as the $L_{0.5}$ norm as loss function. For one dimension it is defined as

$$L_{0.5}(x) = \sqrt{|x|}. \tag{5.28}$$

Within the regularization term, the $L_{0.5}$ norm would not only penalize changes in displacement proportionally to their absolute value, but would even penalize large changes disproportionally less than small ones. This property is useful for allowing for the modeling of saccadic motion. However, as is the case with the L_1 norm introduced before (equation (5.24)), the function is not continuously differentiable.

To overcome this problem while still retaining some of the positive features of $L_{0.5}$, we propose to approximate the $L_{0.5}$ norm using a function $L_{0.5,\epsilon_{0.5}}$. The function is defined as

$$L_{0.5,\epsilon_{0.5}}(x) = \sqrt{\sqrt{x^2 + \epsilon_{0.5}} - \sqrt{\epsilon_{0.5}}}, \tag{5.29}$$

where $\epsilon_{0.5}$ is a positive parameter that controls how well the function should approximate the $L_{0.5}$ norm. The smaller $\epsilon_{0.5}$, the closer the approximation. However, the closer $\epsilon_{0.5}$ goes to zero, the more extreme the values of the derivative of the loss function become close to $x = 0$. In practice, $\epsilon_{0.5}$ is empirically chosen to be 0.1.

One potential problem in this no-reference registration approach is the modeling of pathological solutions within the optimization. A particular problem here is that it is relatively cheap to model a displacement consistent with out-of-plane motion (see section 3.4.6). If the displacements over one B-scan accumulate to a one A-scan shift in the out-of-plane direction, the same B-scan of the input volume can be used twice for the construction of the registered volume. Of course there can be eye motion that causes exactly this effect. However, what can also happen is that such a one A-scan out-of-plane motion is modeled multiple times to replicate a certain B-scan because that would maximize the similarity to the other volume. This can cause pathological solutions which do not lead to a real motion corrected output.

In order to discourage the modeling of such pathological solution an additional factor s_{OOP} is employed within the regularizer. The factor linear scales the regularizer penalty that is associated with the respective out-of-plane component of the displacement fields. Therefore, for the displacements belonging to an **XFAST** volume s_{OOP} is applied after applying the loss to the time-derivative of $D_y(t)$. Conversely, s_{OOP} is used to scale the contribution originating from $D_x(t)$ for **YFAST** volumes. Given an adequate $s_{OOP} > 1$, displacements modeling out-of-plane motion can still be modeled but has to be justified better by the underlying volume data. This helps to stabilize the optimization and avoids pathological solutions.

5.5.3 Mean Displacement Term

In addition to the data similarity term and the regularization term, another term to help with finding a good solution is proposed. The issue addressed here is that the *mean displacement*, i.e. the mean of the displacement functions at all time points in all volumes, is not restricted in value by the formulation used so far. While the regularization term penalizes the time-derivative of the displacement functions, this leaves one degree of freedom per displacement function. This degree of freedom is that the regularization will not penalize any value that is added to the displacements at *all* time points.

The mean displacement in x-direction, corresponding to $D_x(t)$, is defined as

$$F_x(D_x) = \frac{1}{N_v \cdot w \cdot h} \sum_{t_l} D_x(t_l) \tag{5.30}$$

and analogously for the other two displacement functions F_y and F_z. The area in which the volume data is defined is limited. Therefore, a significant mean displacement in one of the dimensions means that the interpolation function I samples a certain fraction of the values from outside of the defined volume area. This is undesired. Up to motion effects and a shift in transverse area covered *between* the two volumes, the data that goes into the residual computation should stem from the valid area of the input volumes as much as possible.

In order to enforce this, an additional mean displacement term is added to the objective function, called F, defined as

$$F(D_x, D_y, D_z) = L(F_x(D_x)) + L(F_y(D_y)) + L(F_z(D_z)). \tag{5.31}$$

F can then be incorporated into the objective function by adding the term weighted with another factor β.

5.5.4 Combined Objective Function

Except for an additional term that is employed in the context of tilt compensation (see section 5.5.10), these three terms are combined to form the final objective function O. Based on equation (5.17) it is defined as

$$O(D_x, D_y, D_z) = S(D_x, D_y, D_z) + \alpha E(D_x, D_y, D_z) + \beta F(D_x, D_y, D_z), \tag{5.32}$$

with the corresponding optimization problem

$$D_x, D_y, D_z = \underset{D_x, D_y, D_z}{\operatorname{argmin}} O(D_x, D_y, D_z). \tag{5.33}$$

5.5.5 Displacement Field Parametrization

For solving the optimization problem it is only necessary to find suitable values for the displacement functions at a finite number of points in time t_l, where $l = 1, ..., N_v \cdot w \cdot h$ (see equation (5.16)). As long as the displacement functions can still be evaluated at these fixed points in time, they can be expressed in a number of different ways. In particular, they can be defined as functions of a *parameter set* $P = \{p_1, p_2, ..., p_{N_P}\}$ that can be evaluated at least at the times t_l. N_P denotes the number of parameters in the parameter set. To show the dependence on a parameter set, the displacement functions $D_x(t)$, $D_y(t)$ and $D_z(t)$ are denoted as D_x^P(t), $D_y^P(t)$ and $D_z^P(t)$, respectively. The optimization problem becomes

$$P = \underset{P}{\operatorname{argmin}} O(D_x^P, D_y^P, D_z^P). \tag{5.34}$$

Three different parametrizations are used in this work. The first and most straightforward one is to make every displacement value at every t_l for each of the displacement functions a parameter. We call this parametrization direct (per-A-scan) parametrization. The parameter set is denoted P^{dir} in this case and the number of parameters is $N_P = 3 \times N_v \times w \times h$. The parametrization can be expressed as

$$\begin{aligned} D_x^{P^{dir}}(t_l) &= p_{1+3\cdot l}, \\ D_y^{P^{dir}}(t_l) &= p_{2+3\cdot l} \quad \text{and} \\ D_z^{P^{dir}}(t_l) &= p_{3+3\cdot l} \end{aligned} \tag{5.35}$$

for all t_l. Additional parametrization types will be described in the context in which they are used in the sections on multi-resolution (see sec: multi-resolution optimization), multi-stage optimization (see section 5.5.9) and tilt compensation (see section 5.5.10).

5.5.6 Volume Interpolation

The volume interpolation function I is used within the objective function to sample volume data of the input or preprocessed volumes which are defined as fixed regular grids in the scanner coordinate system. The displacement function values control where exactly sampling is done. The interpolated intensities from corresponding grid points of the two volumes are then subtracted to compute the residual (see equation (5.12)) and a loss function is applied. In the concrete implementation, I is implemented using cubic spline interpolation [Keys 81]. Specifically, based on having to interpolate three dimensional data, tri-cubic Hermite spline interpolation is used [Leki 05].

Compared to, for example, tri-linear interpolation, this type of interpolation function offers continuous first derivatives. The theoretical requirements of the numerical optimization method used (see section 5.5.7) state that the objective function has to be twice continuously differentiable. Since every other part of the objective function is n-times continuously differentiable, the degree of continuity of the interpolation function determines the overall continuity. Therefore, the requirement is not fulfilled.

The effect of using a twice continuously differentiable interpolation function was also tested. For this purpose, cubic b-spline based interpolation was used [Thev 00]. The drawback of using this method is that a pre-filtering step on the data is required in order to make the function interpolate the original data points. Also, in practice switching to b-spline based interpolation did not improve the optimization results, despite better theoretical justification. Therefore, it was decided that cubic Hermite spline interpolation is preferable based on the run-time advantage it provides.

For performing unconstrained optimization, the interpolation function needs to be defined in all of $\mathbb{R}^3$. However, the volume grid on which intensity data is defined is finite. Therefore, the finite grid is extended to an imaginary infinite grid on which the interpolation function operates. This is governed by the *boundary conditions* of the interpolation.

The rules for extending the grid are to repeat the last grid point in the transverse directions. In the axial direction, the mapping is different. Here, the grid is continued such that the topmost position is followed by grid points from the most deep axial positions. Three grid indices i, j and k that can be out of bounds are mapped to three indices $\hat{i}$, $\hat{j}$ and $\hat{k}$ that are in bound according to the following rules:

$$\hat{i} = \begin{cases} i, & \text{if } 1 \leq i \leq w \\ 1, & \text{if } i < 1 \\ w, & \text{if } i > w, \end{cases} \tag{5.36}$$

$$\hat{j} = \begin{cases} j, & \text{if } 1 \leq j \leq h \\ 1, & \text{if } j < 1 \\ h, & \text{if } j > h \quad \text{and} \end{cases} \tag{5.37}$$

$$\hat{k} = \begin{cases} k, & \text{if } 1 \leq k \leq d \\ d - i, & \text{if } i < 1 \\ i - d, & \text{if } i > d. \end{cases} \tag{5.38}$$

Data associated with the grid position given by $\hat{i}$, $\hat{j}$ and $\hat{k}$ is then used to produce the interpolated value.

5.5.7 Optimization Strategy

The objective function is obviously a non-linear function. For any but very carefully constructed input data the objective function is also not convex. Therefore, iterative numerical methods are used to optimize.

Depending on the concrete parametrization, the volume size and number and other factors the objective function can depend on many parameters. For example, given two $400 \cdot 400$ A-scan OCT volumes as input and a direct, per A-scan parametrization (P^{dir}, see section 5.5.5), $400 \cdot 400 \cdot 2 \times 3 = 960000$ parameters need to be optimized for.

In addition to the objective function being dependent on a potentially large number of parameters, it can also be relatively expensive to compute. This mainly stems from the fact that interpolation of two whole volumes has to be performed based on the parameters. Acceleration techniques are being used (see section 5.9). It seems prudent to minimize the number of times the objective function has to be evaluated.

Based on these considerations, the choice was made to employ a gradient-based Quasi-Newton optimization method, namely limited-memory Broyden - Fletcher - Goldfarb - Shanno (L-BFGS) [Noce 80, Noce 99]. For use of this method the gradient of the objective function with respect to the parameters needs to be available. A finite difference scheme is not advised for this purpose. This is because of the potentially large number of parameters combined with the expensiveness of computing the objective function value. Instead, the gradient is computed analytically.

On a high level, the class of gradient-based iterative optimization methods consists of two major steps. First, a *descent direction* is chosen based on the current and previous gradient and function values. Second, a *line search* is performed. This means that based on the current parameter vector, a line is traced along the descent direction and the minimal objective function value along this line is to be found. This minimum is then the parameter vector for the next iteration.

In practice, a so-called inexact line search is performed [Noce 99]. In this case the actual minimum along the line does not have to be found. Instead, it is sufficient that the position that is found satisfies the so-called Wolfe conditions [Wolf 69]. Here, the use of L-BFGS offers a potential advantage with respect to for example the related conjugate gradients (CG) algorithm [Liu 89]: The L-BFGS method incorporates scaling into the descent direction estimate, such that a position that satisfies the Wolfe conditions can often be found without performing an actual line search [Liu 89]. Therefore, the use of L-BFGS can save time when optimizing

because the number of objective function evaluations needed for a similar optimization performance can be less.

5.5.8 Multi-Resolution Optimization

Given a non-convex objective function, iterative numerical optimization methods can in general only reach a local minimum. This is opposed to the global minimum with the lowest overall objective function value over the whole space spanned by the parameters and is undesired. Therefore, as is the case in many image registration techniques, a multi-resolution (MR) optimization technique is employed [Zito 03].

The concept of MR optimization is based on having a sequence of subsequently simpler optimization problems that are derived from the original optimization problem. In this context, an optimization problem is the combination of input data, objective function and its parametrization. The optimization problems need to be related in such a way that the solution to a simpler problem can be mapped to an initialization for the next more complex problem. Also, the mapped simple solution that becomes the initialization for the next problem needs to be a good initialization, i.e. it needs to be close to an optimal solution. The optimization then starts out with optimizing the simplest problem in the sequence. The result is mapped to the next more complex problem which is also optimized. This process repeats until the original problem has been optimized. Figure 5.3 shows a schematic view of the process.

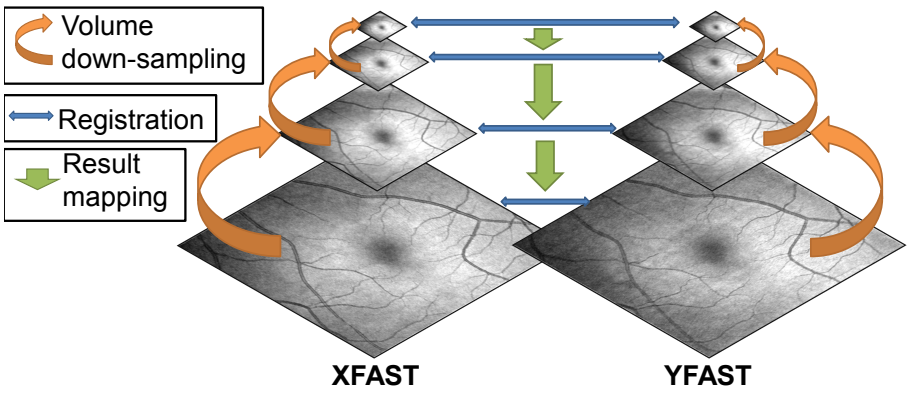

Figure 5.3: Multi-resolution optimization schematic.

For this particular problem, the successive simplification is realized by construction of a *volume pyramid*, i.e. a sequence of recursively down-sampled versions of the original volumes. From each (pre-processed) input volume, $N_{pyr} - 1$ versions are constructed by successively down sampling the volume such that the size in each dimension of halved. The down-sampling method itself is the same as the down-sampling in axial direction used in pre-processing (see section 5.4.4). Also, the intensities of each resulting volume are normalized such that among the volumes of one level the mean is zero and the variance is one (see section 5.4.5). This leads to a volume pyramid consisting of N_{pyr} volumes for each input volume.

Note that intensity data from different B-scans is combined during this process. This represents a potential problem as there might have been motion between the acquisition of the data that gets combined, which can lead to data from far away locations on the object to be combined. From a physical point of view this is not meaningful. On the other hand, the data in the pyramid only needs to serve as an approximation in order to eventually initialize the full resolution problem which does not suffer from this potential problem.

Along with the volume data itself, the time information as given by $\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$ needs to be mapped to lower resolution volume data. Here, successive levels are generated by treating the time structure information as 2D images and down-sampling them similar to the volume data. Similar to the volume down-sampling case the combination of sampling times from different B-scans is not a meaningful operation. However, the overall structure that grid points along the fast direction are close in time while they are more separated along the slow scan direction stays intact. From this perspective, the down-sampling serves as a useful approximation. In keeping with treating the time structure as a 2D image, the low-pass filtering before down-sampling also minimizes aliasing effects.

As a final part of MR optimization the solution of a low-resolution level needs to be mapped to the next higher resolution level. Also, the parametrizations of the problems that needs to be mapped in between might be different. We approach this problem in the following way: If the source parametrization is not direct, per-A-scan it is first converted to direct. Since all parametrization need to be evaluable to a 3D displacement at every A-scan grid point, and the set of these 3D displacements constitutes a direct parametrization this can readily be achieved. As a second step, the displacement field is treated as a three channel 2D image (see equation (5.15)), one channel for each displacement dimension. Each of these channel images is up-sampled to the size corresponding to the A-scan grid of the destination level using bi-linear interpolation. Finally, if the destination parametrization is not a direct per-A-scan parametrization, a least-squares fit is performed to map the displacement fields to the target parametrization.

Within each MR level, the amount of time spent optimizing the respective problem is restricted and can be configured. Instead of waiting for convergence of the optimization (i.e. the magnitude of the gradient going to zero), the maximum number of objective function and gradient evaluations is restricted to $N_{ev,m}$ for the respective multi-resolution level m.

In order to improve stability of the optimization, an additional parametrization is introduced which is intended to be used for the lowest resolution problem representations. Instead of assigning a parameter to every displacement over every A-scan, the parametrization associates a single parameter for every B-scan and displacement function. The parametrization P^b is expressed for all t_l as

$$\begin{aligned} D_x^{P^b}(t_l) &= p_{1+3\cdot T(t_l)}, \\ D_y^{P^b}(t_l) &= p_{2+3\cdot T(t_l)} \quad \text{and} \\ D_z^{P^b}(t_l) &= p_{3+3\cdot T(t_l)}, \end{aligned} \tag{5.39}$$

where $T(t)$ is a function that associates a time t with an index of the B-scan that was acquired at that particular time and $T(t_1) = 0$. The single parameter per B-scan and dimension helps to model only larger scale and per-B-scan consistent motions. This constraining of the degrees of freedom is useful for the lowest resolution representation as they represent a significant approximation of the original problem and might not be reliable in estimating per A-scan displacement values.

5.5.9 Multi-Stage Optimization

In order to further increase the robustness of the motion correction algorithm, the multi-resolution optimization can optionally be combined with a multi-stage approach. In this particular implementation, two stages are employed. Figure 5.4 shows a schematic of the two stage registration pipeline. In the first stage, the input volumes are registered in axial direction only. This is accomplished using the same objective function together with a parametrization that only allows for non-zero displacements in axial direction.

The parametrization P^{ba} is expressed for all t_l as

$$\begin{aligned} D_x^{P^{ba}}(t_l) &= 0, \\ D_y^{P^{ba}}(t_l) &= 0 \quad \text{and} \\ D_z^{P^{ba}}(t_l) &= p_{T(t_l)}, \end{aligned} \tag{5.40}$$

where $T(t)$ is defined as before. In addition to only having axial degrees of freedom in displacement, the parametrization also only has one displacement parameter per B-scan, as opposed to per A-scan. This design reflects the coarse nature of the first stage registration.

The optimization in the first stage is also carried out in a multi-resolution fashion. Once the first stage has finished, the resulting displacement fields are then applied to the original input data to produce a set of roughly registered volumes (see section 5.6). This set of volumes is then used as the input for the second optimization stage. Again MR optimization is used. This time however, transverse degrees of freedom are also optimized. Finally, the resulting displacement fields are applied to the input data of the second stage to produce registered volumes. Subsequently, these can be merged into a single, higher quality volume (see section 5.7).

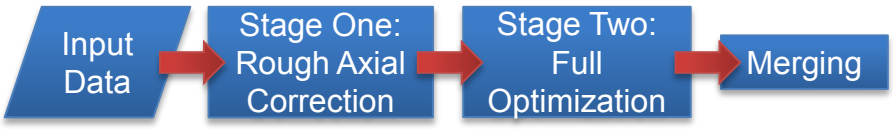

Figure 5.4: Multi-Stage optimization pipeline schematic. In the first stage, the input data is subjected to a rough axial correction. The solution of this optimization is used as the input to a second stage that performs full optimization. Finally, merging of the results from stage two is performed.

One positive effect of employing multi-stage optimization as described is related to non-linear optimization itself. As mentioned previously, the iterative numerical methods that are used can in general only find the nearest local minimum. Multi-stage optimization tries to alleviate this problem. The MR optimization of the simpler, axial-only problem and the generation of roughly registered inputs for stage two improves the starting position for the second stage optimization. Misalignment in axial direction that would have to be modeled by displacements has already been removed by the first stage.

A second effect is that through axial alignment and in particular through tilt compensation (see section 5.5.10) the coupling between axial and transverse degrees of freedom is reduced in the second stage. This is beneficial for finding a good solution. The following scenario serves as an example for this effect. There is a transverse mis-alignment between A-scans of corresponding anatomical locations of one A-scan spacing. In one scenario, there is little axial position difference of where the retina starts in these A-scans. In another there is a significant difference due to axial motion. In both cases, the optimization has to model a transverse displacement to match the locations. In the first case however, a significant axial displacement has to be modeled in addition. If the volumes are already roughly aligned as in the second example, this *coupled* axial displacement is much less. Due to the regularization, there is also more "resistance" to model the additional axial motion, making it less likely in the coupled case for the solution to be found.

5.5.10 Tilt Compensation

So far, the description does not explicitly deal with the effects of tilting of the retina due to alignment (see section 3.4.5). If there is a difference in tilting between the input volumes, the optimization has to model the different tilt using axial displacements that are dependent on the transverse position (i, j). The magnitude of the derivative of these tilt-compensating displacements with respect to time can be significant. However, this is what is penalized by the regularizer. Therefore, compensating a significant difference in tilt in order to register the data is made difficult because the necessary displacements are penalized as axial motion by the regularizer.

In order to alleviate this problem, a difference in tilt can be compensated within the first stage of optimization. For this purpose, a parametrization that is axial-only and explicitly models tilting per B-scan is introduced. It contains two parameters per B-scan. One parameter specifies a constant axial displacement over the whole B-scan, consistent with axial motion. The second parameter is used to model tilting of the respective B-scan in axial direction. It acts as a slope that is multiplied by the A-scan index within the B-scan. The parameter set is denoted P^{bat} and leads to the displacements

$$\begin{aligned} D_x^{P^{bat}}(t_l) &= 0, \\ D_y^{P^{bat}}(t_l) &= 0 \quad \text{and} \\ D_z^{P^{bat}}(t_l) &= p_{1+2\cdot T(t_l)} + V(t_l) \cdot p_{2+2\cdot T(t_l)}, \end{aligned} \tag{5.41}$$

for all t_l and where $V(t)$ relates a time t with an index for the A-scan within the current B-scan at that particular time. When using this parameter set, the regularization is adapted to penalize the time-derivative of the per B-scan constant parameter and of the per B-scan tilt parameter, separately. This follows the notion that the change in slope induced by tilt from one B-scan to the next is expected to be small. Using this parameter set combined with the changes in regularization, tilting of one volume with respect to the other can be compensated as part of the rough axial compensation in the first stage.

However, the registration of a set of these volumes according to these degrees of freedom does not specify what the combined tilt of the registered retinas should be. It might be optimal from an objective function value point of view to make one volume's tilt the same as the other one's. Alternatively, the combined tilt could be an average of the tilt of the inputs. In pathological examples this formulation might lead to an optimization result that *creates* tilt when there was none in the original data. This can happen because the volume similarity term might be optimal for a high-tilt configuration.

In order to minimize coupling between transverse and axial degrees of freedom (see section 5.5.8), it would be preferable to produce a registration result where the retina content is not tilted. In order to guide the optimization in this direction, an additional data term is introduced.

Assuming that background voxels tend to be dark while retina pixels have high intensities, the amount of tilt of the retina influences the shape of the histogram of intensity versus axial depth. The (unnormalized) axial histogram function $H(k, D_x, D_y, D_z)$ associates a discrete axial pixel index k with the sum of the interpolated intensities at the corresponding axial position over all transverse positions and volumes. This can be expressed the following way

$$\begin{aligned} &H(k, D_x, D_y, D_z) = \\ &\sum_i \sum_j \Big(I(\mathbf{X}, x_i + D_x(\mathbf{T}^{\mathbf{X}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{X}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{X}}_{i,j})) + \\ &\qquad I(\mathbf{Y}, x_i + D_x(\mathbf{T}^{\mathbf{Y}}_{i,j}), y_j + D_y(\mathbf{T}^{\mathbf{Y}}_{i,j}), z_k + D_z(\mathbf{T}^{\mathbf{Y}}_{i,j})) \Big) \end{aligned} \tag{5.42}$$

It should be noted that due to the fact that interpolated intensities are used as weights within the histogram, they are not normalized in this case. This ensures that the weights that go into the histogram are all positive.

The normalized axial intensity histogram function $\hat{H}(k, D_x, D_y, D_z)$ is then defined as

$$\hat{H}(k, D_x, D_y, D_z) = \frac{H(k, D_x, D_y, D_z)}{\sum_{\hat{k}=1}^{d} H(\hat{k}, D_x, D_y, D_z)}. \tag{5.43}$$

This scales the values such that the sum over all entries equals one. Since the histogram is normalized, $\hat{H}(k, D_x, D_y, D_z)$ can be seen as a probability distribution over k.

If the retina is not tilted, the axial distribution of intensities will be mostly concentrated on a few bins. On the other hand, if there is high tilt or the two volumes are not aligned axially, the intensities will be distributed more uniformly

over depth. An indication for how concentrated the high intensities are axially is the *variance* of the probability distribution given by the axial histogram. High variance indicates high tilt or bad alignment of the volumes. Based on the algebraic formula of variance, the function expressing the variance $Var_{\hat{H}}(D_x, D_y, D_z)$ can be calculated as

$$Var_{\hat{H}}(D_x, D_y, D_z) = \sum_k \left(\hat{H}(k, D_x, D_y, D_z) \cdot k^2\right) - \left(\sum_k \hat{H}(k, D_x, D_y, D_z) \cdot k\right)^2 . \tag{5.44}$$

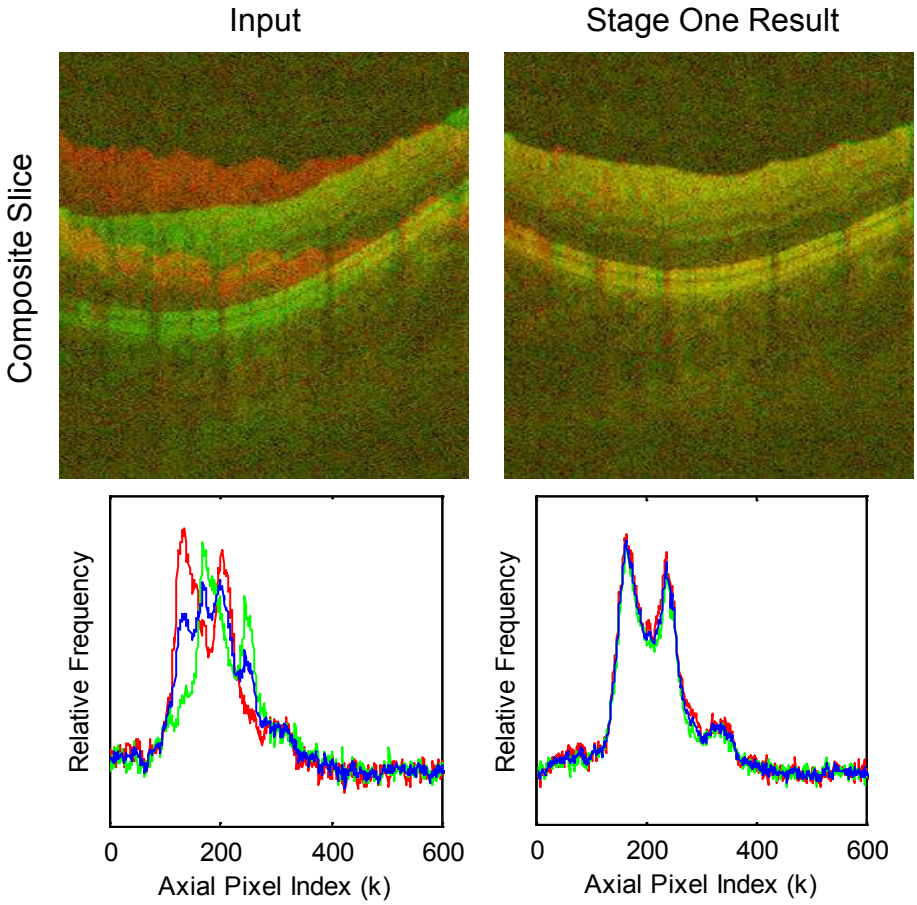

Figure 5.5: Schematic showing the effect of removal of tilt and alignment to the axial intensity histograms.

Figure 5.5 shows the effect that aligning two OCT images and the removal of tilting has on both the individual and the combined axial intensity histograms. In the composite views the two OCT images are shown before and after alignment and tilt compensation. The first image is shown in the red color channel, the second in the green one. Below, the red and geen curves show the individual histograms of the corresponding images. The blue curve shows the combined axial intensity histogram. It can be seen that the correction step causes the histograms

to overlap better. Therefore, the combined histogram has a more pronounced and narrower peak, leading to a reduced variance.

When performing tilt compensation, this term is incorporated into the objective function using another weighting factor γ. Starting from equation (5.32), the objective function becomes

$$\begin{aligned} O(D_x, D_y, D_z) =& S(D_x, D_y, D_z) + \\ & \alpha E(D_x, D_y, D_z) + \\ & \beta F(D_x, D_y, D_z) + \\ & \gamma Var_{\hat{H}}(D_x, D_y, D_z). \end{aligned} \tag{5.45}$$

$Var_{\hat{H}}(D_x, D_y, D_z)$ can also be expressed in terms of the un-normalized histogram by using equation (5.43). Equation (5.44) becomes

$$\begin{aligned} Var_{\hat{H}}(D_x, D_y, D_z) =& \frac{1}{\sum_{\hat{k}=1}^{d} H(\hat{k}, D_x, D_y, D_z)} \cdot \\ & \left(\sum_k \left(H(k, D_x, D_y, D_z) \cdot k \cdot k \right) - \right. \\ & \left. \left(\sum_k H(k, D_x, D_y, D_z) \cdot k \right)^2 \right). \end{aligned} \tag{5.46}$$

Calculating the derivative of this function with respect to the displacement field values at all time points $\mathbf{T}^{\mathbf{X}}_{i,j}$ and $\mathbf{T}^{\mathbf{Y}}_{i,j}$ is relatively complicated. This is mainly because the normalizing factor that is a denominator in the term depends on all displacement values. As the nominator terms are also dependent on the displacement values, the quotient rule would have to be used.

However, the normalization term basically is the sum of the intensities of the interpolated volumes. We can therefore approximate the term by assuming that regardless of the concrete displacements, the overall sum should stay approximately constant. The sum is therefore calculated once for the initial displacement field values and then treated as a constant $v_{\sum}$. The formula then simplifies to

$$\begin{aligned} Var_{\hat{H}}(D_x, D_y, D_z) =& \frac{1}{v_{\sum}} \cdot \\ & \left(\sum_k \left(H(k, D_x, D_y, D_z) \cdot k \cdot k \right) - \right. \\ & \left. \left(\sum_k H(k, D_x, D_y, D_z) \cdot k \right)^2 \right), \end{aligned} \tag{5.47}$$

which avoids having to use the quotient rule for computing the gradient.

In practice, the histogram is not composed from the sum of intensities over all transverse grid points. Instead, in order to speed up the evaluation of the term, only a subset of transverse grid locations (i, j) is taken for constructing the axial

histogram. In the concrete implementation, every 20est grid point in x and y direction is considered. In addition, every grid point that lies on one of the four outer edges of the transverse area is also used for the calculation. This is such that for every B-scan of every input volume there are at least two interpolated A-scans that can be used to estimate the corresponding parameter set parameters. Also, the contents of the A-scans that are taken into account can only move up and down in axial direction. This is due to the lack of transverse degrees of freedom in the parameter set that is used to compensate tilt (see equation (5.41)). Because of this, full coverage of all interpolated A-scans is not necessary. The intensity content of an interpolated A-scan can not move to a neighboring interpolated A-scan that is not taken into account in the variance term. Overall, compared to using all grid points, this leads to a major speedup in evaluating this term and its gradient.

5.5.11 Registration of more than two input volumes

So far, the description was limited to the case of registering two input volumes with each other, i.e. $N_v = 2$. However, the method can be extended to handle more than two input volumes in a relatively straightforward way. In the general case, there are N_v input volumes denoted $\mathbf{V}_n$, where n is the volume index with $n = 1, ..., N_v$. Likewise, sampling time is associated with each A-scan of every volume via $\mathbf{T}^{\mathbf{V}}_{n\,i,j}$.

The regularization, mean displacement and tilt data term can be adapted in a straightforward way. For regularization, the time derivative is just computed and summed over all time points from all N_v volumes. The same holds for mean displacement. To compute the axial histogram from equation (5.42) for more than two volumes, intensities are just summed over all N_v volumes, instead of just $\mathbf{X}$ and $\mathbf{Y}$.

The volume similarity measure $S(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z)$ (see equation (5.13)) computes the similarity between the volumes $\mathbf{X}$ and $\mathbf{Y}$ in the transformed state. In the case that there are are more than two input volumes, the similarity measure is computed between multiple pairs of volumes.

The computation of the similarity term is symmetric, this means that

$$S(\mathbf{X}, \mathbf{Y}, D_x, D_y, D_z) = S(\mathbf{Y}, \mathbf{X}, D_x, D_y, D_z). \tag{5.48}$$

In addition, the similarity term of a volume with itself is zero, i.e.

$$S(\mathbf{V}, \mathbf{V}, D_x, D_y, D_z) = 0. \tag{5.49}$$

Therefore, the maximum number of distinct volume similarity terms that can be computed from N_v input volumes is

$$\sum_{v=2}^{N_v} v - 1 = \frac{N_v \cdot (N_v - 1)}{2} = O(N_v^2). \tag{5.50}$$

The number of volume pairs of which the similarity can be considered grows with the square of the number of input volumes. Also, an individual similarity term

between two volumes can be relatively expensive to compute. Combined, these circumstances can make it prohibitively expensive from a computational point of view to consider all possible pairs of volumes within the objective function.

In order to avoid this problem, we chose only to consider a subset of all possible volume pairs. In the case of more than two input volumes, the combined similarity function is defined as

$$S_{multi}(D_x, D_y, D_z) = \sum_{v_1=1}^{N_v} \sum_{v_2=v_1+1}^{N_v} \theta(v_1, v_2) \cdot S(\mathbf{V}_{v_1}, \mathbf{V}_{v_2}, D_x, D_y, D_z), \tag{5.51}$$

where θ is a function that is one if the particular pair of volumes given by the indices v_1 and v_2 should be considered for the combined similarity term, and zero if not.

We choose to use only of order $O(N_v \cdot \log N_v)$ pairings, resulting in substantial computational savings. For this subset, pairings between volumes with orthogonal fast scan axes are preferred. The pairings are chosen in such a way that the undirected graph formed by the volumes (nodes) and pairings (edges) is fully connected. If the volumes are not connected through the objective function, they will not be registered to a common space.

5.6 Output Volume Generation

Once the multi-resolution, potentially multi-stage registration process has finished, a set of N_v registered output volumes are generated. These volumes are denoted by $\hat{\mathbf{V}}_n$ and are the same size as the input volumes. The volume data itself is produced by interpolating the respective input volume offset by the displacements as given in the final displacement field. This process is the same as the interpolation used when calculating volume residuals (see equation (5.12)). The output volume data is given by

$$\hat{\mathbf{V}}_{n,i,j,k} = I(\mathbf{V}_n, x_i + D_x(\mathbf{T}^{\mathbf{V}}_{n\,i,j}), y_j + D_y(\mathbf{T}^{\mathbf{V}}_{n\,i,j}), z_k + D_z(\mathbf{T}^{\mathbf{V}}_{n\,i,j})). \tag{5.52}$$

As input for the interpolation, the original data of the respective stage is used. This means that the data has not been subjected to pre-processing. The output generation step of stage one takes the original input data. The output step of stage two take the output of stage one as input.

In addition to generating the registered volumes themselves, the output volume generation step also generates a volume that signals whether the data contained in the registered volume at this particular voxel can be considered as valid data. Data is considered as invalid if the interpolation function had to sample the respective input volume outside of the range as defined by the underlying scan-

ner coordinate system grid. The sample validity volume corresponding to $\hat{\mathbf{V}}_n$ is denoted by $\mathbf{Z}_n$. Its data content is defined as:

$$\mathbf{Z}_{n,i,j,k} = \begin{cases} 0, & \text{if} \quad x_i + D_x(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) < x_1 \\ & \text{or} \quad x_i + D_x(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) > x_w \\ & \text{or} \quad y_j + D_y(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) < y_1 \\ & \text{or} \quad y_j + D_y(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) > y_h \quad \forall n,i,j,k \\ & \text{or} \quad z_k + D_z(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) < z_1 \\ & \text{or} \quad z_k + D_z(\mathbf{T}^{\mathbf{V}}_{n\,i,j}) > z_d \\ 1, & \text{otherwise,} \end{cases} \tag{5.53}$$

The information contained within these volumes is used within the weight generation step of the volume merging process.

5.7 Volume Merging

Given a set of registered volumes $\hat{\mathbf{V}}_n$ with $n = 1, ..., N_v$, a single, merged volume can be generated that is both motion-corrected and has higher signal quality than the individual input volumes. The merged volume is denoted $\mathbf{M}$ and is a per-voxel, convex weighted sum of the intensities of the registered volumes:

$$\mathbf{M}_{i,j,k} = \sum_{n=1}^{N_v} \left(\hat{\mathbf{V}}_{n,i,j,k} \cdot \mathbf{W}_{n,i,j,k} \right), \tag{5.54}$$

where

$$\mathbf{W}_{n,i,j,k} > 0 \quad \forall n,i,j,k \quad \text{and} \quad \sum_n \mathbf{W}_{n,i,j,k} = 1 \quad \forall i,j,k. \tag{5.55}$$

For now, let us assume that anatomical locations are registered with each other and that speckle noise (see section 3.4.1) is uncorrelated between voxels at corresponding grid points in the registered volumes. Also, the intensities of the registered volumes are in log-scale. Given this, the mean log-intensity over all registered volume voxels, i.e. $\mathbf{W}_{n,i,j,k} = 1/N_v$, will lead to an SNR gain of up to $\sqrt{N_v}$ (see section 4.2.1).

In practice though, not all interpolated voxels contain valid information of the supposed anatomical location. For instance, a particular voxel in a registered volume might have been produced by sampling outside of the scanner coordinate system grid of the respective input volume. This can happen if the overlap of the imaged areas is small. Such samples should not be considered when calculating the final intensity. Note that $\mathbf{Z}$ contains the required information for this for each volume.

5.7.1 Sampling Density

Another effect is related to motion artifacts. Due to motion, a certain anatomical location might not have been imaged in one volume but in another. For example,

due to out of plane motion (see section 3.4.6) whole transverse regions of the retina might have been skipped during imaging of one volume. Therefore, the scan coordinate grid of one of the input volumes might not contain a certain object space location. However, a certain grid point in the registered or object space can correspond to such a location that is missing. In a volume where the respective location was sampled, the displacements will ideally be set such that the correct location is taken from the input volume. In the case that a location was not sampled to begin with there is no correct location in the input volume to sample from. The optimization has to sample from somewhere, though.

In the volume with the missing information, the displacements for this particular grid point will likely be set to sample from a neighboring location that has similar intensities as the interpolated A-scan from the other volume to which it is matched to. However, the particular location in the input volume that has been taken to fill this gap in the data is also likely to be matched to its real anatomical counterpart in the other input volume. The result of this example is that one particular location in one input volume is used more than once to interpolate from.

Therefore, if for every A-scan, we look in the registered volumes where the sample was taken from in the corresponding input volume and how often it was used, this can give an indication whether this particular location in the registered volume had no data available. If a location in an input volume was often sampled from, the corresponding locations in the registered volume are likely to contain no valid data. We call this quantity *sampling density*.

In order to estimate sampling density, Parzen density estimation is used [Parz 62]. As a kernel function, an isotropic 2D normal distribution with standard deviation σ_p, which was empirically chosen to be 0.5 pixels, is used. The 2D normal probability density function centered at (c_x, c_y) is denoted by $\mathcal{N}(x, y, c_x, c_y)$ and defined as

$$\mathcal{N}(x, y, c_x, c_y) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{\frac{(x-c_x)^2+(y-c_y)^2}{2\sigma_p^2}}. \tag{5.56}$$

Using this formula, the sampling density $\mathbf{SD}^{\mathbf{V}}_{i,j}$ for a volume $\mathbf{V}$ and its corresponding displacement field $\mathbf{D}^{\mathbf{V}}$ is

$$\mathbf{SD}^{\mathbf{V}}_{i,j} = \sum_{a=1}^{w} \sum_{b=1}^{h} \mathcal{N}\left(x_i, y_j, x_a + \mathbf{D}^{\mathbf{V}}_{1,a,b}, y_b + \mathbf{D}^{\mathbf{V}}_{2,a,b}\right). \tag{5.57}$$

Before weight generation, the sampling density is calculated for every registered volume $\mathbf{V}_n$. In practice, the estimation is discretized in order to save run-time. The density is estimated on a grid that is four times the size as the original A-scan in each direction. This is done to enable sub-pixel accuracy. For estimation itself, the kernel is discretized and truncated such that it has finite support and then added to the grid, centered on each sample location. The final look-up of the sampling density then consists of a bi-linear interpolation look-up at the necessary points.

5.7.2 Weight Generation

Using sample validity and sampling density information, the weight generation for merging follows the following principles:

- If sample validity for a sample is low compared to the other samples, assign a low weight
- If sampling density for a sample is high compared to the other samples, assign a low weight
- If there are no differences in the first two factors, assign evenly distributed weights.

As a first step, the weights are constructed using the formula

$$\mathbf{W}_{n,i,j,k} = \frac{\mathbf{Z}_{n,i,j,k}}{(1 + \mathbf{SD}_{i,j}^{\mathbf{V}_n})^8}. \tag{5.58}$$

The denominator is always greater one and depends on the sampling density raised to the eight power. This ensures that the weights are very sensitive to changes in sampling density. Therefore, even if the sampling density for one volume is only slightly higher than for the other(s), the weight will be disproportionally much lower. Currently the power that is used is somewhat arbitrary. In the future it might be worthwhile to perform additional investigation on the weight generation formula. Subsequently, the weights are normalized such that $\sum_n \mathbf{W}_{n,i,j,k} = 1$.

5.8 Processing of Additional Data Channels

In addition to registering and merging the intensity data that the algorithm operates with to find the displacement fields, additional data channels that are inherently registered with the intensity information can also be motion-corrected, registered and merged. Examples of additional data channels include functional OCT channels such as Doppler shift information [Baum 11a], speckle decorrelation flow information [Jia 12] and polarization sensitive OCT information [Baum 12].

In order to process additional channels, the displacement fields that were estimated need to be applied to the functional data. Depending on the type of information that is stored in the channel, the interpolation needs to be carried out in a different way. Speckle decorrelation information for example, is similar to OCT image intensity in meaning and can be interpolated in the same way.

On the other hand, Doppler shift information is phase-like and has a 2π ambiguity that needs to be considered when interpolating. One way to solve this problem is to map the phase to a corresponding complex number, a *Phasor* to calculate the weighted directional mean [Fish 95]. A single phase volume is mapped to two volumes this way: One volume contains the real part of the Phasor, the other the imaginary part. The two volumes can then be independently interpolated. Subsequently, the Phasor as given by the interpolated real and imaginary parts can be

mapped back to the phase-like Doppler shift. The combination of multiple phases in this way results in correct treatment of the 2π ambiguity.

Since interpolation fundamentally consists of a weighted combination of data samples, the weighted combination step when merging registered functional information is analogous. Note that when merging additional data channels, the weights are still computed based on the original intensity information and then applied to the other data.

5.9 GPU Acceleration

From a computational perspective, the most expensive operations are the the computation of the similarity measure and its gradient (see equation (5.13)) and median filtering within pre-processing (see section 5.4.1). In order to optimize the run-time of the method, these two parts have been accelerated using the Compute Unified Device Architecture (CUDA) programming language for programming on massively parallel GPUs [Nick 08].

For median filtering, both 1D and 2D filtering were implemented on the GPU. One thread was created for each output voxel and executed in parallel. The necessary data was collaboratively fetched among a group of threads using shared memory. In order to compute the median element, bubble sort was used. While not being the most efficient sorting algorithm, it can be implemented using two nested, fixed size loops and only max and min operations. This avoids branching altogether.

The combined computation of the similarity measure and its gradient with respect to a direct 3D per A-scan parametrization was achieved in two steps. First, one thread is spawned per transverse grid point that sets up the interpolation. Here the data pointers that are needed for interpolation (boundary condition handling) and the corresponding coefficients for interpolation and for the derivative of interpolation with respect to the three spatial dimensions are computed. In the second step, one thread is generated per voxel with the threads of one work group being arranged along the axial direction. Each thread interpolates two volumes, computes the residual value and its derivatives with respect to the 6 displacement dimensions for the two volumes. The values are then accumulated along the axial direction. In the end, one similarity value and six derivative values are produced per A-scan. If the actual parametrization of the displacement field is not direct 3D per A-scan the gradient can be mapped to any other parametrization using the chain-rule. If needed, this step is currently performed on the CPU.

5.10 Summary

Within this chapter, a detailed description of the method to motion-correct, register and merge a set of 3D-OCT volumes was given. The method is formulated as a specialized kind of registration problem where there is no reference and all volumes have to be transformed in order to register the volumes. A problem-specific regularization based on the time-structure of the OCT scanning process is

employed. Furthermore, in order to cope with potential inconsistent intensity information several methods are introduced: These include illumination correction and a pseudo-Huber-norm loss function within the data similarity term.

For the optimization of the objective function, multi-resolution and multi-stage methods are employed. In addition, differences in alignment related tilt between the input volumes can be compensated within the first optimization stage via specific modeling of the effect of tilt on the volume and an additional data term that ensures that tilt is removed during registration.

Once the volumes have been registered, the set of registered volumes can be merged into a single higher-quality volume. For this purpose an adaptive weighting scheme based on the concepts of sample validity and sampling density is introduced. In addition to intensity data, functional data channels that are associated can also be motion corrected and merged. Finally, in order to improve the run-time performance of the method, key parts were optimized using GPU programming.

CHAPTER 6

Evaluation Approach

In this chapter, the methodology for the evaluation of the proposed motion correction method is described. Parts of this chapter have been published in prior publications of the author [Krau 12, Krau 14].

6.1 Introduction

The motion correction method removes motion artifacts and registers multiple volumes with orthogonal scanning together by applying a transformation to *each* input volume. This is because all acquired volumes can be distorted by motion, there is no motion free reference available.

This poses a challenge for evaluating the method. As opposed to standard image registration, it is not sufficient to show that anatomical locations are mapped onto each other after registration. The volumes could be registered onto each other but the registered space itself might be distorted. Because there is no fixed reference, this effect can occur.

One option would be to use a secondary modality that does not suffer from motion artifacts to evaluate how close the motion correction result is to the actual morphology of the retina. Other modalities that are used for imaging the eye include Ultrasound (US) , Magnetic Resonance Imaging (MRI) , Fundus Photographs and Confocal Microscopy. However, comparing between OCT and these modalities is difficult due to much lower resolution of the other modalities (US, MRI), no or limited availability of 3D information (Fundus, Confocal) and different distortion of the images of the different modalities due to the imaging process. For example, an Ultrasound volume is distorted by refractive effects at interfaces

between different tissues where the speed of sound changes. Sources of (static) distortion in OCT are due to the angular scanning and optical distortions. In addition, use of a secondary modality would naturally necessitate that the device is available (which would for example be hard in the case of MRI) and that additional imaging is performed.

In order to give a good indication of the effectiveness of the method in a real-world scenario, the evaluation also needs to be large-scale. This means that data sets from many subjects need to be acquired and used in the evaluation. Also, the set of subjects needs to be a realistic cross-section of subjects that would be encountered in a clinical context. Specifically, subjects with retinal pathologies and older subjects need to be included in the evaluation as these kinds of subjects would also be encountered in actual clinical practice.

6.2 High Level Approach

Based on these considerations, it was decided against using additional imaging modalities for evaluation. Instead, the quantitative evaluation is based on three main pillars:

- Evaluate the ability of the registration to map corresponding locations onto each other, maximize the similarity of the volumes after registration and remove obvious motion artifacts (registration performance).
- Evaluate how well registered, merged volumes correspond to the actual anatomy. Multiple, disjoint sets of OCT volumes of the same location are used as input for motion correction and merging. After processing, the similarity and the reproducibility of quantitative measurements from of the multiple merged volumes is evaluated (reproducibility).
- Evaluate whether signal quality improves through motion correction and merging (signal improvement).

Good motion correction will show both high registration performance and high reproducibility of the results. Also, given that the input volumes are well registered by the motion correction, we expect the signal quality to improve through the merging step.

In addition to the quantitative evaluation, visual inspection of example sets of volumes before and after application of the algorithm is used. Furthermore, the run-time of the method represents an important measure that determines the acceptance in clinical practice and is therefore also evaluated.

6.2.1 Registration Performance

For evaluating the registration performance, the similarity between the input volumes is compared with the similarity between the set of registered volumes that are produced by the algorithm. Figure 6.1 shows a high level overview of the registration performance evaluation work flow. Given a pair of input volumes with

orthogonal scan patterns (shown in the green box) the evaluation proceeds by first calculating the similarity of the two volumes before motion correction. Also, motion correction is performed which yields two registered volumes. Similarity is also assessed for these registered volumes. The more the similarity increases by performing motion correction, the better the registration performance.

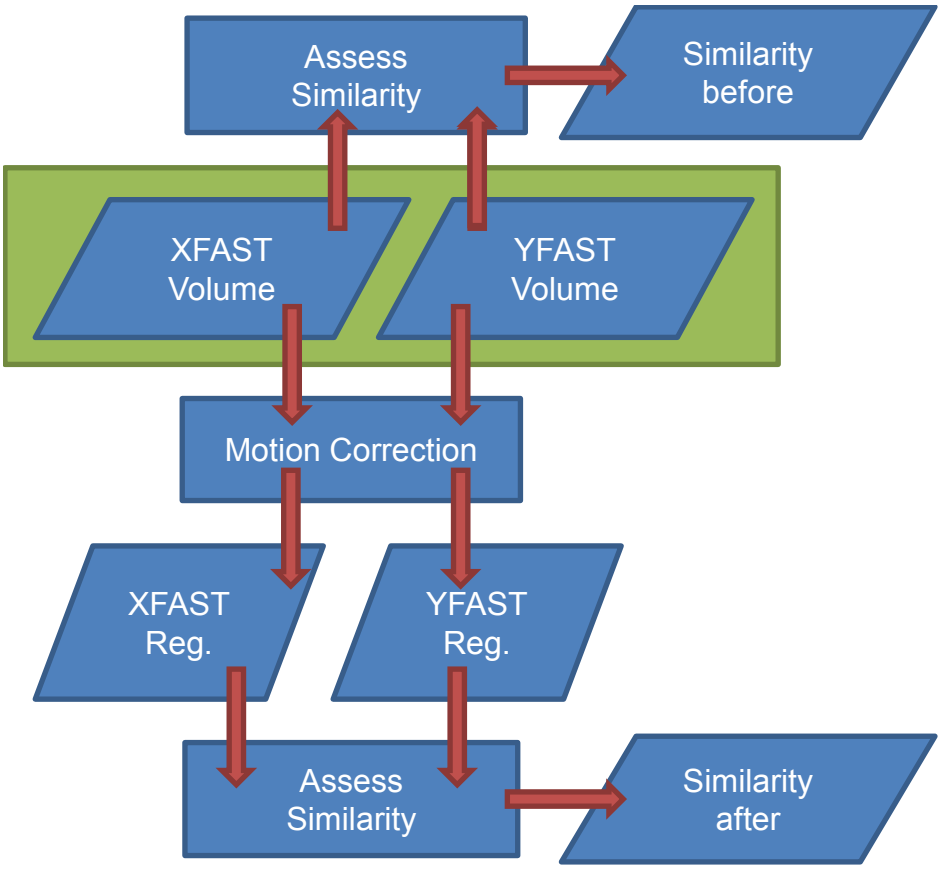

Figure 6.1: Registration performance evaluation work flow. Input data is contained within the green box.

6.2.2 Reproducibility Performance

The idea of the reproducibility evaluation is that if multiple volumes of an area of an eye are acquired in direct succession, the actual imaged object, i.e. the retina, can be assumed to remain the same. The volume data itself however will differ between the volumes mainly because of noise, illumination, motion artifacts and alignment related effects. Alignment related effects are modeled as a 3D translation of the whole volume content, a possible rotation around the optical axis and a tilt in the other two directions [Krau 12]. If a set of volumes from the same location has no motion artifacts or motion artifacts have been removed by motion correction it should be possible to align the set of volumes with each other using quasi-rigid registrations that model said alignment related effects. The higher the similarity between the volumes at this point, the better the alignment.

Figure 6.2 shows a high level overview of the work flow. As input, *two* pairs of orthogonal input volumes from the same location are needed. Each pair of input volumes is then motion corrected and merged, yielding two merged volumes. Subsequently, quasi-rigid registration is performed on these merged volumes in order to align them. This step yields two quasi-rigidly registered, merged and motion corrected volumes. Finally, the similarity between these two volumes is evaluated. The higher the similarity, the better the reproducibility performance.

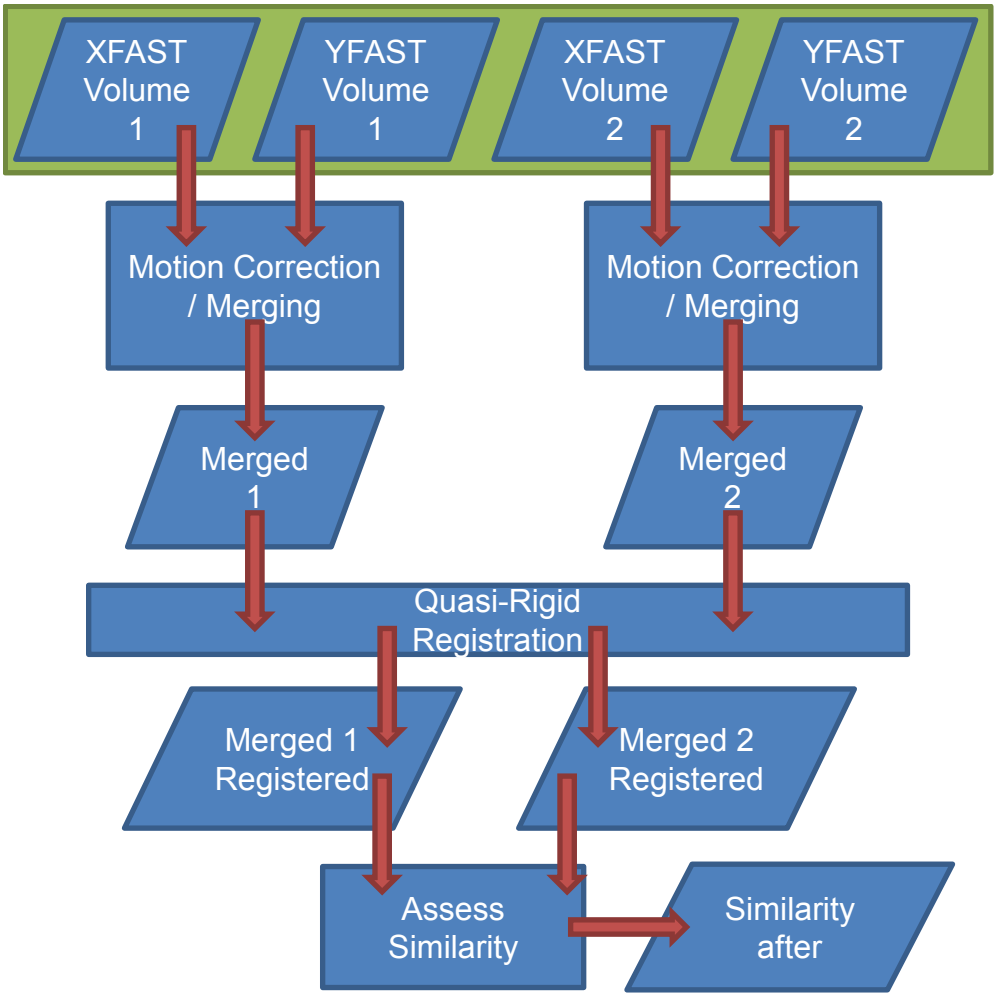

Figure 6.2: Reproducibility evaluation work flow. Input data is contained within the green box.

6.2.3 Signal Quality

Evaluating whether signal quality is improved is non-trivial, once again because there is no ground through information available. This prevents an evaluation such as used by Mayer et al. which calculated peak signal-to-noise ration (PSNR) with respect to a high quality reference OCT image obtained from an ex-vivo sample [Maye 12]. Instead, a no-reference measure of image quality called the Q-metric is employed [Xian 10] (for details, see section 6.6.6). The higher Q, the higher quality the underlying image.

Figure 6.3 shows a schematic view of the work flow to quantitatively evaluate whether signal quality improves using the Q-metric. A pair of input volumes is motion corrected, producing a set of registered volumes. The Q-Metric is calculated on this set to get an indication of the signal quality of the input volumes. Since there are two registered volumes, the mean of the Q-metrics of the individual volumes is the resulting quantity of this step. Note that while the registration step will un-distort the volumes it should not change the signal quality itself. The set of registered volumes is then used as input for the merging process, producing a single merged volume. Again, the Q-metric is calculated of this merged volume. The merging combines two (or more) volumes, basically by averaging of the intensities. Therefore, it would be expected that the signal quality of the merged volume is improved relative to the registered volumes if the volumes are well registered, similar to angular compounding (see 4.2.1). The change in Q-metric between registered and merged volumes is used as a measure of this change in signal quality.

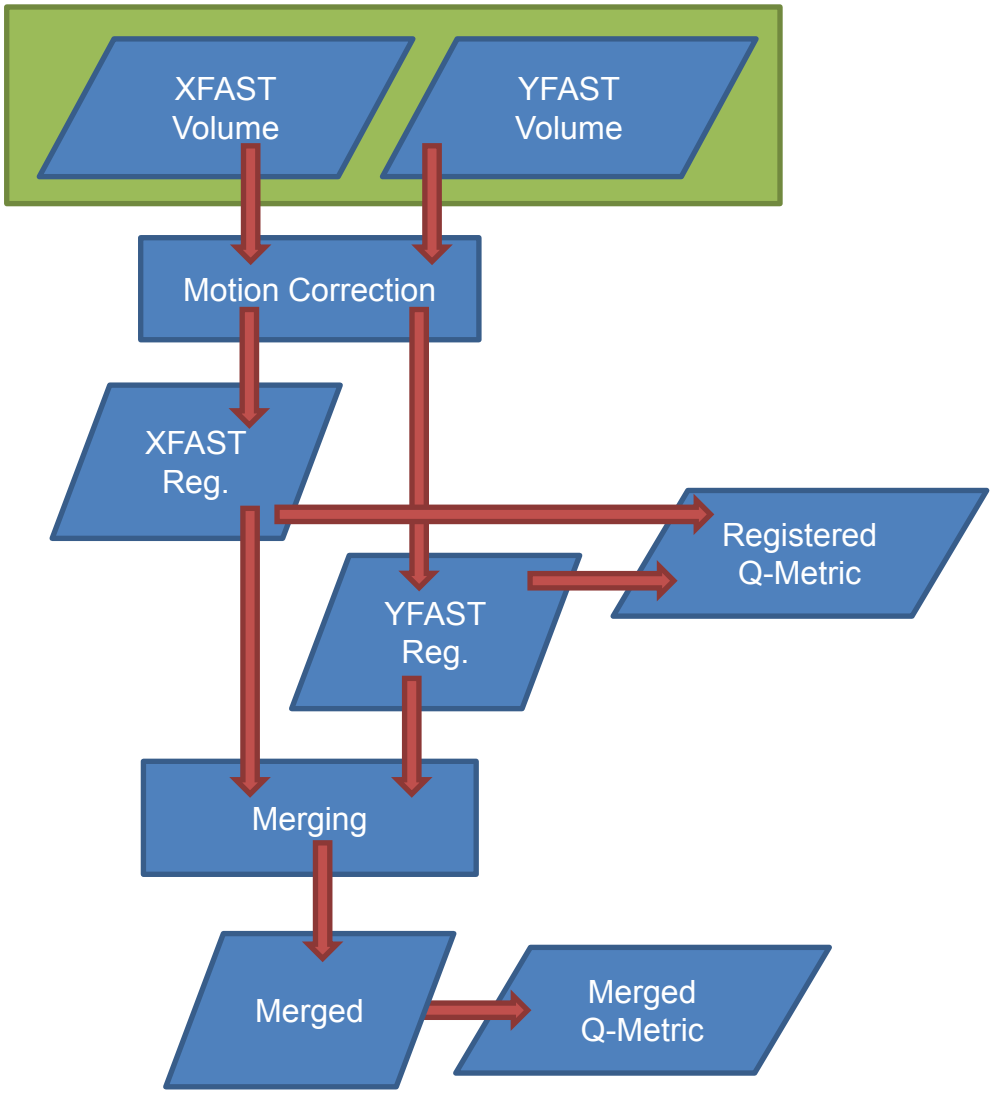

Figure 6.3: Signal improvement evaluation work flow. Input data is contained within the green box.

6.3 Quantitative Similarity Measures

Both for evaluating registration and reproducibility performance the similarity of volumes that are mapped into a common space by a transform has to be assessed. For registration performance, the two volumes are the XFAST and YFAST input volumes and the transform is the displacement field for each volume that is used to motion correct and register the two volumes. For reproducibility evaluation, the two volumes are two merged algorithm outputs with disjoint sets of XFAST/YFAST volume pairs as input.

6.3.1 Mutual Information

Mutual information (MI) is used as one way to quantitatively evaluate the similarity of two volumes. 128 bins equally spaced between the overall minimum and maximum intensities of the volumes are used for the intensity histograms. The intensity histograms of the individual volumes are denoted $p(a)$ and $q(a)$, respectively, where $a = 1, \ldots, 128$ is a discrete bin index. Also, the 2D joint histogram $h(a, b)$ is calculated, with $b = 1, \ldots, 128$ again being a bin index. The mutual information MI is then calculated according to the formula

$$MI = \sum_{a=1}^{128} \sum_{b=1}^{128} \left(h(a,b) \log \frac{h(a,b)}{p(a)q(b)} \right). \tag{6.1}$$

6.3.2 Segmentation-based Similarity Assessment

The second approach for assessing the similarity of two volumes uses the similarity of per A-scan segmentation maps. Figure 6.4 shows a schematic view of the process. Segmentation maps are extracted from the volumes in their original space. The segmentation maps measure the positions of retinal layers boundaries, the thickness of specific retinal layers and whether a blood vessel is present at a particular A-scan. The transform that is obtained by either the motion correction process or by quasi-rigid registration is then applied to the segmentation maps in order to map them to a common space. Once in a common space, the measurements at corresponding locations are compared. The lower the absolute difference between these measurements the more similar the underlying volumes from which they were obtained. Correspondingly, the presence of differences points to a problem in this pipeline. This could be an error of the segmentation algorithm. Also, the transform might not be able to map the data into a common space such that the comparison does not look at the same anatomical positions. Therefore, the (absolute) difference here can also be seen as an error.

6.4 Data for Evaluation

For the main quantitative evaluation, $6 \times 6\,mm$ 200×200 A-scan 3D-OCT volumes of both the ONH and the macula region were acquired using software modified

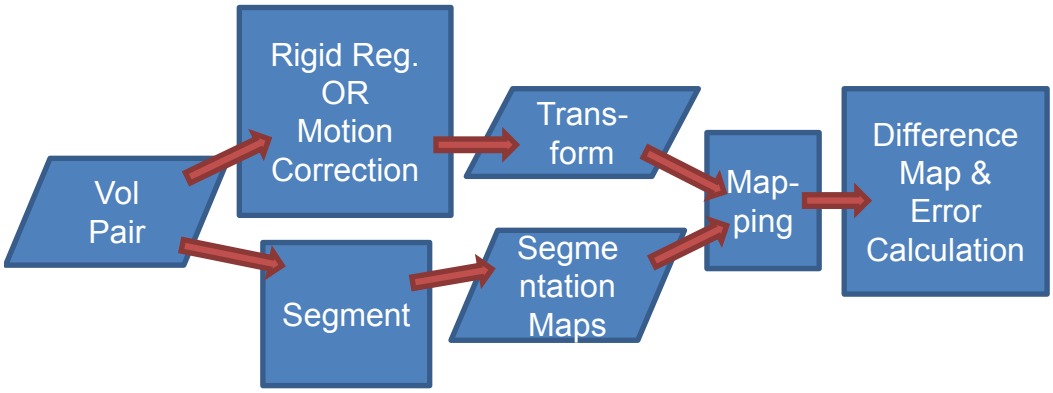

Figure 6.4: Segmentation based quantitative evaluation work flow.

Optovue RTVue devices at New England Eye Center and the University of Pittsburgh Medical Center. The software modification enabled the scanning of YFAST type volumes in addition to the standard XFAST type raster scan. The axial pixel spacing of the devices was 3.1 μm/pixel in tissue in axial direction and 30 μm/pixel in the transverse directions. The study protocol was approved by the Investigational Review Boards of the New England Medical Center, University of Pittsburgh Medical Center and the Massachusetts Institute of Technology. Written informed consent was obtained from all subjects before imaging. The study was approved by the ethics boards of the involved institutions.

For each subject, one eye was chosen at random for imaging. Each subject was imaged three times at two scan regions centered at the macula and ONH, respectively. Each time a set of two orthogonally scanned volumes was acquired. Between repetitions, the device was reset and the subject re-aligned to the device. In some of the subjects, multiple volumes were acquired per scan region and volume type. In this case, the first volume that did not have blinks and where the retina was mostly pertained within the imaging range in axial direction was chosen as the input volume. In general, subjects were instructed to fixate on the internal fixation target, no artificial motion artifacts were induced. With 73 subjects being imaged, 876 input volumes were acquired in total and used as input for the evaluation.

A subject qualified as a normal subject if they had a normal Humphrey 24-2 visual field, intraocular pressure (IOP) at or below 21 mmHg, no history of diabetes, no family history of glaucoma and a normal ocular examination. Glaucoma suspect eyes were defined as those with IOP at between 22 and 34 mmHg, asymmetrical ONH cupping or an abnormal appearing OHN, all in the presence of normal visual field test results. The contralateral "healthy" eye of an unilateral glaucomatous eye was defined as glaucoma suspect. This subgroup includes eyes that may manifest ocular hypertension, increased cupping or asymmetrical cupping. The third group of eyes, namely glaucomatous eyes was defined as those with at least one of the following features: Glaucomatous visual field defect, IOP > 35 mmHg in the presence of ONH cupping or a nerve fiber layer defect on biomicroscopy. Table 6.1 shows statistics about the study population. Three Subjects of the 73 were

excluded from the evaluation due to data acquisition problems or extremely bad volume quality due to blinking, retina out of range, etc.

Group	Count	Age
Normal Subjects	38	$39.9 \pm 16.4 (21 - 74)$
Glaucoma Subjects	26	$63.0 \pm 10.7 (33 - 83)$
Glaucoma Suspects	6	$63.0 \pm 10.6 (33 - 83)$
Excluded	3	$61.7 \pm 9.0 (51 - 73)$
Total	73	$52 \pm 19.0 (21 - 89)$

Table 6.1: Evaluation data set population statistics.

6.5 Algorithm Profiles Evaluated

In the course of the evaluation performing no motion correction is contrasted with two different profiles of the correction method. These two profiles are denoted ”basic” and ”advanced” and correspond approximately to two different stages in the evolution of the method. The ”basic” profile corresponds approximately to the method of the initial journal publication [Krau 12]. The ”advanced” profile represents the latest version of the method [Krau 14]. The main differences between these two profiles can be found in the use of different loss functions for similarity measure and regularization, the use of two stage registration and tilt compensation and illumination correction. Table 6.2 shows a comparison of the two profiles.

Note that the second pass of the advanced profile matches the first pass of the basic profile, especially in the budget of function evaluations that may maximally be used per multi-resolution level. In addition, the advanced profile employs a first pass that performs rough axial alignment and tilt compensation per B-scan. Compared to the second pass, the function evaluation budget is set relatively low here. This is in order to save on algorithm run-time. Also, it is justified by the notion that the number of degrees of freedom when using the per B-scan Axial+Tilt parametrization is much less than when using a fully 3D per A-scan parametrization. Therefore, the optimization problem can be seen as being easier which means that a good solution can likely be found in much fewer iterations in the optimizer. This is in accordance with practical observations.

In addition to the two algorithm profiles that are evaluated and contrasted with performing no motion correction, the influence of the regularization factor α is also evaluated. Table 6.3 shows the different values that were tested in the evaluation. Several other parameters that control the behavior of the algorithm are set to common values for all of the configurations tested. They are shown in table 6.4. These parameters were empirically set based on the manual inspection of a small subset of the data. In total, two different profiles are evaluated, each with five different settings for α, resulting in ten different configurations being evaluated. For 70 subjects, 6 volume pairs per subject, 5 α settings and two compared methods, $4200 = 70 \times 6 \times 5 \times 2$ motion corrected and merged volumes were generated as input for the subsequent analysis. Experiments were performed on a Core i7-2600k

Profile	Basic	Advanced
Illumination Correction:	No	Yes
Similarity Measure Loss Function:	Square (L_2)	Pseudo-Huber (L_{H,ϵ_H})
Regularizer Loss Function:	Square (L_2)	Pseudo $L_{0.5}$ ($L_{0.5,\epsilon_{0.5}}$)
Two Pass Optimization:	No	Yes
Tilt Compensation:	No	Yes
Regularizer out-of-plane scale s_{OOP}	16	16
First Pass:		
Axial Downscaling N_{down}:	1	2
Median Filtering (Size):	2D (3×3)	1D (7)
Number of Multi-Resolution Levels N_{pyr}:	5	5
Parametrization:	Per-A-scan 3D P^{dir}	Per B-scan Axial+Tilt P^{bat}
Evaluation Budget (lowest to highest res) $N_{ev,m}$:	20000, 10000,5000,250,80	50,40,30,20,10
Second Pass:		
Axial Downscaling N_{down}:	-	1
Median Filtering (Size):	-	2D (3×3)
Number of Multi-Resolution Levels N_{pyr}:	-	5
Parametrization:	-	Per-B-scan 3D P^{b} (lowest resolution level) Per-A-scan 3D P^{dir} (other levels)
Evaluation Budget (lowest to highest res) $N_{ev,m}$:	-	20000, 10000,5000,250,80

Table 6.2: Setting comparison between basic and advanced algorithm profiles.

Regularizer weighting α
0.001
0.01
0.1
1.0
10.0

Table 6.3: Regularizer weighting factors evaluated.

Description	Symbol	Value
Mean Displacement Term Factor	β	1.0
Tilt Data Term Factor	γ	1.0

Table 6.4: Common algorithm parameter values.

CPU with an NVIDIA GeForce GTX 580 GPU and 16 GB of RAM running C++ and CUDA code, respectively.

It would be desirable to also evaluate the effect of different settings in additional parameter dimensions, such as those shown in table 6.4. However, given the large-scale nature of the evaluation, it is not feasible to also evaluate the effect of additional parameters due to the combinatorical explosion of results that need to be generated and evaluated.

6.6 Evaluation Components

Several auxiliary techniques are used within the evaluation, which will be described in the following subsections.

6.6.1 Quasi-Rigid Registration

In order to account for alignment related effects between successively acquired volumes, quasi-rigid registration is employed. The quasi-rigid registration is performed by fixing one volume and transforming the other in 3D. The sum of the pseudo Huber loss function L_{H,ϵ_H} applied to the difference of image intensity between the volumes is minimized for the parameters of the transform . The transform is parametrized by translation in all three directions, rotation around z and a tilt in x and y direction. The tilt parameters model a linear shift in axial position as a function of x and y, respectively (see section 3.4.5). This is used to model a tilt in the volumes that appears when the beam passes through a different position on the pupil plane. Because of these tilt parameters, the registration is strictly speaking not rigid, but affine. In the scenario of OCT imaging of the retina we deem this transform to be able to model the alignment related changes (global translation, tilt, rotation around the optical axis) between two volumes in a correct way given the scenario which is why we call it quasi-rigid [Krau 12].

In the concrete implementation of the quasi-rigid registration, multi-resolution optimization as well as illumination correction for pre-processing are employed.

6.6.2 Layer Segmentation

In addition to the abstract similarity measure of mutual information, layer segmentation is employed to assess similarity in the quantitative evaluation. To evaluate performance over the whole transverse field it is helpful to use two dimensional segmentation maps which associate a measurement with every A-scan of the respective volume. The segmentation component itself should be as reliable as possible. Reliability in this context means that segmentation errors need to be minimized. At the same time, the segmentation needs to be spatially accurate on a per A-scan basis.

Based on these considerations, layer segmentation was performed using an algorithm based on Chiu et al. [Chiu 10] which segmented the positions of inner limiting membrane (ILM) , inner segments (IS) and RPE and based on these, retinal thickness (defined to be $RPE - ILM$). In order to use additional knowledge and improve robustness, Chiu et al.'s algorithm was extended so that multiple layers could be found in a single Dijkstra shortest path search. Whereas in Chiu's algorithm a node in the graph corresponds to a single pixel in the B-scan (x, z) we extend this concept such that a node corresponds to a combination of two or more axial positions within an A-scan of the B-scan $(x, z_1, z_2, ..., z_n)$ with $z_1 < \ldots < z_n$. This enables a node to model multiple layer positions at the same time. A path through the graph then denotes the segmentation of multiple layers within the B-scan. Connections between the nodes of this "multi-graph" are based on the combination of the possible transitions between individual layer positions from one A-scan to the next. Additionally, transitions that lead to layer positions that are too close or too far apart are discarded. The corresponding connection costs are the sum of the axial gradient based costs of the individual layers.

Compared to the original approach by Chiu et al., the graphs are much larger with this method because of the combinatorial explosion of the possible number of nodes. Therefore, special care had to be taken in order to optimize graph construction and shortest path search. The number of nodes per A-scan is $\mathcal{O}(N_a^{N_l})$, where N_a is the number of possible distinct axial positions on which a layer can be located and N_l is the number of layers that are searched simultaneously. In order to keep the graph size within reasonable limits N_a had to be restricted by either limitation of the axial search space or only considering every n-th axial position, or both.

In the concrete case, segmentation was performed individually for each B-scan of each volume in the following sequence: First, candidate locations for the ILM and RPE were found simultaneously using the multi-graph concept and considering every fifth axial pixel. Then, the RPE estimate was low pass filtered using a Gaussian filter and a reference layer was defined to be slightly above (17 pixels) this line. The B-scan was then shifted in the axial direction such that the reference layer would become a straight line. Subsequently, the IS and RPE layer positions were searched for simultaneously below the reference line. The ILM was searched for as an individual layer above said reference line. By applying this segmentation algorithm to every B-scan in a volume, 2D en face layer position and thickness maps for ILM, IS, RPE and retinal thickness were obtained. In order to enable

maximum spatial resolution of the resulting maps, no further smoothing was performed. Figure 6.5 shows an example application of the segmentation step on a motion corrected and merged ONH volume.

To evaluate the algorithm, a manual segmentation study was performed. Three human graders segmented 158 randomly chosen B-scans from the available data. Comparison with the automatic algorithm showed the mean absolute difference between human observers to be only slightly lower than between human and automatic observers (ILM: 2.32 pixels vs. 2.96 pixels, RPE: 2.96 pixels vs. 3.48 pixels).

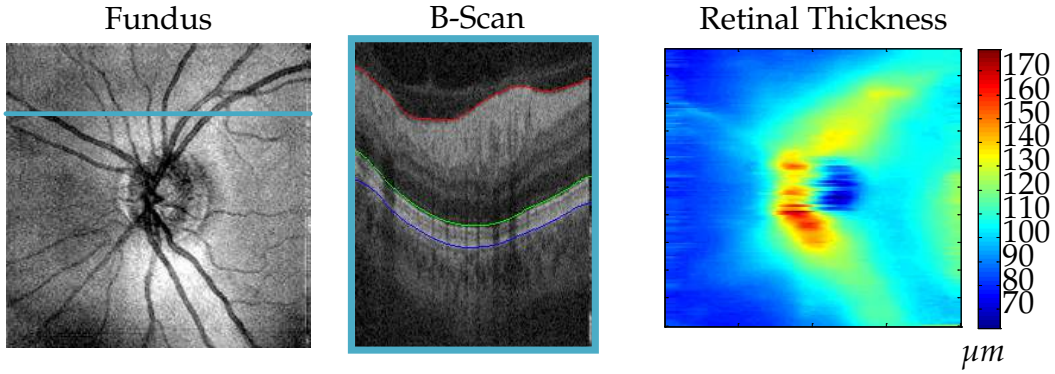

Figure 6.5: Layer segmentation example of a merged ONH dataset. Left to right: Volume fundus projection, segmented B-scan corresponding to the line in the fundus projection and the obtained retinal thickness map. The red, green and blue lines mark the ILM, IS and RPE boundaries that were segmented, respectively.

In addition, NFL thickness was automatically segmented per A-scan on all volumes with another method based on adaptive thresholding [Gabr 07, Ishi 06]. When comparing segmentation maps obtained this way, the difference is sensitive to transverse distortion between the volumes for all maps. Furthermore, the individual layer maps (ILM, IS, RPE) are also sensitive to axial distortion while retinal and NFL thickness are not.

6.6.3 Blood Vessel Segmentation

As an additional type of measurement that is sensitive to transverse distortions, blood vessel likelihood maps were generated from the volume data. Based on the retinal layer segmentation, the mean intensity between the IS and RPE layer was plotted into a 2D map. This map was subsequently illumination corrected by applying a bias field that was obtained using a large standard deviation Gaussian filter (20 pixels) [Hou 06]. In addition, the map was scaled such that the median intensity was 0.5. Subsequently, a Hessian based multi-scale vesselness measure was applied [Fran 98], yielding a 2D vessel likelihood map. The maps were thresholded and scaled such that the maximum likelihood of 1.0 was reached when the vesselness response was at or above 0.0001. Figure 6.6 shows an example of applying the blood vessel map generation method to three OCT volumes of the same eye and location. The volumes show slight signal loss in the top left corner due to part of the retina being outside of the axial imaging range. Nevertheless, the method is able to successfully extract blood vessel maps from the volumes.

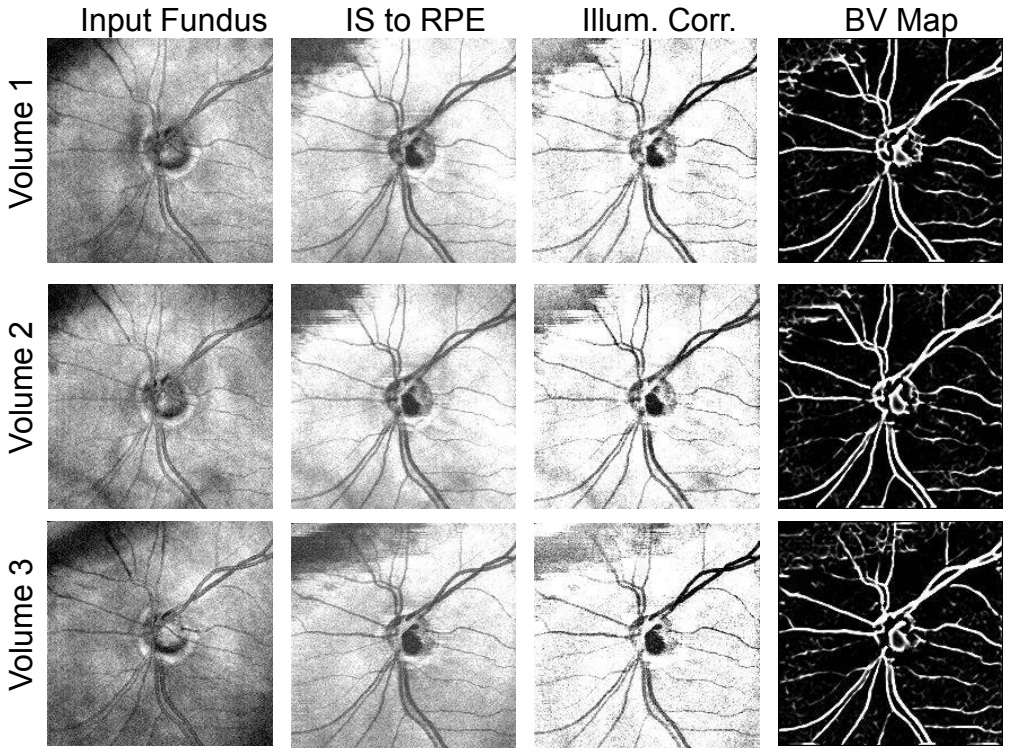

Figure 6.6: Example of blood vessel map generation for three corresponding uncorrected 3D-OCT volumes. Top to bottom: Maps corresponding to three XFAST input volumes from one subject. Left to right: Input fundus view, average projection from IS to RPE layer, illumination corrected average projection and last the resulting blood vessel maps.

6.6.4 Segmentation Map Mapping

As part of the evaluation, measurements in the form of 2D per A-scan segmentation maps from two volumes are mapped into a common space and then compared. The maps are produced by the layer and blood vessel segmentation steps.

Transforming a segmentation map is achieved by offset interpolation. The offsets at every grid position are given by the transform. For the interpolation itself, bi-linear interpolation is used. If the type of segmented quantity corresponds to a 3D position, such as is the case for segmented layer positions of the ILM or RPE, the axial component of the transform is used to offset the value accordingly. On the other hand, if the quantity is not affected by an axial translation applied to the volume, only 2D offset interpolation is needed for mapping. This is the case for thickness measurements of the retina and nerve fiber layer as well as for blood vessel maps.

6.6.5 Difference Map Computation

Given two segmentation maps that have been mapped into a common space, the evaluation is concerned with how similar the measurements contained within the two maps are. For this purpose, absolute difference maps are computed. For every pixel within an area of interest, excluding a border of 10 pixels, i.e. 5 percent of the transverse area on each side in these mapped segmentation maps, several

absolute difference maps were calculated. The border was introduced to account for a slight lack of overlap of acquired volumes due to changes in fixation or motion. The absolute difference maps simulate a varying tolerance to an uncertainty in lateral position in the maps. This is achieved by taking the minimum absolute difference of a reference pixel in the first map $a(x, y)$ and the values in a neighborhood $n(x, y, tol)$ of $\pm tol$ pixels in each direction around the corresponding position in the second map such that

$$\text{absdiff}_{tol}(x, y) = \min_{(x_2,y_2)\in n(x,y,tol)} |a(x, y) - b(x_2, y_2)|. \tag{6.2}$$

Values of the positional tolerance of 0, 1 and 2 pixels (corresponding to a tolerance of up to $\pm 90\mu m$) were evaluated, where $tol = 0$ is the standard absolute difference operation. Figure 6.7 shows a schematic of the difference computation in relation to the spatial tolerance.

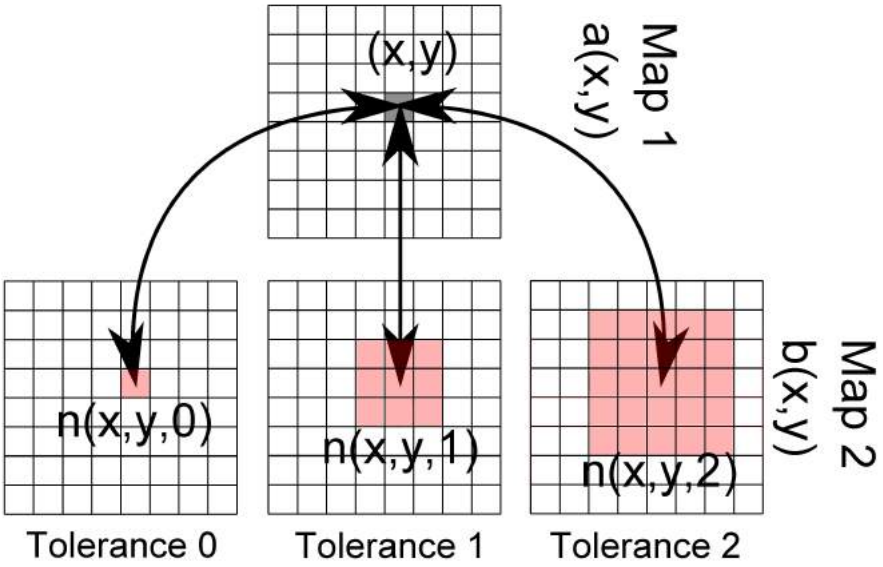

Figure 6.7: Spatial tolerance schematic. The minimum absolute difference between a value in map 1 and a neighborhood in map 2 is calculated. The size of the neighborhood depends on the spatial tolerance.

Finally, the mean and median values of each absolute difference map were calculated. Lower values corresponded to lower error in segmentation maps obtained from the volumes and are therefore indicative of better registration of the volumes and/or higher similarity between the volumes and/or better segmentation performance.

6.6.6 Q-Metric

The Q-metric is computed based on the singular value decomposition of the local image gradient matrix on selected anisotropic patches in an image[Xian 10]. Anisotropic patches are patches in which there is a dominant gradient direction and are automatically selected from the image based on statistical testing. The underlying assumption is that such patches contain structure such as edges. Noise or

blurring of the image would lower the anisotropy of the gradient in such patches. Therefore, the amount of patches that were selected together with their measured anisotropy contribute to the metric. For the concrete evaluation, a MATLAB implementation of the metric was used (`http://users.soe.ucsc.edu/~xzhu/doc/metricq.html`). The patch size was set to 8×8 pixels. The number of patches was not fixed but was set based on the statistical detection of anisotropy. Patches were selected individually for every 2D input image.

For quantitatively evaluating signal quality of a given volume, Q is evaluated in central 2D cross-sections in the X-Z and the Y-Z plane. The mean of Q over the these cross-sections of the input volumes, the registered volumes and the merged output volumes is computed and compared. A change in Q from the registered to the merged volumes is indicative of a change in image quality (see section 6.2.3).

6.7 Significance Testing

Within the quantitative evaluation, measurements of similarity and signal quality are computed over different sets of volumes and for different settings and correction methods (no correction, basic correction, advanced correction). Subsequently, quantities such as the mean and standard deviation of these individual results are computed over all volumes in question. In order to check whether there is a significant difference between the quantities between different subgroups of the available data and/or for different methods, non-parametric statistical significance testing is employed. For comparing errors where there was a pairing between data sets, i.e. when comparing the different methods for the same set of input volumes, a Wilcoxon signed rank test (significance level 0.01) was used to check whether the distributions were significantly different. For independent sets of input volumes, e.g. when comparing different subgroups of the population in the study for differences, the Mann-Whitney-U test was employed.

6.8 Summary

In this chapter, a detailed description of the methodology used to evaluate the motion correction was given. A key issue in evaluating the algorithm is that the registration operates without a fixed reference image. Therefore, it is not sufficient to show that anatomical features are registered to each other to prove motion correction success. Instead, a three pronged approach is used. First, the aforementioned ability to register volumes onto each other is evaluated. Second, reproducibility of the produced output volumes is evaluated to assess whether the results are reliable. Third, signal improvement is evaluated using automatic and quantitative methods.

In order to evaluate the first two criteria, quantitative similarity measures between OCT volumes are used. On the one hand, these consists of the information theoretic measure of mutual information. On the other hand, segmentation of features of the volumes are compared for similarity.

A large body of 3D-OCT data acquired from two different clinical sites is available for evaluation. In order to facilitate reproducibility evaluation, three pairs of orthogonal volumes of each eye and location were acquired. The study population consists of normal subjects, glaucomatous subjects and glaucoma suspects.

Two different motion correction algorithm profiles are evaluated. These correspond to two different steps in the evolution of the method. The main differences between these two profiles can be found in the use of different loss functions for similarity measure and regularization, the use of two stage registration and tilt compensation and illumination correction.

Several auxiliary techniques are required in order to perform the evaluation. These include quasi-rigid registration, retinal layer and blood vessel segmentation, the mapping of segmentation maps and the generation absolute difference maps of segmentation maps. Furthermore, the Q-metric which is used as part of the evaluation of signal improvement is described. Finally, as part of the evaluation, quantitative measurements are tested for statistical significance.

CHAPTER 7

Results and Discussion

In this chapter, results of the evaluation according to the methodology described in chapter 6 are presented. The results are separated according to the aspects of registration performance, reproducibility performance, signal quality and run time. Where applicable, visual inspection is used to give an intuition of the results, followed by different quantitative measures. Also, some cases of algorithm artifacts and their potential reasons are shown. Subsequently, the results are discussed. Parts of this chapter have been published in prior publications of the author [Krau 12, Krau 14].

7.1 Registration Performance

The following subsections are concerned with evaluating the registration performance, i.e. how similar the volumes are after motion correction and registration and how well motion artifacts are corrected.

7.1.1 Visual Inspection

In order to visually examine the evaluation pipeline and the effect of advanced motion correction on volume data reliability, two pairs of XFAST volumes and their corresponding YFAST volumes from a random subject was selected. Figure 7.1 shows fundus images of the two pairs. The volumes are from the ONH region of a glaucomatous subject. As indicated by the red arrows, all input volumes contain saccadic transverse motion artifacts.

Figure 7.2 shows merged fundus projections using the first volume pair shown in figure 7.1 as input data. Results for both the basic and advanced correction algorithm are shown for all tested α settings. Several observations can be made here.

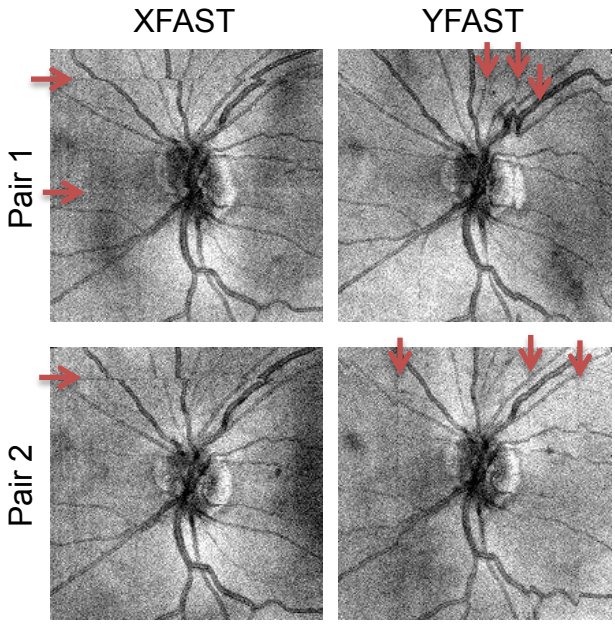

Figure 7.1: Fundus views of pairs two Optic Nerve Head input volumes. Red arrows indicate motion artifacts.

First, for the lowest amount of regularization ($\alpha = 0.001$) and especially for the basic algorithm, a significant amount of distortion can be seen in the merged fundus. This can be explained as follows. Very low regularization hardly restricts the modeled displacements to conform to the time structure of the OCT sampling process. The primary goal of the optimization becomes maximizing similarity. Therefore, while features are registered well to each other, the displacement fields of the solution might model unrealistic motion and lead to distorted output volumes.

On the other hand, on the high end of the regularization spectrum ($\alpha = 1.0$ and $\alpha = 10$) the amount of motion that can be modeled is very much restricted by the regularization. This causes the optimization to fail to register anatomical structures such as blood vessels onto each other. In the corresponding merged fundus projections of figure 7.2 this shows up as duplicate vessels.

In the middle of the α range, especially $\alpha = 0.01$, both methods are able to register most anatomical locations onto each other without causing an apparent distortion of the volumes. Compared to the input fundus views in figure 7.1 there are no obvious motion artifacts visible. For this single case, visual inspection indicates that the advanced algorithm leads to a more stable result over the α range (see especially $\alpha = 0.001$).

Based on visual inspection of this case, it is not obvious which algorithm produces the better results. Besides a single example not being sufficient, this points to the general problem that registration performance alone is not sufficient for evaluating the performance of the algorithm. This is due to there being no reference volume within the registration and no ground truth available. We will address this issue in the second part of this evaluation (section 7.2).

Figure 7.3 shows composite views of the central slices along y-direction from the same volumes as are shown in figure 7.2. Here, the XFAST volume is shown in

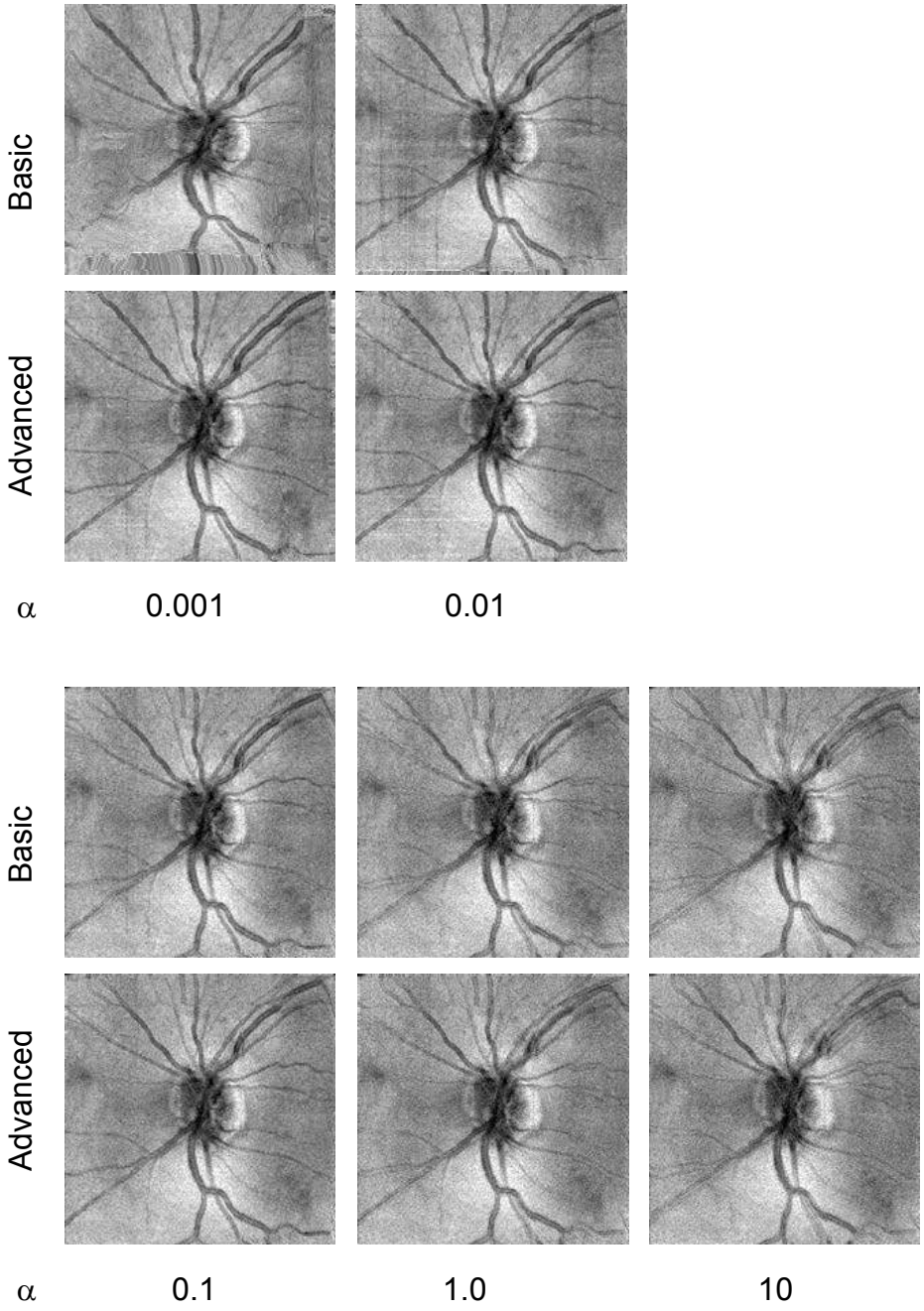

Figure 7.2: Comparison of merged fundus views for the two methods and different α.

red color while the YFAST volume is shown in the green. Therefore, a good registration between the two volumes will show as yellow color, while misalignment between features can be seen as green or red areas in the images. It can be seen that the input volumes are not aligned at all. All output slices show that the large-scale axial misalignment between the input volumes is removed. As α increases, the alignment of small details gets worse, leading to significant mis-registration

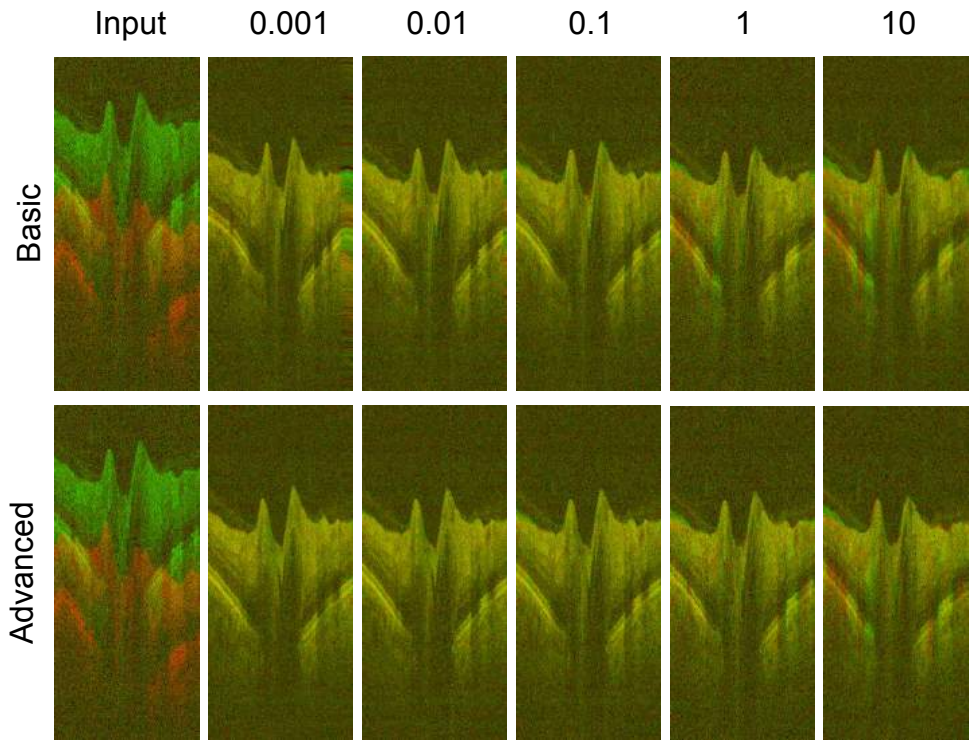

Figure 7.3: Comparison of composite images of central slices along the y-direction for the two methods and different α.

of features for $\alpha = 1$ and $\alpha = 10$. This is consistent with what can be seen in the fundus views from figure 7.2.

7.1.2 Mutual Information

The first quantitative measure of registration performance is the difference of MI through motion correction and registration. MI is calculated between the two registered volumes and between the two input volumes and subtracted to yield the difference. Subsequently, for each of the two correction profiles, the mean MI difference is calculated over the set of available XFAST/YFAST pairs.

Figure 7.4 shows this mean increase over all data, for the two methods and for the different α values that were evaluated. It can be seen that the mean MI over all data steadily decreases as α increases. This can be explained by the balance of similarity measure and regularization. As regularization strength is increased, volume similarity becomes relatively less important leading to lower similarity of the volumes after motion correction. In addition, the advanced algorithm leads to higher similarity than the basic algorithm, regardless of α. This difference was found to be significant for all α.

7.1.3 Segmentation Measures

Figure 7.5 shows the mean absolute segmentation map error over all data for three maps (retinal thickness, ILM and blood vessels). Here, each XFAST and YFAST

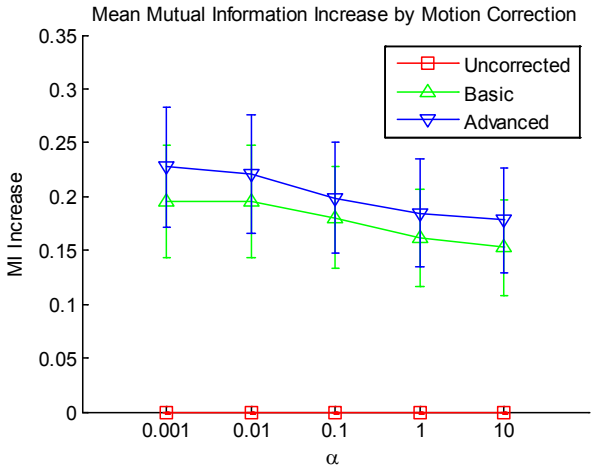

Figure 7.4: Comparison of information theoretic measures of registration performance. Mean mutual information increase through registration over all registered data sets for different for different regularization strength for no correction, basic algorithm and advanced algorithm. Dashed lines of corresponding color indicate $\pm$ one standard deviation around the mean.

input volume has been segmented. The segmented maps have then been mapped to the common motion corrected space via the the displacement fields that were found during motion correction. The figure shows the results for different α and for the three compared methods. For all α, the advanced algorithm leads to the lowest errors followed by the basic algorithm. Performing no correction at all leads to the highest error between the segmentation maps, indicating low similarity of the underlying input volumes. Significance testing using a signed rank test revealed that for each α and each possible pairing of three methods in question, the values were significantly different. Consistent with the visual inspection results, it can be seen that as α increases, the segmentation based registration error tends to increase slightly.

Also, it is notable that the ILM position error is much more affected by performing any motion correction than the other two map types. This can be explained with the ILM position being the only map type that is sensitive to axial motion. Furthermore, the axial motion and misalignment related error seems to dominate the ILM position error. Performing any motion correction will very likely lead to a coarse alignment of the volumes in the axial direction (see figure 7.3). This leads to a more severe reduction in error compared to the other two map types. Nevertheless, the advanced algorithm still leads to a lower ILM position error. As with all the other map types and for each α, this difference was found to be statistically significant.

7.2 Reproducibility Performance

This section is focused on assessing the reproducibility performance of the different settings and methods on the available data. Fundamentally, the question is how well the motion correction output volumes from independently acquired

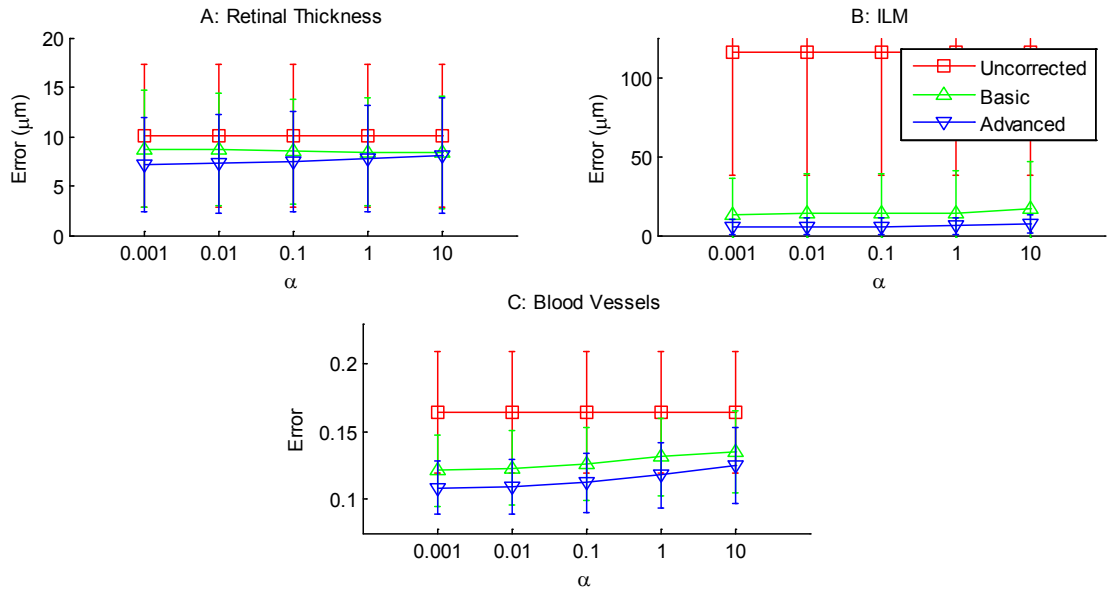

Figure 7.5: Comparison of the mean of mean absolute segmentation map errors between XFAST and YFAST volumes for different α and the three different methods. Error bars indicate $\pm$ one standard deviation around the mean. A: Retinal thickness error. B: ILM position error. C: Blood vessel map error.

input volume pairs of the same area correspond to each other, how similar they are.

7.2.1 Visual Inspection

In order to get a visual indication of the reproducibility performance at least two independent pairs of orthogonal input volumes of the same region have to be used. Visual inspection will therefore use both pairs of volumes shown in figure 7.1. First, each pair of input volumes is corrected using the different algorithm/α combinations. In order to correct for alignment related differences between the volumes, a quasi-rigid registration step is then performed between the possible pairs of output volumes that resulted from processing the independent pairs of volumes from the same subject and area.

Figure 7.6 shows central cross-sections of the volumes resulting from applying the different algorithms to the two pairs of volumes from figure 7.1. Composite slices before and after quasi-rigid registration are shown for the three methods and for $\alpha = 0.1$. Compared to the uncorrected result, the rigidly registered motion corrected volumes from both correction algorithms correspond much better. This indicates that the motion correction improves the consistency of volume data. Visual inspection also reveals that for these particular volumes, the advanced algorithm leads to better similarity after quasi-rigid registration. This can be observed in particular when looking at the blood vessel shadows and comparing the left of the two cross-sections between the two methods. In addition, application of the advanced algorithm leads to the retina being aligned more horizontally than in the corresponding slices for no and basic correction. This can be explained with

the additional tilt compensation that is performed only in the advanced algorithm (see section 5.5.10).

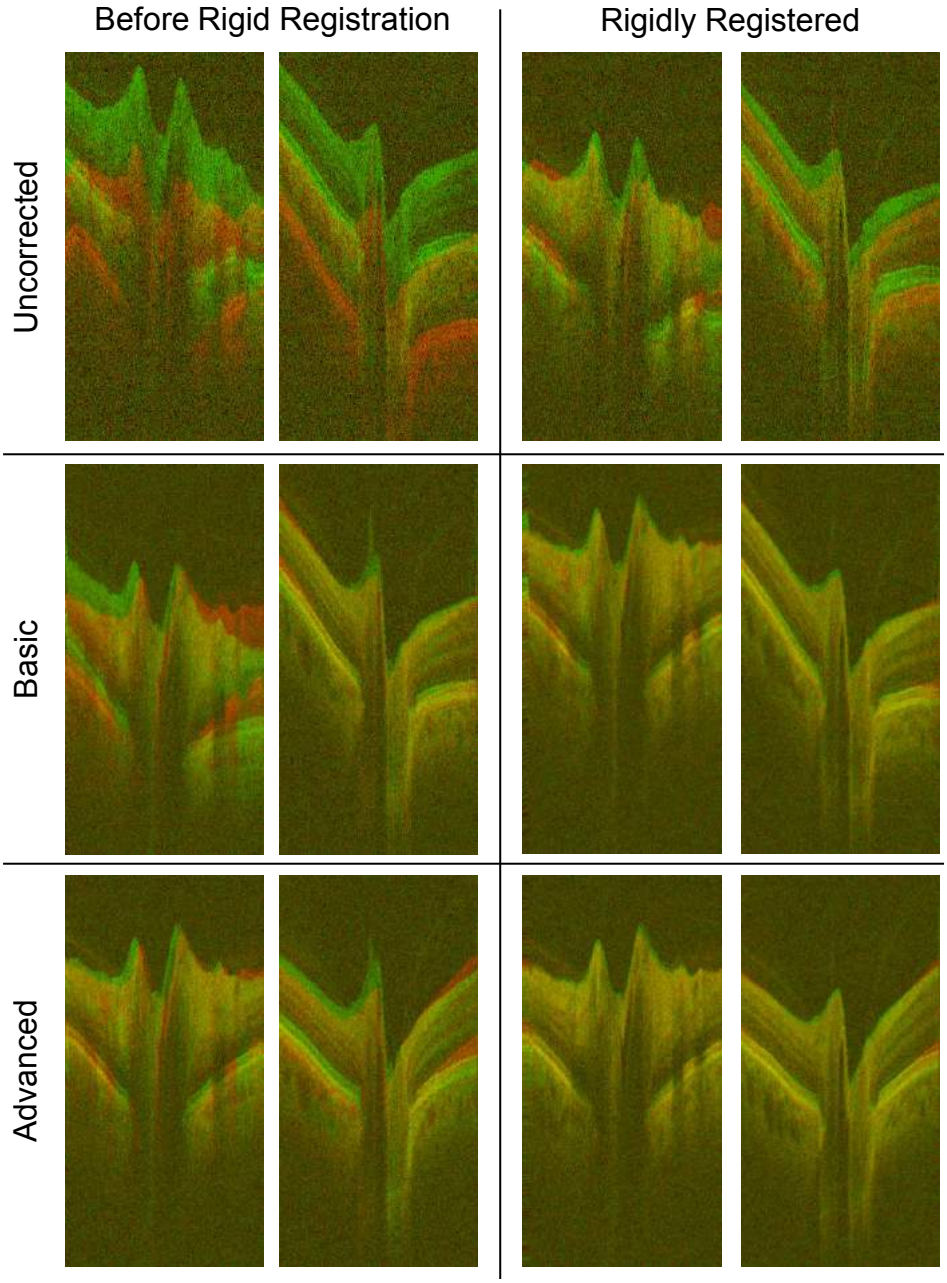

Figure 7.6: Example comparison of quasi-rigid registration performance between uncorrected and motion corrected data with $\alpha = 0.1$.

After segmentation of the output volumes and quasi-rigid registration, the quasi-rigid transform obtained is used to map the segmentation maps into a common coordinate system. Subsequently, absolute difference maps are calculated (see equation (6.2)). Figure 7.7 and figure 7.8 show the individual transformed segmentation maps for blood vessels and retinal thickness and the corresponding difference maps for uncorrected, basic and advanced corrected data for $\alpha = 0.1$.

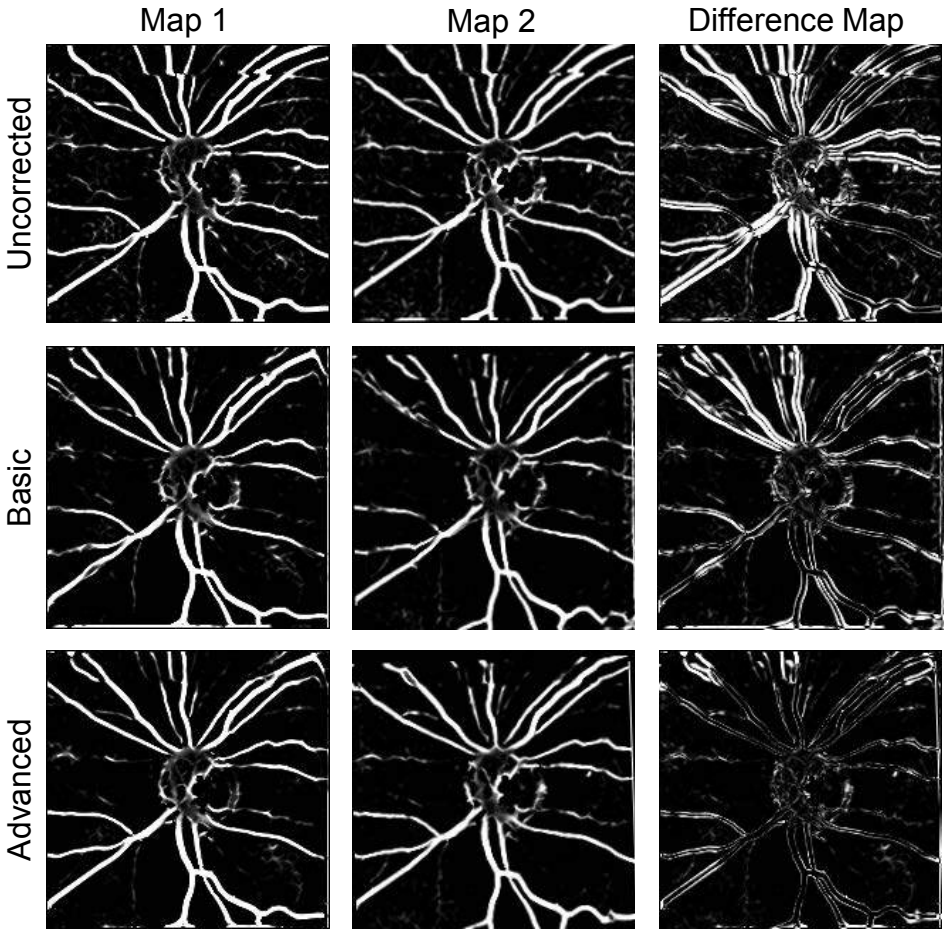

Figure 7.7: Example comparison between mapped blood vessel maps and the corresponding difference maps.

In figure 7.7, it can be seen that the blood vessel likelihood map segmentation works well for all output methods of the different methods. The segmentation corresponds well to the OCT fundus images from the same volumes. For the case of performing no correction (top row) these are the two XFAST fundus images in figure 7.1. For basic and advanced correction they can be found in figure 7.2 under $\alpha = 0.1$. In the absolute difference maps, a mismatch of the blood vessel positions can be seen as bright areas, corresponding to a large absolute difference in blood vessel likelihood. The uncorrected case shows the largest amount of mismatch. Here, the mismatch shows up characteristically as a double vessel pattern. For the basic algorithm the mismatch is significantly reduced overall. Double patterns are only visible in the top left and center part. In the other areas, the mismatch areas are at the edge of the vessels. This indicates that here the mismatch is less than the diameter of the respective vessels. Finally, the difference map for the advanced method shows no double patterns except for the very edge of the transverse area. Also, overall the amount of mismatch seems lowest. The numbers of the mean absolute difference (see section 6.6.5) also support this: It is 0.19 for the uncorrected case, 0.11 for the basic correction algorithm and 0.06 for the pair resulting from using the advanced correction algorithm.

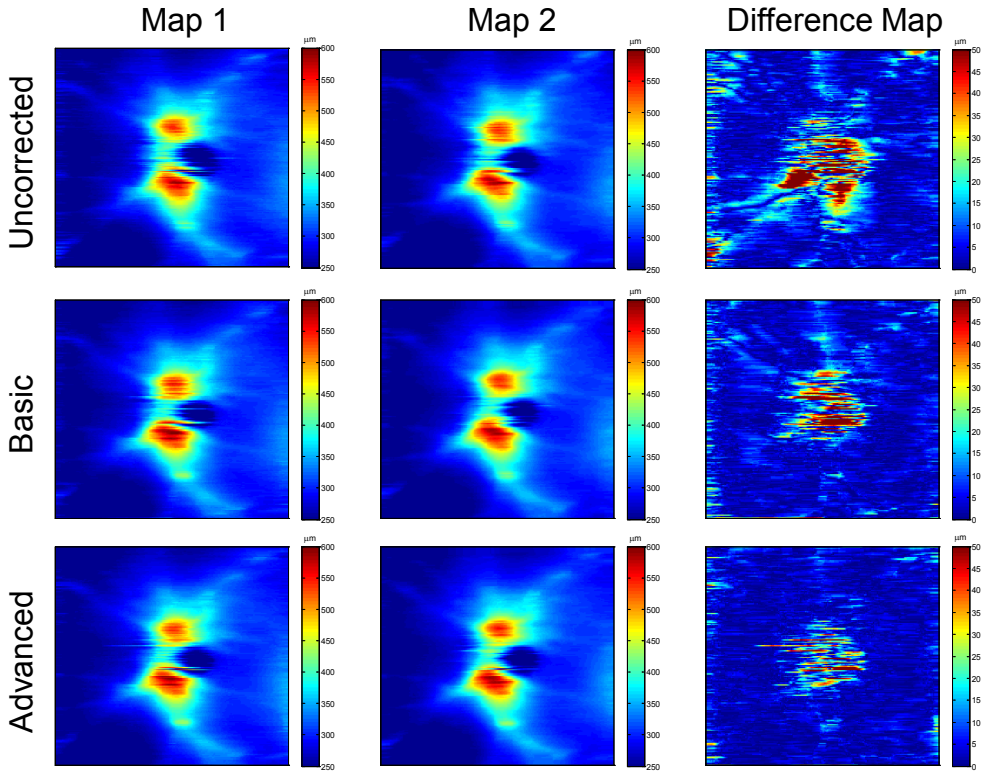

Figure 7.8: Example comparison between mapped retinal thickness maps and the corresponding difference maps.

Figure 7.8 provides a similar comparison for segmented retinal thickness. The segmentation looks to be able to segment all the volumes reasonably well. Segmentation errors can mainly be observed around the cup of the ONH itself. Importantly, there is no apparent bias in the value of the segmented thickness among the different methods. The absolute difference or error maps follow a similar trend as in the blood vessel case. It can be observed that the highest errors occur directly at the ONH. This can be explained with the aforementioned segmentation errors. Retinal thickness is most variable around the ONH. Therefore, small mismatches between anatomical locations will result in a comparatively large error. Also, the overall error is largest in the uncorrected case. Again, basic correction leads to a reduction. As before, advanced correction leads to the apparently lowest error. The numbers of the mean absolute difference in this case are 9.32 μm for no correction, 6.17 μm for basic and 4.28 μm for the advanced correction algorithm.

Overall, visual inspection indicates that motion correction, especially using the advanced algorithm, leads to a higher reproducibility of the volume data and the resulting segmentation maps for the set of volumes that were inspected. The following sections cover different quantitative measures of reproducibility.

7.2.2 Mutual Information

Figure 7.9 shows the mean mutual information after quasi-rigid registration of all possible corresponding pairs of output volumes for the three methods. Both mo-

tion correction algorithms lead to higher similarity after quasi-rigid registration compared to applying no correction. In addition, the advanced algorithm shows consistently higher similarity compared to the basic algorithm. Both of these differences were found to be significant regardless of α.

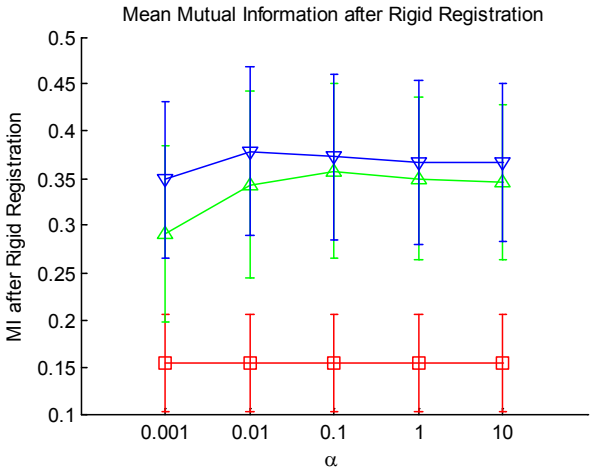

Figure 7.9: Comparison of information theoretic measures of reproducibility performance. Mean mutual information after quasi-rigid registration of pairs of disjoint results of result volumes for no correction (red), basic algorithm (green) and advanced algorithm (blue). Error bars indicate $\pm$ one standard deviation around the mean.

7.2.3 Segmentation Measures

Figure 7.10 compares the mean absolute error over all pairs of data sets between the three methods for different segmentation maps. Lower numbers indicate better reproducibility of the values contained in the respective maps. In addition, retinal thickness and blood vessel likelihood map errors are not sensitive to axial motion artifacts, while ILM and RPE position are. Applying no correction consistently leads to the largest error for all four types of maps. The advanced algorithm tends to produce the lowest errors, followed by the basic algorithm.

Considering the quantitative results for both registration and reproducibility performance in conjunction, $\alpha = 0.1$ can be considered an optimal parameter setting. Especially the blood vessel likelihood map reproducibility errors and their dependence on α give an indication of this. Here, and for other measures $\alpha = 0.1$ produces the lowest overall errors for the advanced algorithm without the choice disadvantaging the basic algorithm. Therefore, the remaining evaluation will be performed with α at 0.1. For this regularization strength and most others (except 0.001), reproducibility errors are significantly lower for retinal thickness, ILM, NFL thickness and blood vessel maps. The mean blood vessel map reproducibility error is reduced to 69% of the uncorrected error for the basic and to 47% of the uncorrected error for the advanced algorithm, which was found to be statistically significant.

7.2.4 Sub-group Analysis

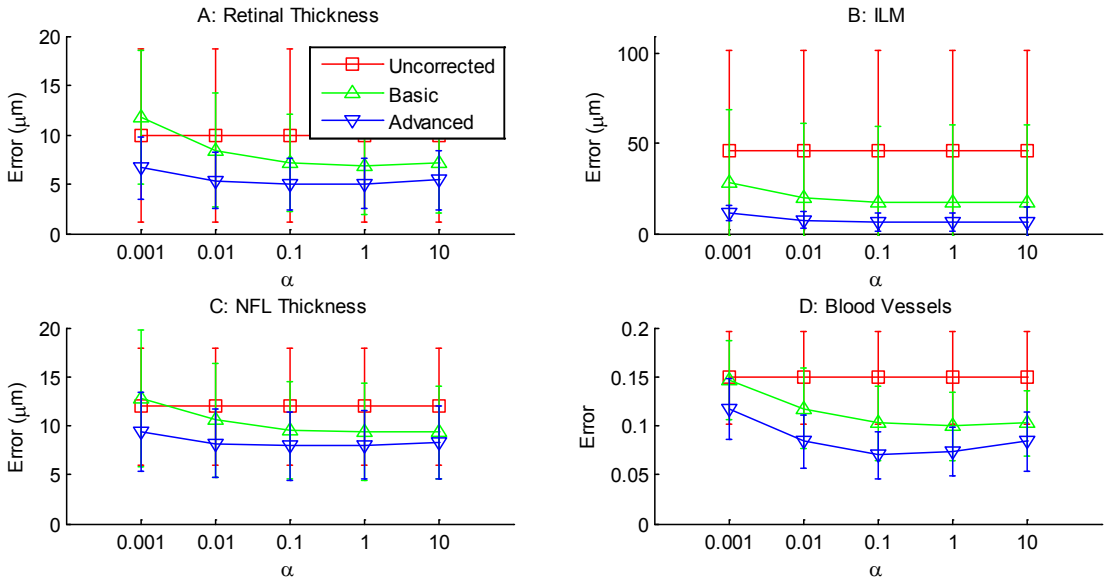

Figure 7.10: Comparison of the mean of mean absolute segmentation map errors between all possible pairs of output volumes from one scan region for the three methods and different α. Error bars indicate $\pm$ one standard deviation around the mean. A: Retinal thickness error. B: ILM position error. C: NFL thickness error. D: Blood vessel map error.

The previous results considered the mean reproducibility error over all available data; however, it is also interesting to look at potential differences in reproducibility performance for subgroups of the data. Figure 7.11 shows a box plot comparison between the errors over all data sets versus all normal and versus glaucomatous subjects and glaucoma suspects combined. For all groups and the four map types, the advanced algorithm always has the lowest errors, followed by the basic algorithm and no correction. In addition, the normal subject group shows slightly lower errors than the combined glaucoma suspect and glaucoma groups.

Figure 7.12 shows the same type of box plots, but this time grouped according to anatomical location into all data sets, ONH only and macula only. Again, the advanced correction shows best reproducibility performance. In addition, the macula subgroup shows lower errors. This can be explained by the fact that the area around the ONH has both more variability in retinal thickness and contains more blood vessels. Also, NFL thickness varies more around the ONH and the NFL thickness segmentation algorithm is not very reliable within the optic disc. For the same amount of transverse distortion, this leads to higher errors for mismatched maps. Using advanced correction, the mean reproducibility error consistently drops below two axial pixels (6.2 μm) for retinal thickness for all subgroups in the two figures.

Finally, figure 7.13 compares the reproducibility errors over all data sets for different spatial uncertainty tolerance values *tol* of 0, 1 and 2. As expected, a higher spatial tolerance consistently leads to lower errors, when other parameters

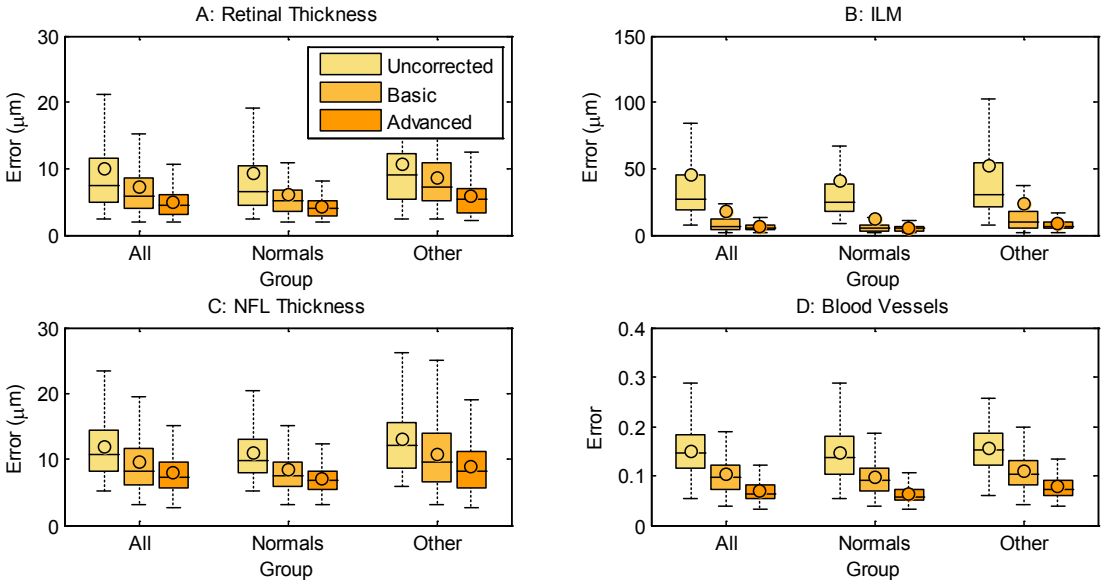

Figure 7.11: Box plot segmentation error comparison between Normal and Other Subjects (glaucoma + glaucoma suspects) for the three methods and $\alpha = 0.1$. A: Retinal thickness. B: ILM position. C: NFL Thickness. D: Blood vessel maps.

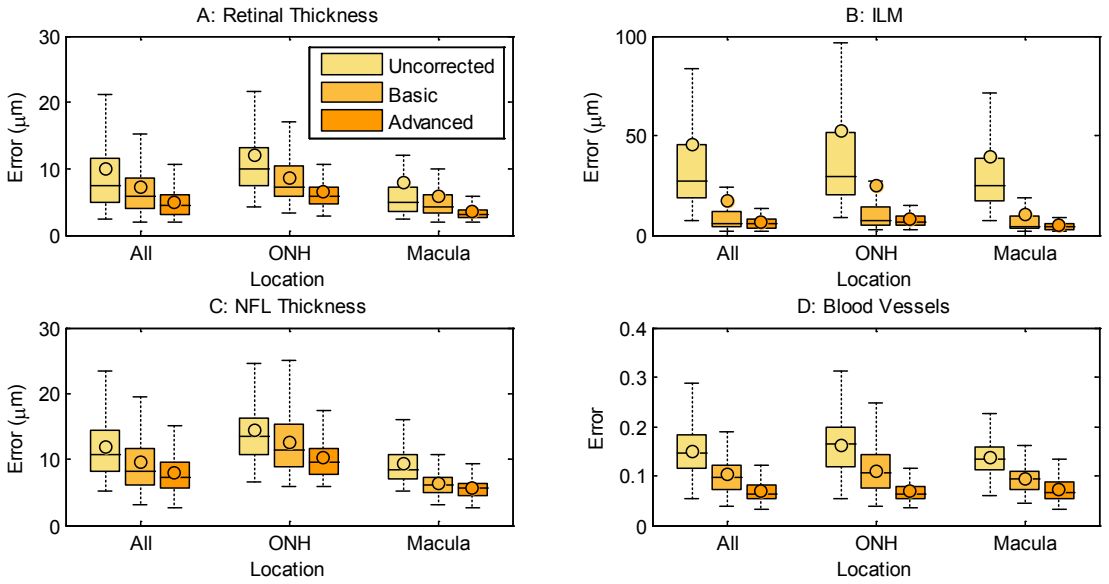

Figure 7.12: Box plot segmentation error comparison between for ONH and Macula region volumes for the three methods. A: Retinal thickness and $\alpha = 0.1$. B: ILM position. C: NFL Thickness. D: Blood vessel maps.

are kept fixed. In addition, regardless of the tolerance value, the advanced algorithm leads to the lowest error, followed by the basic algorithm, with performing no correction being the worst.

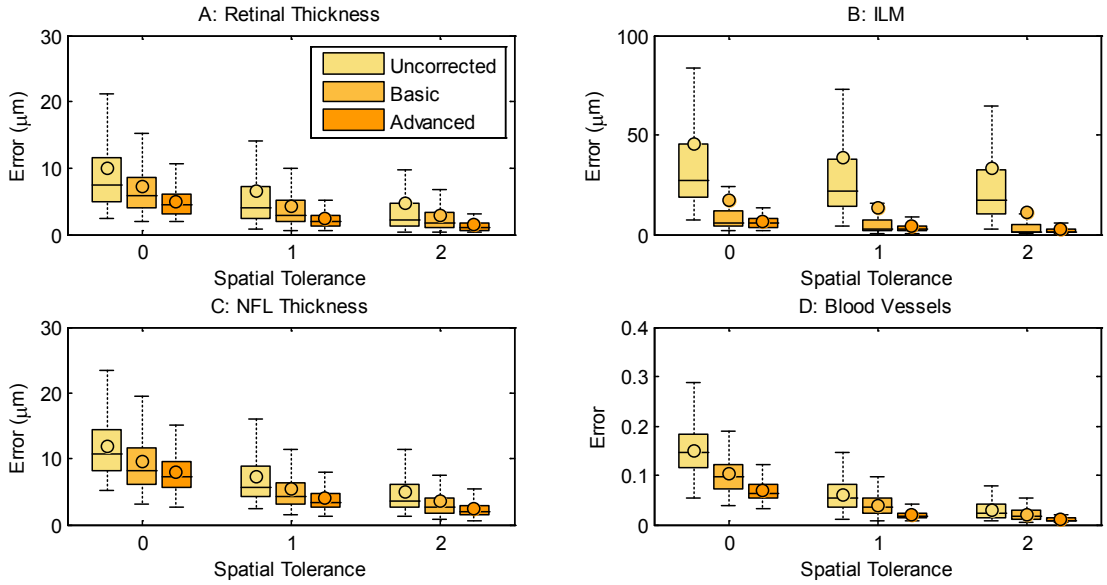

Figure 7.13: Box plot segmentation error comparison for different spatial uncertainty tolerances *tol* for the three methods and $\alpha = 0.1$. A: Retinal thickness. B: ILM position. C: NFL Thickness. D: Blood vessel maps.

7.3 Signal Quality

Since multiple volumes are registered to each other in the motion correction approach there is the opportunity to improve signal quality by merging the registered volumes. In order for the merged volume to be not blurred, the volumes need to be registered well. Then, speckle noise can be reduced without a loss in apparent resolution of the images. As with the other aspects that were evaluated visual inspection is used to give an intuition first, followed by quantitative evaluation based on the Q-metric (see section 6.6.6).

7.3.1 Visual Inspection

Figure 7.14 shows an example of signal quality change through motion correction and merging for three example data sets. One of them (top row) was processed using the basic algorithm, the other two were processed using the advanced algorithm. In all cases α was 0.1. In the columns, from left to right, central cross-sections along the x-direction are shown from the registered XFAST volume, the registered YFAST volume and the merged volume. The three examples were picked to show a range of outcomes regarding Q-metric and signal quality change. In the first example which shows a data set from the macula region, the registered YFAST data set shows an artifact that is marked with a red arrow. Here, the same data from the input volume was repeated multiple times to generate the registered result. This shows as a kind of repeating pattern in the image. Correspondingly, because there is no real data at this place and because the two slices are not well registered, the merged slice is blurry and of low quality. Looking at the Q-Metric, the mean Q of the (four) registered central slices is 29.9. Q for the merged cross-

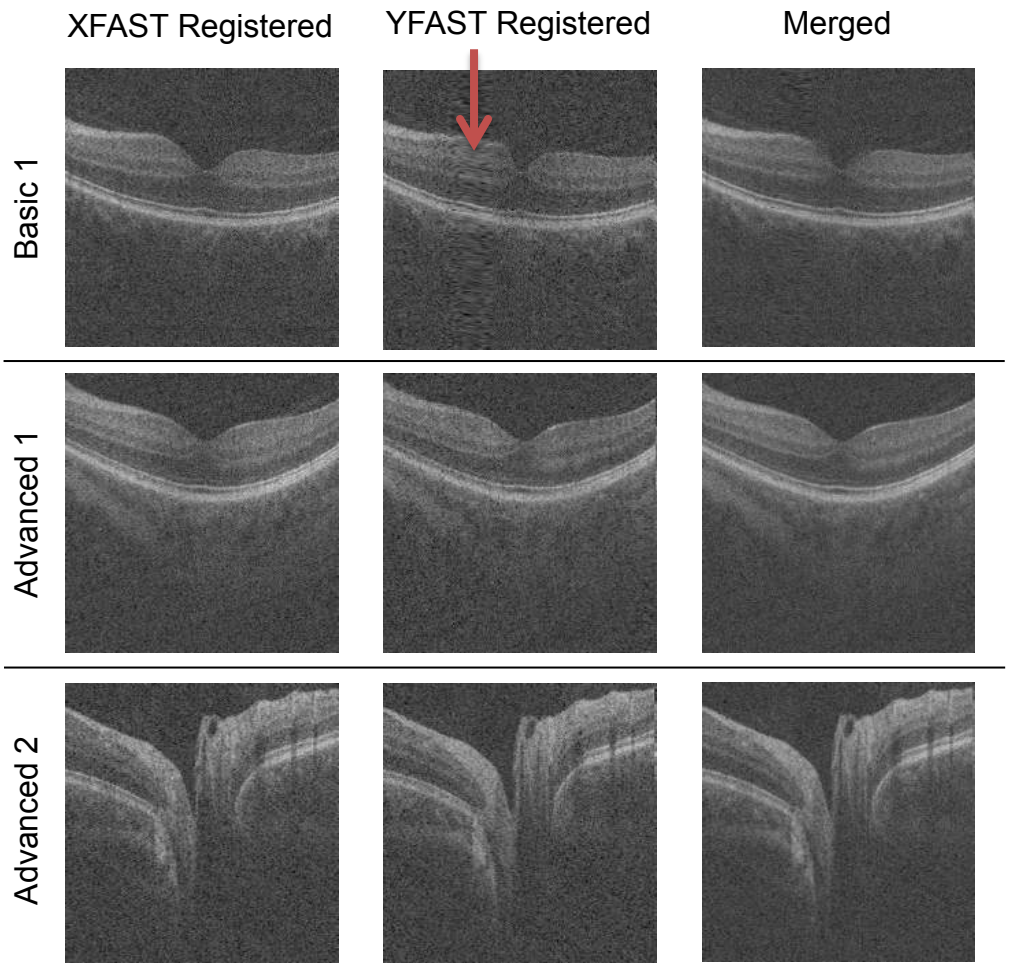

Figure 7.14: Signal quality change through merging shown in three example data sets.

sections regresses to 12.9. Relative to the registered Q, the merged Q decreases by 57 percent in this case.

On the other hand, in the second and the third case no obvious artifacts are visible. Correspondingly, the Q-Metric shows a 31 percent relative improvement in the first advanced case and a 44 percent relative improvement of Q for the second advanced case. Because the retinas in the registered slices are better registered with each other the merged slice shows no obvious blurring of edges. At the same time, the amount of speckle noise that is visible is decreased. This makes it easier for example to discern retinal layers in the image by visual inspection. Correspondingly, it would be expected that automatic segmentation is also be improved given a higher signal quality input.

7.3.2 Q-Metric Quantitative Results

Figure 7.15 shows different mean Q-Metrics over all data sets and for different α. Again, no correction, basic and advanced correction are compared. Subfigure A shows the mean Q calculated on the registered slices, before merging. B shows mean Q over the merged central slices. Finally, C shows a combination of A and B

that is the mean relative change, relative to the Q value associated with the registered slices. For the uncorrected case the mean Q of the input volumes is shown. Again, the error bars mark $\pm$ one standard deviation around the mean.

In subfigure A, it can be seen that the registered Q of both correction methods is significantly higher than Q calculated on the input volumes. In addition, the mean Q for the advanced algorithm is more constant over α. Also, the standard deviation is lower than for the basic case.

In subfigure B, it can be seen that regardless of α, the Q for the advanced algorithm is higher than for the basic algorithm. This difference was statistically significant for all α. Also, the curves resemble the mutual information measure of registration performance (see Figure 7.4). As is the case there, Q tends to decrease with increasing α.

In subfigure C, the mean relative change of Q from registered to merged is shown. Note that this measure is not independent of the mean Q from the registered data and the mean Q of the merged data from subfigure A and B. In fact, the mean relative improvement in Q is calculated from these measures. As can be seen in subfigure C, the highest improvements can be observed for the advanced algorithm. For $\alpha = 0.1$, the advanced algorithm lead to a 28 percent mean relative improvement in Q-metric compared to 9 percent improvement for the basic algorithm. The differences between the three methods were found to be statistically significant for each α. Also, the standard deviations of the distribution for the advanced algorithm are lower than for the basic algorithm, indicating higher consistency of the results over the body of data of the study. The mean relative change of Q also decreases with increasing α, at least for the advanced algorithm. It is not clear why the mean relative change peaks at $\alpha = 1$ for the basic algorithm, although at a lower level than is the case for the advanced algorithm. One possibility is that registration artifacts in the basic algorithm for low α lead to increased image sharpness, leading to increased Q.

7.4 Run Time

The time window between acquisition and analysis of OCT data that is allocated for motion correction needs to be kept short in order to not interfere with the clinical work flow. Therefore, algorithm run-time is a concern. To reduce run-time, parts of pre-processing and of the evaluation of the objective function and its gradient have been implemented on a GPU using CUDA (see section 5.9).

Table 7.1 shows a comparison of the run time for different input volume sizes and number of input volumes. The basic and the advanced algorithm run times are in seconds and are shown with using the GPU via CUDA (see section 5.9) and using the CPU only with some multi-threading via OpenMP. The complete run time including reading input volumes was measured. No whole registered or merged volume were written. However, during the measured time there was some diagnostic output generated. In order to minimize the time spent reading the volume the algorithm was invoked once before actually measuring the run time in order to cache the volume data.

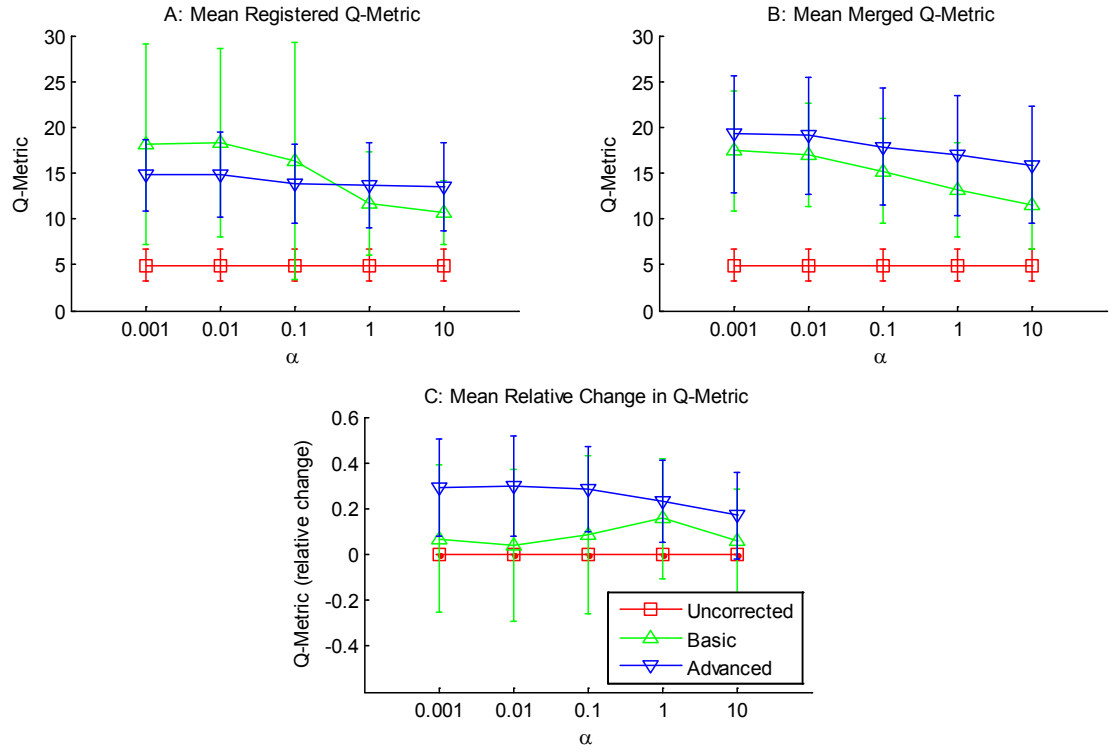

Figure 7.15: Quantitive evaluation of signal quality using Q-Metric. A: Mean Q-Metric over all data of registered slices. B: Mean Q-Metric over all data of merged slices. C: Mean relative change in Q-Metric from registered to merged slices.

The numbers indicate that using the GPU leads to an acceleration factor of approximately three. Also, approximately linear scaling with the total number of voxels in all input volumes can be observed. However, it can also be seen that the advanced correction algorithm is about twice as slow as the basic algorithm. This can be explained by the two stages containing two optimizations, preceded each by pre-processing. Also, two sets of output volumes need to be constructed, once at the end of the first stage and once at the end of the second stage. The optimization in the first stage is allotted much less function evaluations than the second. However, the first stage needs to compute the axial histogram based tilt compensation term, which is currently always implemented in software and therefore rather expensive. In addition, the computation of $L_{0.5,\epsilon_{0.5}}$ and its derivative used in the advanced algorithm are much more expensive to evaluate than the square loss used in the basic one.

7.5 Artifacts

While the quantitative results show clear improvements, in some cases the algorithm produces artifacts. These can be divided into mis-registrations which show up for example as doubled blood vessels and distortions of the registered volumes. Figure 7.16 shows four examples that were picked to showcase artifacts that are generated in certain cases. As such they are not representative for the vast

Dimensions				Times (s)			
				Basic		Advanced	
Width	Height	Depth	#Volumes	GPU	CPU	GPU	CPU
200	200	768	2	23	85	56	154
200	200	640	2	18	84	47	136
200	200	768	4	49	168	114	298
200	200	640	4	48	178	107	258
400	400	1000	2	124	507	274	702
400	400	1000	4	313	982	627	1553

Table 7.1: Algorithm Run-time comparison. Times are in seconds.

majority of volumes that were part of this evaluation. All cases were processed with the advanced profile and $\alpha = 0.1$. The first case shows an ONH volume with significant saccadic motion artifacts in both input volumes. In the lower part of the merged volume the algorithm is not able to register the volume correctly. This leads to a double vessel pattern. The second case shows a pair of input volumes from a macular region with very low signal and inconsistent illumination, likely caused by cataract. As can be seen it is very hard to see features in the input fundus projections. In addition to low signal, the the XFAST input volume also shows saccadic motion in the bottom part. The merged projection shows interpolation artifacts at the corresponding location. In the third case the YFAST input volume was affected by a blink at the right edge, causing total loss of signal. Overall, the algorithm is able to recover from this problem. However, the merged projection shows a narrow black vertical bar where the blink occurred. This is because the merging step used data from the blink region there. The fourth example shows two effects: First, the YFAST volume shows signal loss in the top right corner. This is the result of the retina moving outside of the axial imaging area of the OCT system. Secondly, there is a slight rotation around the optical axis between the two input volumes. The resulting merged volume shows clear artifacts in the area of signal loss. In general, features seem not to be registered well to each other, leading to a blurring of the result fundus projection. This could be caused by the rotation around the optical axis to which the regularization of the algorithm is not adapted.

Based on these four examples and general experience with the algorithm the following features of the input volumes can be identified as potentially causing artifacts.

- Low OCT signal: Caused by cataracts, bad eye optics or a badly aligned OCT system.
- Selective total loss of signal: Caused by blinks or by motion outside of the axial range of the system. These cause inconsistencies where the OCT for the same anatomical location shows the retina in one input and no signal at all in the other.
- Other signal inconsistencies: Signal between corresponding locations can also be inconsistent due to OCT system alignment(illumination) and moving

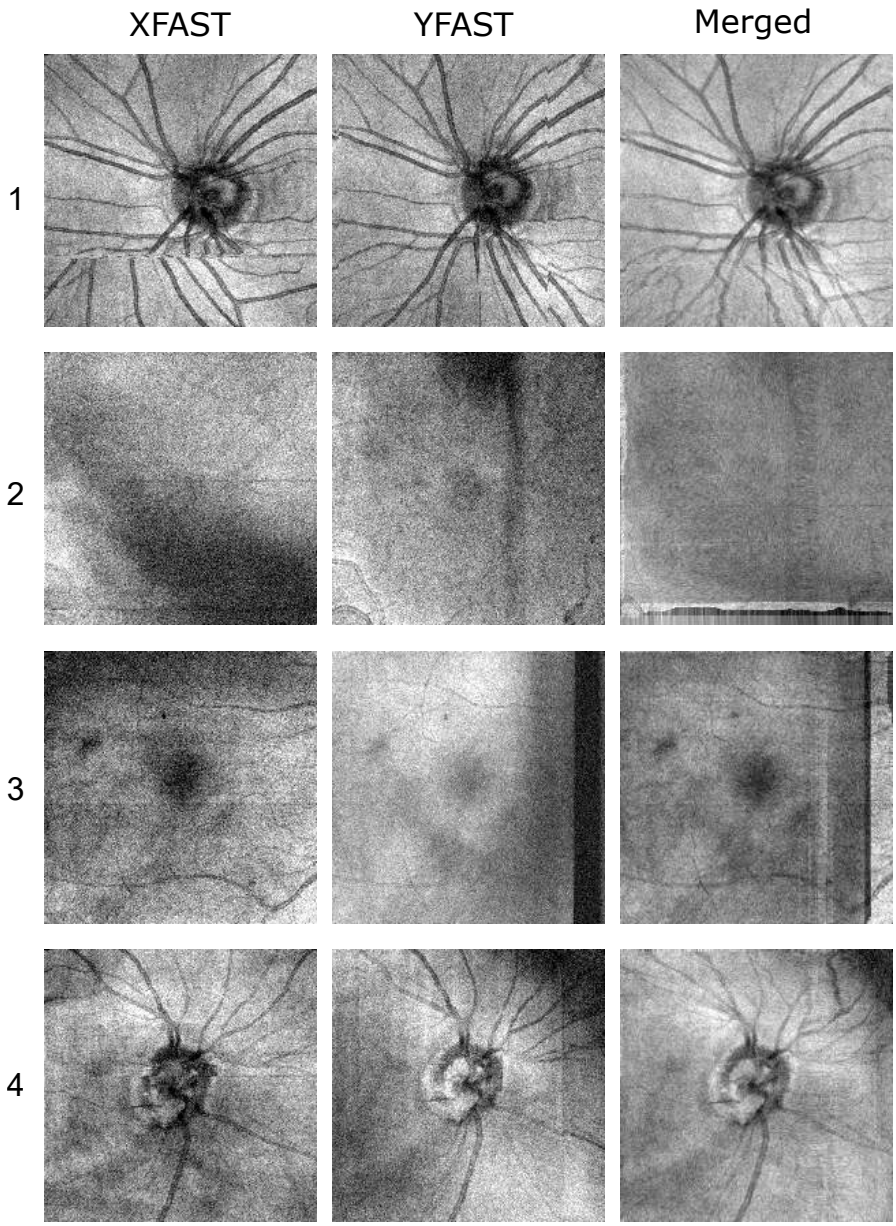

Figure 7.16: Fundus views from four cases where the algorithm produced artifacts. Left column: Fundus projections of the XFAST input volume. Middle column: Fundus projections of YFAST input volume. Right column: Fundus projection of the merged volume.

floaters for example. Illumination correction within the motion correction algorithm tries to alleviate this problem. Sometimes however, recovery is not possible.

- High motion: When the motion level is very high the regularizer might not allow for this much motion to be modeled. This leads to mis-registration of features. The lower the overall OCT signal level, the worse this problem becomes.

- Unsampled areas: Depending on the concrete motion patterns during acquisition of the input volumes it is possible that certain areas of the region of interest on the retina have not been sampled at all in the input volumes. In this case the algorithm cannot find the correct area from the inputs to sample from to create the output volumes. Therefore, output voxels have to be wrong and usually show up as interpolation artifacts where the same area is repeated multiple times.

The development of the advanced algorithm already caused much better robustness, leading to less cases which exhibit artifacts. Future work should be concerned with reducing artifacts even further.

7.6 Discussion

The results show that the advanced algorithm yields significant improvements in the obvious reduction of motion artifacts, the improvement of signal quality and in obtaining reliable quantitative measurements from 3D-OCT volume data. Whereas the basic algorithm already shows significantly lower errors than performing no correction, the advanced algorithm yields even further improvement. It is important to keep in mind that the errors measured in this evaluation are associated with the combined reproducibility of the entire processing pipeline. This pipeline includes the OCT device and its axial and transverse resolution and sampling, its SNR, the presence of motion artifacts in the data, the performance of the motion correction and merging algorithm, layer and blood vessel segmentation and quasi-rigid registration performance. These components also interact with each other. For example, good motion correction counteracts motion artifacts in the input data. In addition, good motion correction and registration followed by merging will improve SNR. High SNR is important for segmentation algorithm performance. Conversely, poor motion correction can introduce additional non-reproducible distortions in the volume. In addition, if the motion correction algorithm fails to register the volumes, the merged volume would have lower SNR and resolution since image information from different anatomical locations would be combined. Based on these interactions, it can be hypothesized that advanced motion correction plays an especially important role. It effectively corrects for motion artifacts and improves SNR compared to the input volume, leading to better performance of the subsequent steps and subsequently the highest reproducibility.

In addition, there is an inherent trade off between the precision of the segmentation and its reproducibility. By applying low pass filtering to the segmentation maps prior to difference map calculation, the reproducibility error would be reduced. However, the segmentation would not be as precise, i.e. losing its ability to capture focal changes. In this evaluation, the aim was to have the segmentation itself be as spatially precise as possible, in part also to be more sensitive to distortions caused by motion artifacts. The spatial uncertainty tolerance was introduced to be able to assess how the reproducibility error would decrease when the requirements for a precise per A-scan segmentation are decreased. In this context, the results show that even for the largest spatial tolerances, when the minimum of

a grid of pixel differences is taken to compute difference maps, advanced motion correction still leads to the best results. One factor that could explain this result is the improved SNR of the merged data. Also, the motion induced distortions in uncorrected data which lead to reproducibility errors might in part be larger than what the spatial tolerance allows for.

The ability of the merging step being able to improve signal quality is directly related to how well volumes are registered by the motion correction step. Therefore, the quantitative measures of signal quality using the Q-metric follow a similar pattern as the measures for registration performance (see section 7.1). Again, the advanced algorithm leads to better results than the basic algorithm. Given that the Q-metric is fully automatic and operates without a reference, one has to be careful not to over interpret the results. However, it is consistent with the other results and the Q-metric numbers that the advanced algorithm also leads to the highest signal quality in the merged volumes. Because of the aforementioned interactions, it is likely that this improved signal quality also contributes to lower the segmentation errors, leading to higher reproducibility performance.

The results regarding algorithm run time indicate that in the current implementation the advanced algorithm has a higher run time cost than the basic algorithm. This can be explained with additional steps that have to be performed. However, it is also true that the current code base makes certain choices in its implementation that might cause more slowdown for the advanced algorithm than is necessary. Therefore, the gap in run time between advanced and basic algorithm can likely be decreased in an optimized implementation. Overall though, the results show that using a GPU, the run time is acceptable for small volume sizes which currently are the clinical standard. With an optimized implementation and as GPU hardware progresses, it is expected that the run time requirements will not pose a serious problem for clinical adoption. In fact, an optimized version of the algorithm is already successfully being used clinically in the Optovue RTVue XR.

7.7 Summary

In this chapter, the method proposed in chapter 5 was evaluated using the methodology described in chapter 6. The evaluation is divided into registration performance, reproducibility performance, signal quality as well as run time.

Visual inspection of registration performance was performed on a sample pair of orthogonal volumes. Merged fundus projections and composite images of the registered volumes were inspected for different α and for basic and advanced motion correction. From these examples it can be seen that an α that is too low leads to a distortion of the anatomy in the resulting volume. However, features tend to be registered well to each other. On the other end of the α spectrum, the effect is reversed. Here, anatomical features will not be registered to each other and there is little distortion. It is hard to draw further conclusions based on the visual inspection. This is due to only looking at a single case and also due to the lack of ground truth information.

For the quantitative evaluation of registration performance, the increase in mutual information through motion correction and registration was examined as a

measure for registration performance. Consistent with visual inspection, registration performance decreases as α increases. In addition, the advanced algorithm leads to significantly better registration performance than the basic algorithm, regardless of α.

As a second quantitative measure of registration performance, segmentation map errors for retinal thickness, ILM position and blood vessel maps before and after registration were evaluated. Results show that the advanced algorithm outperforms the basic algorithm and no correction for all map types, with the differences being statistically significant. Also, it was found that for axial position like maps such the ILM position, even basic correction leads to a major reduction in error, compared to the other two maps where the changes are less severe.

The second major part of the evaluation was concerned with reproducibility performance. For visual inspection two pairs of orthogonal volumes from the same location and subject were used. As part of reproducibility evaluation, a quasi-rigid registration has to be performed in order to correct for alignment differences. Visual inspection compared composite views for the two algorithms before and after this quasi-rigid registration step. Inspection showed that the rigid step could align the advanced algorithm outputs better than the basic algorithm outputs and better than the uncorrected outputs. This indicates improved reproducibility especially for advanced correction. Also, segmentation maps of quasi-rigidly aligned segmentation maps of the blood vessels and of retinal thickness of the algorithm outputs were inspected. The segmentation algorithm itself produces reasonable results. Moreover, the difference maps, which indicate reproducibility errors show the lowest errors for advanced correction, followed by basic correction, with no correction being worst.

The quantitative part of the reproducibility evaluation employed mutual information and segmentation based measures. The first measure was the mean mutual information after quasi-rigid registration over all data sets. It was highest for advanced correction, followed by basic correction with no correction leading to the worst reproducibility performance. In the evaluation of the segmentation based measures, mean absolute errors of segmentation maps after quasi-rigid registration were evaluated for four segmentation map types, three methods and different α. Significant improvements could be observed. For example for $\alpha = 0.1$, the blood vessel map reproducibility error is reduced to 69% of the uncorrected error for the basic and to 47% of the uncorrected error for the advanced algorithm. The ranking of the methods here is consistent with the mutual information results. Overall, $\alpha = 0.1$ tended to lead to the lowest errors.

Based on these results and from registration performance results, α was fixed at 0.1 for the analysis of the segmentation map results for different sub groups in the study data. Again, reproducibility errors for four map types were evaluated. Sub groups that were compared were normal subjects versus Glaucoma and Glaucoma suspects as well as ONH versus macula volumes. For all sub groups the advanced algorithm tended to perform best, followed by the basic algorithm, followed by no correction. In addition, the mean errors for certain maps were different for different sub-groups, reflecting differences between the groups.

Finally, the effect of spatial tolerance on the reproducibility error was evaluated. As expected, a higher spatial tolerance leads to lower errors. In addition, regardless of spatial tolerance, the advanced algorithm led to the lowest error, followed by basic correction, with performing no correction being worst.

Signal quality can be expected to improve when merging multiple volumes which are well registered with each other. For evaluating the effect of motion correction visual inspection and quantitative evaluation based on the Q-metric were employed. In the visual inspection of signal quality, corresponding registered as well as a combined merged cross sectional view of three example orthogonal volume pairs are shown and compared with the Q-metric for these cases. In the case where the volumes are not registered well with each other, the merged cross-section shows blurring. Also, the Q-metric decreases in the merged images, compared to the registered images. On the other hand, if the volumes are registered well, no blurring occurs and the Q-metric increases. In addition, speckle noise is decreased.

In the quantitative evaluation, the main measure was the mean relative change in the Q-metric from registered to merged slices. Again, advanced correction performed best, followed by basic correction, with the differences being statistically significant.

The run times for the two algorithm profiles were compared for different number and size of volumes. In addition, the run times of a CPU only implementation as well as GPU accelerated version was compared for each method. The results indicate an approximately linear scaling of the run time with respect to the total number of voxels in the input volumes. In addition, using a GPU leads to a speedup of approximately three. In the current implementation, the run time for the advanced algorithm is about twice as long as for the basic one.

In some cases the advanced algorithm still produces artifacts in the output volumes. Four example cases for this were inspected. Also, potential reasons for artifact were discussed.

Taken together these results indicate that the advanced algorithm yields significant improvements in obtaining reliable quantitative measurements from 3D-OCT volume data. Signal quality can also be improved. The algorithm run-time is already not prohibitive and bound to improve given an optimized implementation and advances in GPU hardware.

CHAPTER 8

Applications

The last chapter showed that the proposed motion correction algorithm improves the reliability of quantitative measurements that are extracted from 3D-OCT data. In this chapter, applications of the algorithm are described.

8.1 Algorithm Deployment

In order for the algorithm to be easily integrated into the clinical work flow, the corresponding program has to be easy to use without expert knowledge in image processing and be fully automatic. Since the sites of the collaborators are not local to the author, the need for constant interaction to perform motion correction would add a significant delay which would lower acceptance of the technique.

Therefore, a drag-and-drop front end to the command line based motion correction and merging program was developed. The concept is that input volume data (either a single slice image file of each volume or the volume files themselves if the format is one file per volume) is dragged and dropped onto the front end. The fronted will identity the input data (i.e. find all the image files belonging to one volume and the type of scan pattern based on file naming conventions) and combine this information with preset settings. These are read from an accompanying file. Then, a command line is generated and fed to the main program that will invoke the algorithm on the corresponding data. This enables very easy use of the algorithm. In addition, the configuration can still be customized according to use case and OCT device via the accompanying settings file.

This combination of main program, drag-and-drop front end and settings file was deployed at the sites of multiple collaborators. These are:

- Prof. James G. Fujimoto's group at MIT
- Prof. Wolfgang Drexler's group at the Medical University of Vienna
- New England Eye Center (NEEC) (Dr. Jay S. Duker)

- University of Pittsburgh Medical Center (UPMC) (Dr. Joel Schuman)
- Oregon Health and Science University (David Huang, MD, PhD)

In addition to being in use at several research sites the technique has also recently been commercialized. A joint patent application between the University of Erlangen-Nuremberg and MIT has been filed [Krau 11]. Optovue Inc. has incorporated it into the latest Optovue AVANTI™ RTVue XR system under the name SMART™ Motion Correction. Through this the technique will find very widespread use in the coming years.

8.2 Diagnostic Structural Imaging

The most common application for 3D-OCT is within diagnostic structural imaging in ophthalmologic practice and research. Here, the advantage of motion correction and merging is two fold. First, motion correction itself enables more reliable data, both for qualitative visual inspection and for quantitative measurements. Second, the merging of two or more motion corrected volumes improves the signal quality. As 3D-OCT volume quality from eyes that are encountered in clinical practice can be relatively low, any improvement here is appreciated.

Figure 8.1 and figure 8.2 show examples of motion correction applied to clinical 3D-OCT volumes at NEEC. Both sets of volumes were acquired on a 100 kHz A-scan rate prototype OCT system developed by MIT and using a swept source laser operating at 1060 nm central wavelength. Motion correction and merging was performed using the advanced correction method and $\alpha = 0.1$ (see section 6.5). In figure 8.1, a $3 \times 3\,mm$ transverse field was scanned with 500×500 A-scans. The subject was a 82 years old female with mild non-proliferative diabetic retinopathy and pseudoexfoliation glaucoma. Two volumes were used as input. As can be observed from the en face fundus projections of the XFAST and YFAST input volumes, there are significant transverse motion artifacts present which scramble the appearance of the ONH. The merged en face fundus projection shows that after motion correction and merging, the artifacts are removed and the ONH morphology as depicted in the volume data is intact. The composite views of slices along all three dimensions before and after registration paint a similar picture.

Figure 8.2 shows views of data that was acquired with a wide-field scan pattern. In this case a $12 \times 12\,mm$ transverse field was scanned with 500×500 A-scans. The subject was a 74 year old male exhibiting severe non-proliferative diabetic retinopathy with diabetic macula edema and also asteroid hyalosis. Compared to figure 8.1, transverse motion artifacts are less obvious in the input en face fundus projections. This is related to the fact that for a smaller scanned transverse field the same amount of motion leads to more severe relative motion artifacts than for a larger field. Nevertheless, the merged fundus projection shows that these motion artifacts are removed. The composite slice views show motion artifacts and misalignment before registration that are removed. The registered data is in very good alignment and no motion artifacts can be seen.

These results are consistent with the general experience of using the advanced motion correction algorithm, especially in conjunction with the 1060 nm based

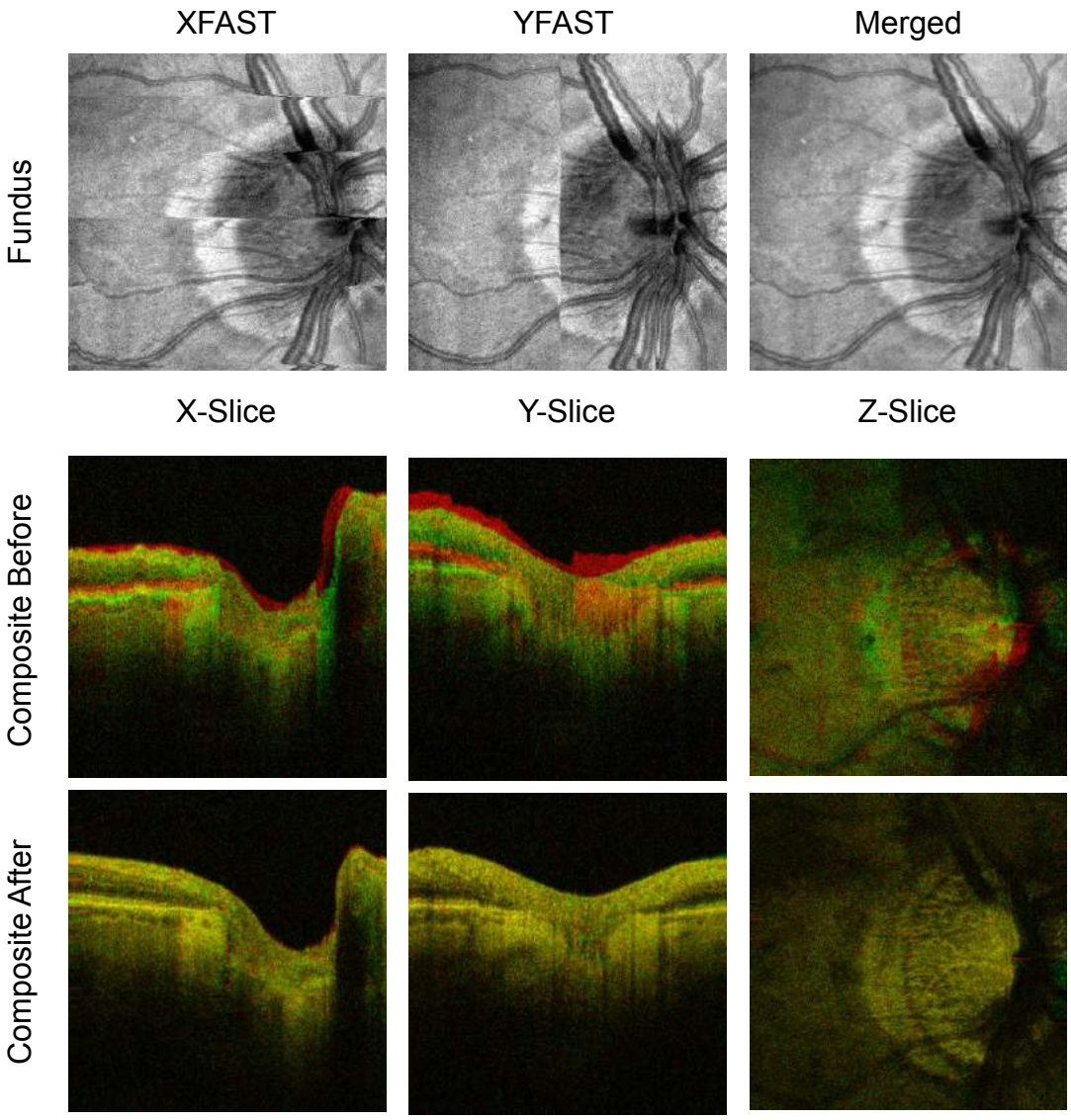

Figure 8.1: Example data showing the effect of motion correction on a 3x3 mm field.

swept source system at NEEC and UPMC. Algorithm failures that lead to a decrease in volume quality relative to the input data are rare, with there being a major reduction in motion artifacts and improvement in quality in most cases.

Within the context of the collaboration with Prof. James G. Fujimoto at MIT, Dr. Jay S. Duker at NEEC and Dr. Joel S. Schuman at UPMC the technique has been used in multiple (pre-)-clinical studies so far. Adhi et al. performed analysis on the choroid using 3D-OCT data obtained with a swept source OCT prototype system that used the motion correction technique [Adhi 14]. Ferrara et al. used the same system combined with motion correction to look at en face features of the RPE and choroid in eyes with chronic central serous chorioretinopathy [Ferr 14]. On a similar system and also using motion correction, Wang et al. at UPMC investigated the lamina cribrosa micro-architecture in healthy and glaucomatous eyes [Nadl 13, Wang 13, Wang 14a]. Furthermore, Alasil et al. used motion correction for en face imaging of the choroid in Polypoidal Choroidal Vasculopathy[Alas 15].

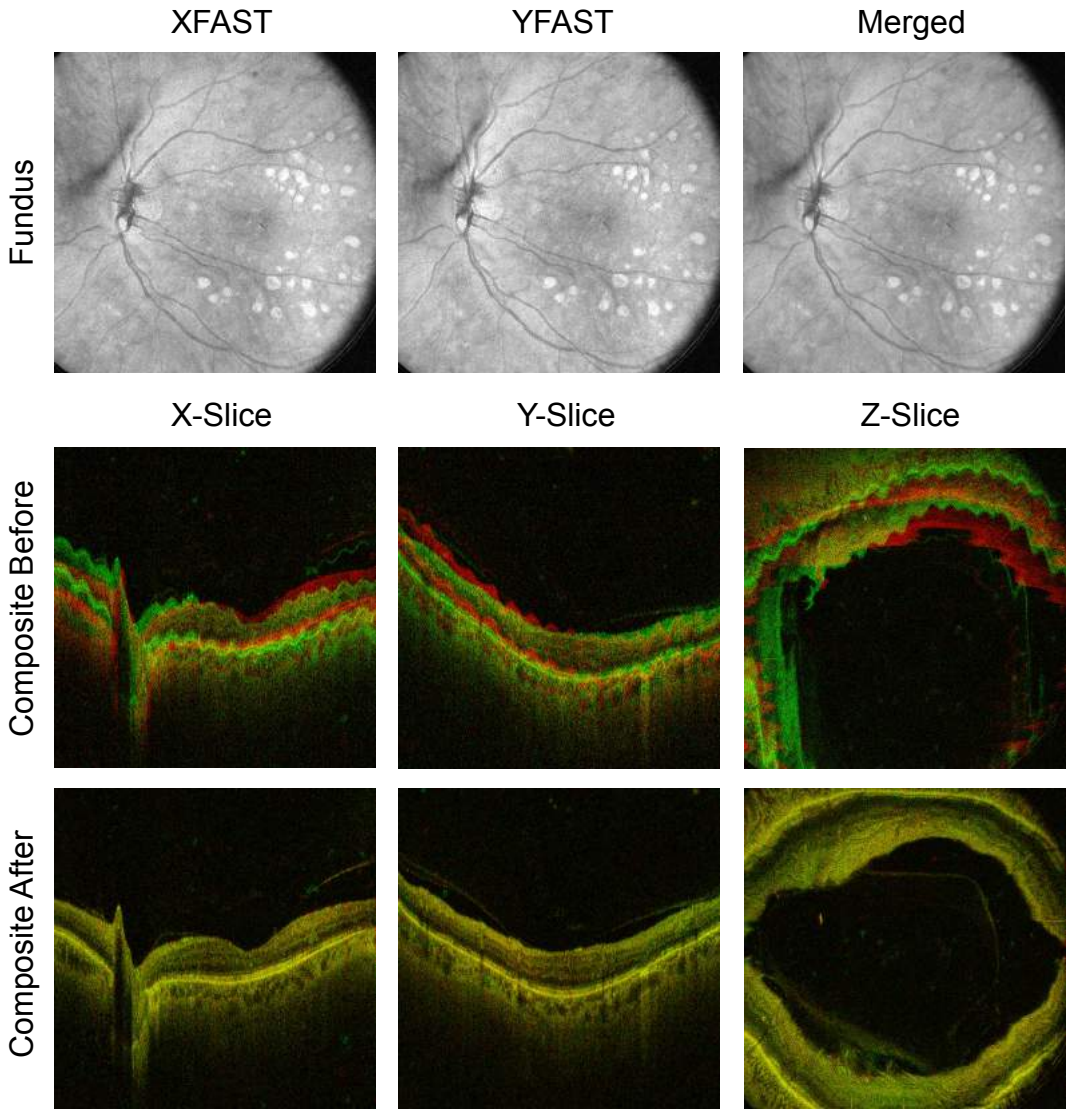

Figure 8.2: Example data showing the effect of motion correction on a 12x12 mm field.

As part of an ongoing collaboration with the Medical University of Vienna, Kajic and Esmaeelpour et al. used the motion correction algorithm to improve wide field imaging of patients with a 1060 nm based prototype OCT device [Kaji 13, Esma 14].

Within the context of these papers the improved ability to create en face visualizations of the 3D data due to motion correction was especially useful. Figure 8.3 shows and example of the possibility to create high quality en face visualizations of the data from figure 8.2. Multiple single axial pixel en face slices that are relative to the RPE layer of the retina are shown. These were generated by first segmenting the RPE in the merged volume using an automatic algorithm (see section 6.6.2). Subsequently, the segmentation was manually corrected in areas where the automatic segmentation failed. The RPE layer segmentation was then used as a reference layer to *flatten* the volume relative to it. Flattening was achieved by shifting each A-scan of the volume in axial direction such that the RPE position after shifting was at the same axial depth for the whole volume. The figure shows several en

face slices of this flattened volume at and below the RPE layer. Due to the spatial continuity of the underlying merged and motion corrected data and its high signal quality, artifact free en face views corresponding to individual anatomical layers can be generated.

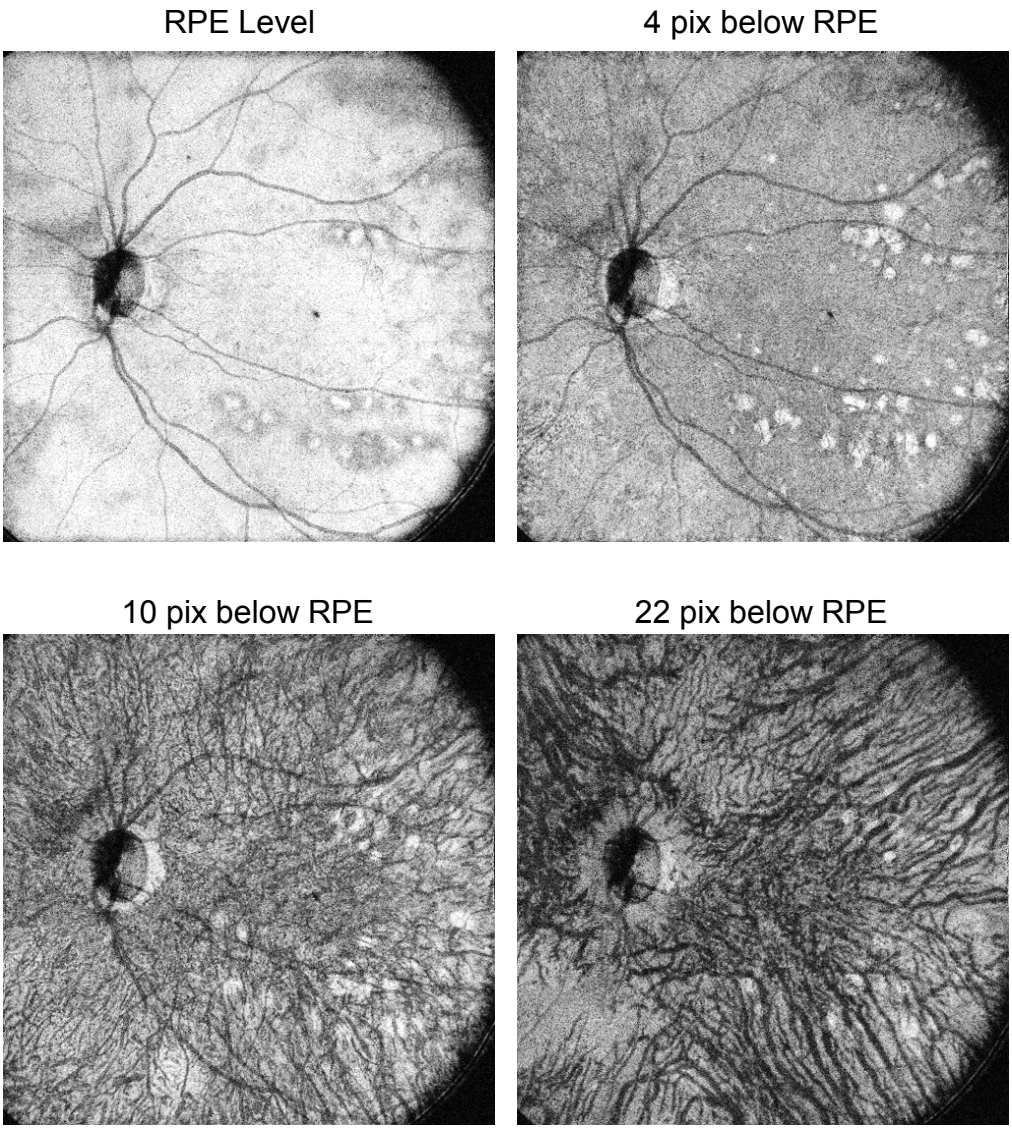

Figure 8.3: En Face visualizations of motion corrected data from figure 8.2.

8.2.1 Hand-held OCT

So far, OCT devices can mostly be found in ophthalmologic clinics. This is due to the size and cost of such devices. For screening and other specialty purposes such as the imaging of infants it would be desirable to have a low-cost, small, hand held device. However, when the device itself is not fixed but held by a hand that itself is not fixated, additional motion is introduced during imaging. This increases the need to perform motion correction.

Lu et al. designed and built two prototype hand held OCT instruments based on a Micro-Electro-Mechanical Systems (MEMS) scanning mirror and ultrahigh speed swept source OCT [Lu 13]. Figure 8.4 shows example data from a healthy young subject before and after correction that was acquired using one of these prototypes in combination with a nm vertical cavity surface emitting laser (VCSEL) based swept source OCT system running at 350 kHz A-scan rate. Because the MEMS mirror that is used is not able to scan a high speed linear raster, a sinusoidal raster was used instead. One sinusoidal B-scan consisted of 1350 A-scans, with there being 400 B-scans per volume over a 6x6 mm transverse field centered on the Macula. After acquisition, each B-scan was linearized using re-sampling. This resulted in 400×400 A-scan input volumes. A single pair of orthogonal volumes was used as input. As can be seen in the figure, the algorithm is also able to register and motion correct this kind of data.

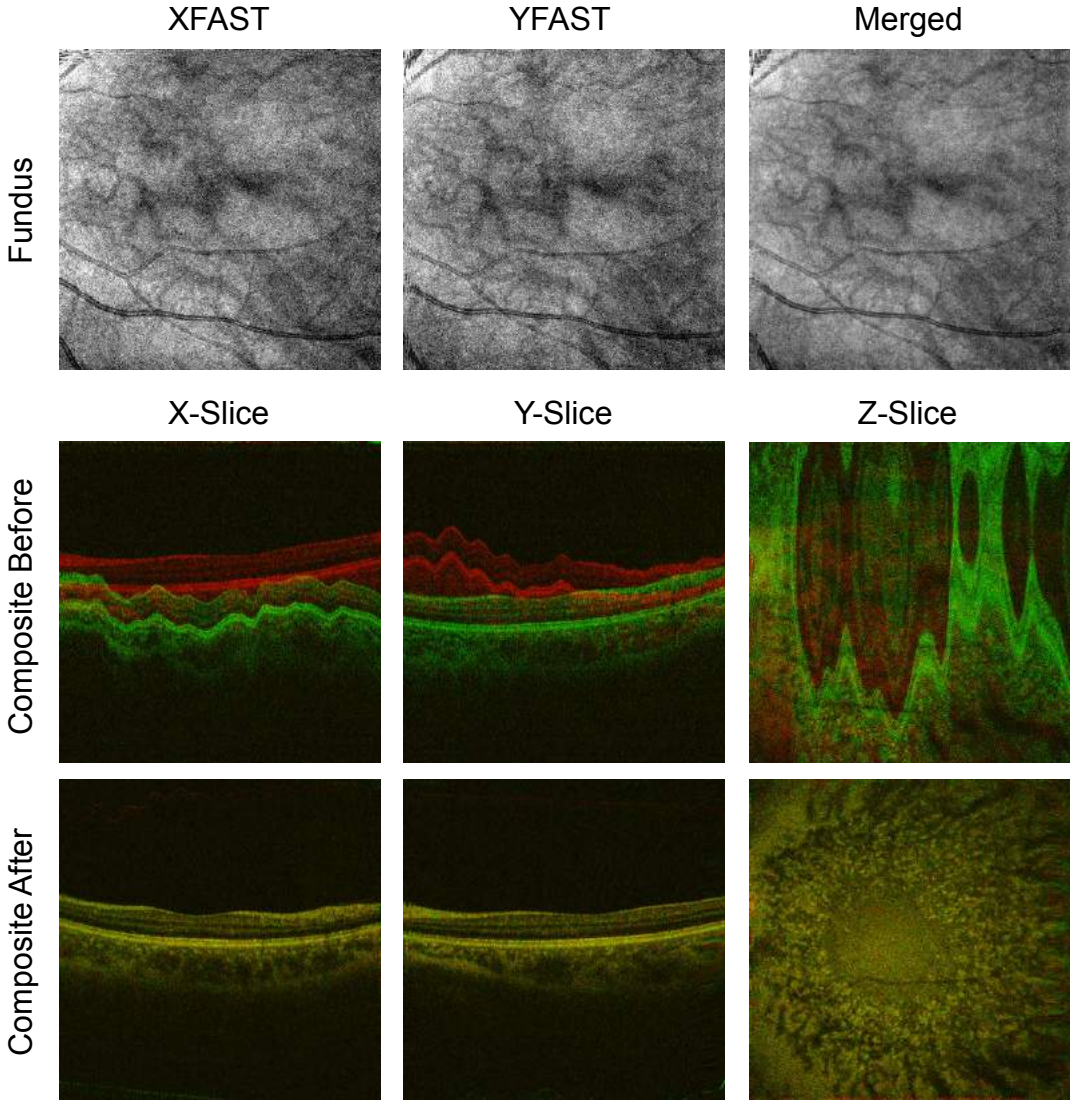

Figure 8.4: Example showing motion correction and merging of OCT volumes acquired with a prototype handheld OCT system.

8.2.2 Small Animal Imaging

In addition to being used for the imaging of humans with OCT, the motion correction algorithm was also employed for in-vivo imaging of rodent eyes [Liu 13b]. Our motion correction algorithm was then used to correct the acquired volumes retrospectively.

8.3 Enhancement of Auxiliary Data

In addition to improving structural imaging using motion correction, functional OCT can also benefit from motion correction. In this context, functional OCT is defined as a type of OCT device providing additional data channels beyond structural intensity data. Motion correction can then be performed on the structural data. The resulting transform can be used to map the additional functional channels, effectively motion correcting them, too (see section 5.8). Subsequently, the functional channel data can also be merged, increasing SNR.

Several ocular and systemic pathologies are associated with abnormal blood circulation [Flam 02, Schm 99a]. Therefore, being able to visualize and quantify blood flow in the retina promises to offer advantages for diagnosis and understanding disease. There are several techniques to provide additional functional OCT data that relates to flow. A detailed discussion of these is beyond the scope of this thesis. However, motion correction has been used in conjunction with two specific techniques that will be described below.

8.3.1 Intensity based Angiography

The first technique is based on the idea that when the same location is imaged multiple times within a short time frame, static tissue will tend to show the same intensity level while areas where there is blood flow will show widely fluctuating intensity levels. This is because as the blood cells move through the vessels, the speckle pattern changes. Speckle variance [Mari 10] and amplitude decorrelation, specifically split-spectrum amplitude decorrelation (SSADA) [Jia 12] are two representatives of such the intensity change based technique.

Figure 8.5 shows an example of motion correction and merging in conjunction with amplitude decorrelation based angiography (in this case not split spectrum) 3D-OCT data. The data was acquired at NEEC using a 1060 nm center wavelength swept source VCSEL based system operating at 400 kHz A-scan rate. A 6x6 mm area on the retina centered at the macula was scanned. The subject here was a 78 year old male with wet AMD. To obtain decorrelation information, five B-scans at the same location were scanned back-to-back. Each B-scan contained 500 A-scans and 500 slow direction transverse positions were sampled.

As part of pre-processing, the multiple available B-scans were rigidly registered and decorrelation information was extracted. In addition, the structural intensity information was combined by averaging. Therefore, the input volumes for motion correction and merging were 500×500 A-scans in size and contained an

intensity and decorrelation channel each. Two pairs of orthogonal volumes of this kind were then motion corrected and merged.

As can be seen in figure 8.5, the two intensity input fundus projections show significant transverse motion artifacts. Also, a large-scale atrohpic area is visible as the large bright area in the fundus projections. After motion correction and merging, the merged intensity fundus projection shows that motion artifacts have been successfully removed. In the corresponding angiography fundus projection, the vasculature, in this case mostly of the choroid, shows a high degree of continuity and can be nicely appreciated.

The combination of motion correction and merging and decorrelation based angiography holds promise for non-invasively assessing perfusion in the retina. Jia et al. used SSADA in conjunction with motion correction and merging of four input volumes to perform angiography of the optic disc in Glaucoma [Jia 14a, Jia 14b]. In these publications, a flow index was defined as a mean decorrelation value over the optic disc area. It was found that the flow index was reduced significantly for glaucomateous subjects compared to normal subjects. Also, the repeatability of the calculation was found to be good. This might offer a useful way to detect Glaucoma and measure progression.

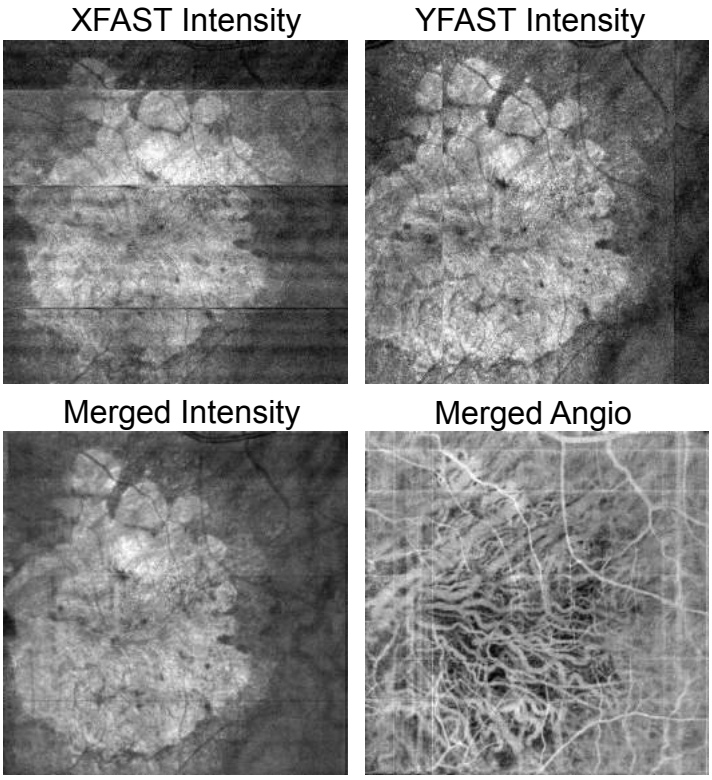

Figure 8.5: Example showing motion correction and merging of amplitude decorrelation angiography OCT volumes.

8.3.2 Doppler Imaging

The second major flow sensitive functional OCT technique is Doppler OCT [Leit 03b, Whit 03, Baum 11a]. Here, the change in phase between consecutively acquired and overlapping A-scans is used to measure the speed of flow along the axial direction.

Liu et al. used Doppler OCT and motion correction and merging to image the circulation in the ONH of a rat in-vivo [Liu 13b].

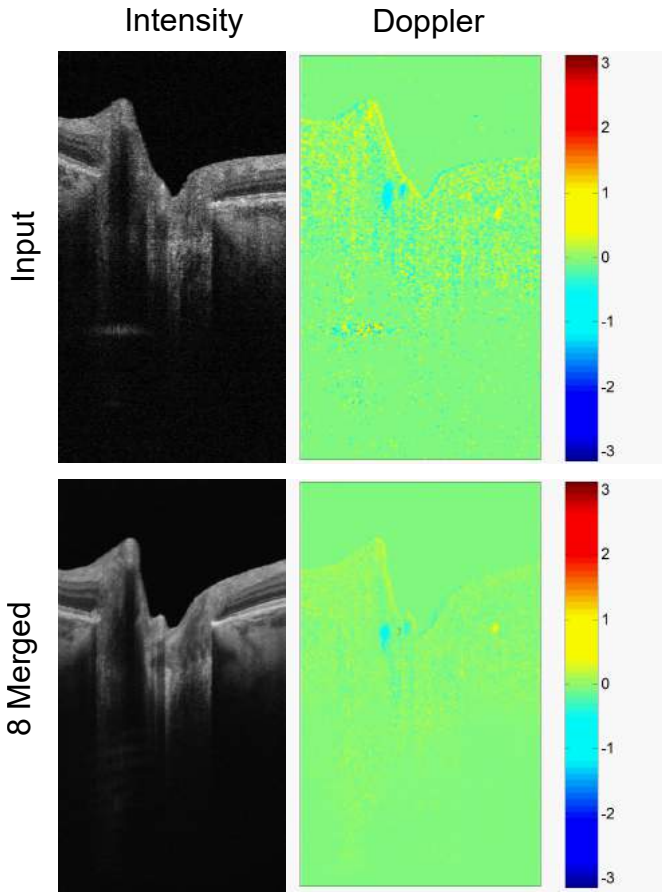

Figure 8.6: Comparison of intensity and associated Doppler shift channel before and after motion correction and merging.

Motion correction and merging can also be applied to retinal Doppler OCT data of humans. So far, only preliminary experiments have been performed in this direction, though. Figure 8.6 shows corresponding intensity and Doppler slices of OCT volumes of a human ONH before and after motion correction and merging of eight volumes. 1000×200 A-Scans were sampled over a 3x3 mm area for each volume. The volumes were then downsampled to 200×200 A-Scans while extracting Doppler shift information. Eight orthogonal such volumes were then motion corrected and merged. The top row of figure 8.6 shows a the intensity and Doppler channel of a central slice of one of the input volumes. The bottom row shows the result of motion correction and merging. It can be observed that merging eight volumes results in a very high quality, almost noise free structural

image. The Doppler channel shows a similar effect, with background noise being reduced. The three main vessels that can be observed are delineated much clearer in the merged Doppler slice.

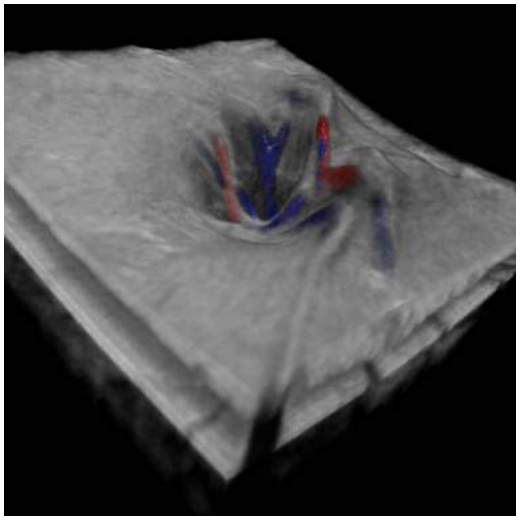

Figure 8.7: 3D Volume rendering of merged volume from figure 8.6.

In order to get a more comprehensive view of the resulting dual channel 3D data, volume rendering techniques can be used. Figure 8.7 shows a volume rendered view of the combined data from figure 8.6. Here, the structural data was rendered in gray scale while positive and negative flow were visualized as red and blue, respectively. In the view one can see the arteries and veins originating from within the ONH.

Due to removed motion artifacts and increased signal quality, motion correction and merging might in the future also be useful in quantifying total retinal flow using Doppler OCT such as in [Baum 11a, Choi 12]. This is beyond the scope of this work though and might be subject of future research.

8.4 Summary

The chapter focused on applications of the developed motion correction algorithm. The algorithm was deployed to several clinical partners in the form of an easy to use, fully automatic package. In addition, the technique has been commercialized and is already being used as part of a OCT system from Optovue Inc.

The most common application is to improve structural imaging for clinical practice and research. The effect of motion correction and merging was shown on several clinical data sets. It could be seen that the technique leads to a tangible improvement in image quality. Furthermore, in conjunction with segmentation, the lack of motion artifacts enables high quality en face visualization. Additional areas where motion correction was employed are hand held OCT imaging and small animal imaging.

In addition to structural imaging, the algorithm was also employed to enhance functional imaging. Functional techniques provide an additional data channel be-

yond intensity. Motion correction and merging can also be applied to such additional channels.

Visualization of blood flow is a key interest in the research community. One OCT method to visualize blood flow is intensity based angiography. An example was shown using angiography, motion correction and an ultra high-speed OCT system together to visualize circulation.

An alternative flow sensitive technique is Doppler OCT, providing quantitative flow information along the axial direction. Motion correction and Doppler OCT were combined for imaging the circulation in the ONH of a rat in-vivo. Also, there is preliminary research into performing Doppler OCT and motion correction in humans. An example showed the improvement in signal quality. Also, volume rendering can be used for a more comprehensive visualization. In the future, techniques based on the combination of motion correction and Doppler OCT might be helpful in improving the quantification of flow.

Part III

Outlook and Summary

CHAPTER 9

Outlook

The work presented in this thesis can form the basis for further research. In this chapter, some potential avenues are explored:

9.1 Algorithm Improvements

While the algorithm as presented here leads to significant improvements in data reliability and signal quality, there are some limitations that might warrant further work.

9.1.1 Modeling Rotation around the Optical Axis

The algorithm currently assumes that the fast scan directions of the two or more volumes are orthogonal *in object space*. Now, when the subject's head rotates around the optical axis during or in between the acquisition of the orthogonal volumes, this assumption can be violated. If there is significant rotation of this kind the algorithm has to model the rotation induced effects as heavy motion. This is likely to cause registration failure. Thankfully, cases with significant head rotation around the optical axis are very seldom. Nevertheless, addressing this problem presents interesting challenges, and might improve overall robustness.

9.1.2 Segmentation Based Similarity Measures

A typical clinical work flow for retinal imaging will involve a segmentation step in order to provide quantification of the image. For this, segmentation algorithms for retinal layers are needed and available. It would be interesting to incorporate segmentation information into the motion correction algorithm itself. Currently, similarity between volumes in the transformed state is calculated based on one to one voxel intensity differences. One the one hand, this assumes no specific object being imaged. On the other hand, it incorporates no domain specific information of the class of objects (i.e. retinas).

Given a reliable segmentation algorithm of retinal layers and blood vessels, one can incorporate the information in order to guide the optimization such that

better registration of anatomical structures is achieved. One way to do this is to modify the pre-processed volume content based on the segmentation information. For example, the background could be blacked out completely. Also, blood vessel shadows could be assigned a special intensity value. In this context, similarity would still be calculated on a voxel basis, but the intensities of the voxels are modified based on the segmentation information.

A second approach would be to add an additional similarity term based on the symbolic segmentation information. For example, the term could minimize the absolute difference between corresponding layer boundaries in the transformed volumes. Another possible term could be to penalize differences in whether a blood vessel was detected at a particular A-Scan. This way, the intensity based similarity information can be augmented using the domain specific information. However, it would be important that the segmentation is robust. In the case that segmentation failure leads to inconsistent segmentation of the input volumes, the addition of segmentation information can actually be harmful for the overall robustness. Also, the segmentation would need to be fast as overall algorithm speed is an issue.

9.1.3 Automatic Failure Detection and Parameter Tuning

As opposed to a normal image registration task, no reference volume is available. This means that similarity of the registered volumes is not sufficient to judge whether the algorithm succeeded. For fully automatic operation of large data sets it would be useful however, to be able to automatically detect algorithm failure. This would also allow changing the algorithm settings (in particular α) and try again.

It might be possible to use the Q-metric (see section 6.6.6) that was used for automatic image quality assessment, for this purpose. Calculating the quality measure on the original volumes, the registered volumes and on the merged volume would provide five feature dimensions from which one could try to judge success or failure. A simple heuristic might be used such as the relative improvement of Q from the registered to the merged volume. Alternatively, this can be treated as a machine learning task. It would be interesting to see how well such an automatic detection would work. Also the robustness with respect to different object types (e.g. Macula versus ONH) would be of interest.

9.1.4 Run Time Improvements

The overall run-time of the motion correction algorithm is of key interest. The shorter the run time, the better the algorithm can be integrated into the clinical work flow. Run time results so far are tolerable (see section 7.4) but improvements would certainly be welcome. Especially an optimized implementation of the better performing advanced algorithm would be useful. Other than that, for the same number and sizes of volumes, the run time is bound to improve as computer hardware, especially GPUs, advance. However, as OCT systems become faster, the av-

erage volume size is also bound to become larger. For this reason, optimization of the algorithm run time remains a constant concern.

9.2 Application Outlook

In addition to core algorithm improvements, the availability of motion corrected data with high signal quality opens up new potential possibilities for applications.

For example, the combination of intensity based angiography and motion correction (see section 8.3.1) opens up possibilities to quantitatively look at perfusion of the optic disk and capillary networks in the retina. This makes optimized segmentation of flow and subsequent analysis necessary. The availability of motion correction and merging might lead to different choices here.

Similarly, Doppler OCT in general and total retinal flow quantification in particular (see section 8.3.2) might benefit from motion correction and merging. Key issues here would be the segmentation of the vessel tree and the handling of pulsatility in the flow. Motion correction and merging provides multiple registered samples from different time points for this information. Therefore, an optimized algorithm could lead to a better quantification of total retinal flow.

CHAPTER 10

Summary

The main focus of this work was the introduction, evaluation and application of a novel motion correction algorithm for 3D-Optical Coherence Tomography.

Part 1 of the work introduced the fundamentals in technical and medical OCT. Chapter 2 starts out with OCT itself. Here, the basic operation principle of OCT was introduced. Low-coherence light reflected from a reference surface and from the sample gives rise to an interference pattern. From this pattern, the back-reflected intensity of the sample along depth can be calculated. Also, key imaging parameters such as speed, axial and transverse resolution and sensitivity were identified. In combination with lateral scanning, multi-dimensional images can be created. Raster scanning represents an important way to generate 3D volumetric OCT data. This work in particular operates with orthogonally scanned raster scans.

Chapter 3 focused on important aspects of OCT imaging in ophthalmologic practice. For this purpose, basic anatomy of the eye was introduced. The eye has optics to focus light on the retina in the back of the eye. The retina itself is a layered structure responsible for sensing the light. Many retinal diseases manifest as changes in this layered structure. Eye motion plays an important role in the sensing process, to prevent an effect called neural adaptation. The retina is scanned using a collimated OCT beam. By varying the incident angle of the beam, different lateral positions on the retina can be imaged. This enables 2D and 3D imaging, with raster scanning playing an important role for 3D. OCT is able to image the retina in 3D and non-invasively and as such is well suited for diagnostics and tracking disease progression. However, several effects such as speckle noise, blinking, illumination effects, floaters, tilt and especially motion artifacts affect this ability. Motion artifacts result from relative motion between the subject and the OCT device and distort the acquired data. Axial and transverse motion artifacts can be distinguished. Also, motion artifacts will manifest differently, depending on the concrete scan pattern. For example, the fast scan direction in a raster scan presents as relatively undistorted, as opposed to the slow scan direction.

Chapter 4 is concerned with the state of the art on motion artifact correction and signal enhancement in OCT. One way to reduce motion artifacts is to increase acquisition speed. Tracking methods represent another way. Here the position where the OCT beam is pointing on the retina is measured and corrected. One can also attempt to correct motion artifacts through post processing. One option here is to relate the OCT images to images from another modality that does not suffer from motion artifacts. In absence of a fixed reference, motion can still be corrected for example by assuming that the structure that is depicted in a volume

is fundamentally smooth. Any high frequency content is assumed to be the result of motion artifacts and can be filtered out. Finally, orthogonally scanned data has been used to correct motion artifacts. Techniques range from having a single or a few orthogonal "guidepost" B-scans to which the full volume data is registered. On the other end of the spectrum, multiple whole orthogonal raster scans have been used for correction. The state of the art in OCT signal enhancement and noise reduction can be divided into physical methods, which need modifications in OCT hardware, and post processing approaches. Multiple images of the same area can be combined to reduce noise. This can furthermore be combined with digital image processing methods, which operate on a single or multiple images.

In part 2, our 3D-OCT motion correction approach using image registration and orthogonal raster scans is introduced, evaluated and applications are shown. Chapter 5 presents a detailed description of the algorithm. The method is treated as a special kind of registration problem without a reference. This means that all input volumes are transformed in order to register the volumes. The problem is regularized with an application specific regularization based on the time structure of the OCT scanning process. The influence of this regularization is controlled using a parameter α. In order to improve robustness, an illumination correction approach can be employed in pre-processing. In order to assess the similarity of volumes within the objective function, a sum of squared differences approach can be used. Alternatively, a pseudo-Huber loss function is employed for improved robustness. For the optimization of the objective function, multi-resolution and multi-stage methods are used. Multi-resolution uses different resolution representations of the input data during optimization. Multi-stage on the other hand uses different parametrization of the parameters of the optimization, leading to a different number of degrees of freedom. As part of the multi-stage approach, differences in alignment related tilt between the input volumes are corrected. This is achieved via modeling of the corresponding degrees of freedom. Furthermore, an additional data term is used in order to remove overall tilt. After optimization, the set of registered volumes can merged into a single higher-quality volume. For this purpose an adaptive weighting scheme based on the concepts of sample validity and and sampling density is used. In addition to intensity data, data channels that carry functional information can also be motion corrected and merged. Finally, GPU optimization techniques were used to optimize key parts of the algorithm.

Chapter 6 presents the methodology that is used to evaluate the presented algorithm. Since the method operates without reference, registration success in the sense of achieving similarity is not sufficient for evaluation. Therefore, a three pronged approach is used. First, registration performance is evaluated. Second, the reproducibility of the output volumes is evaluated assess whether the results are reliable. Third, signal improvement is evaluated using automated methods. In order to evaluate the first two criteria, mutual information is used to assess similarity. In addition, automatically segmented maps of features from the volumes are compared for similarity. For evaluating the third goal, a no-reference image quality measure called Q-metric is used. A large body of 3D-OCT data which allows for reproducibility assessment was acquired at two different clinical sites and is available for evaluation. The population of 73 subjects consists of normal subjects,

glaucomatous subjects and glaucoma suspects. Two different motion correction algorithm profiles (advanced and basic) are evaluated, corresponding two different stages in the evolution of the method. The main differences between these two profiles can be found in the use of different loss functions for similarity measure and regularization, the use of two stage registration and tilt compensation and illumination correction. Several required techniques, such as segmentation, are also described. Finally, quantitative measurements are tested for statistical significance.

In chapter 7, results are presented and discussed. For each aspect, visual inspection as well as quantitative evaluation were performed. The results consistently show improvements through motion correction and merging for the three aspects that are evaluated. More specifically, the advanced motion correction algorithm profile significantly outperforms the basic profile which again outperforms no correction. For the aspect of registration performance for example, the mean increase in mutual information through registration over all data is significantly higher for advanced correction compared to basic correction, for all α. The quantitative evaluation of reproducibility performance also showed significant improvements through motion correction. For example for $\alpha = 0.1$, the blood vessel map reproducibility error is reduced to 69% of the uncorrected error for the basic and to 47% of the uncorrected error for the advanced algorithm. Based on the aforementioned results, α was fixed at 0.1 and different subgroups of the data such as Normal subjects and Glaucoma subjects and Glaucoma suspects were inspected. The results were consistent with the results when looking at all data. Also, looking at different spatial tolerances for the reproducibility error showed a reduction in error for larger tolerances and otherwise consistent improvements through motion correction. For the aspect of signal quality, the mean relative change in the Q-metric from registered to merged slices showed the best results for advanced correction, followed by basic correction. Finally, algorithm run times of the two profiles were compared for different number and sizes of volumes. The current GPU acceleration leads to a speedup of factor three. Also, in the current implementation, the advanced profile run time is about twice as long as for the basic profile.

Chapter 8 focused on applications of the developed algorithm which was deployed to several clinical partners as a fully automatic package. The technique has been commercialized and is already being used as part of a OCT system from Optovue Inc. The most common application is to improve structural imaging for clinical practice and research. The effect of motion correction and merging was shown on several clinical data sets. It could be seen that the technique leads to a tangible improvement in image quality. Furthermore, in conjunction with segmentation, the lack of motion artifacts enables high quality en face visualization. Additional areas where motion correction was employed are hand held OCT imaging and small animal imaging. In addition to structural imaging, the algorithm was also employed to enhance functional imaging. Motion correction and merging can also be applied to additional functional data channels. Visualization of blood flow is a key interest in the research community. An example was shown of using intensity based angiography, motion correction and an ultra high-speed OCT system together to visualize circulation. An alternative flow sensitive technique

is Doppler OCT. Motion correction and Doppler OCT were combined for imaging the circulation in the optic nerved head of a rat in-vivo. Also, there is preliminary research into performing Doppler OCT and motion correction in humans. An example showed the improvement in signal quality. Volume rendering can be used for a more comprehensive visualization.

In conclusion, it is demonstrated that the motion correction algorithm can improve both the visual appearance and the reliability of quantitative measurements derived from 3D-OCT data substantially. This can lead to improved diagnosis and tracking of retinal diseases.

List of Figures

List of Symbols

Abbreviations

OCT	Optical Coherence Tomography	1
SNR	Signal to noise ratio	1
MIT	Massachussetts Institute of Technology	2
TD-OCT	Time Domain OCT	7
SLD	Superluminescend Diode	8
PZT	Piezoelectric Transducer	8
FWHM	Full-width-half-max	8
PSF	point spread function	8
NA	Numerical Aperture	9
FD-OCT	Fourier-Domain OCT	9
ONH	Optic Nerve Head	15
UHR	Ultra High Resolution	19
AMD	Age-Related Macular Degeneration	20
NFL	Nerve Fiber Layer	20
RPE	Retinal Pigment Epithelium	20
FDML	Fourier Domain Mode Locked	30
SLO	Scanning Laser Ophthalmoscope	31
PS-OCT	Polarization Sensitive OCT	43
GPU	Graphics Processing Unit	43
SSD	Sum Of Squared Differences	48
L-BFGS	Limited Memory Broyden-Fletcher-Goldfarb-Shanno	54
CG	Conjugate Gradients	54
MR	Multi-Resolution	55
CUDA	Compute Unified Device Architecture	67
US	Ultra Sound	69
MRI	Magnetic Resonance Imaging	69
PSNR	Peak Signal To Noise Ratio	72
MI	Mutual Information	74
IOP	Intraocular Pressure	75
ILM	Inner Limiting Membrane	79
IS	Inner Segments	79
NEEC	New England Eye Center	107
UPMC	University of Pittsburgh Medical Center	108
MEMS	Micro Electro Mechanical Systems	112
VCSEL	Vertical Cavity Surface Emitting Laser	112
SSADA	Split Spectrum Amplitude Decorrelation	113

Symbols

List of Tables

Bibliography

[Adhi 14] M. Adhi, J. J. Liu, A. H. Qavi, I. Grulkowski, C. D. Lu, K. J. Mohler, D. Ferrara, M. F. Kraus, C. R. Baumal, A. J. Witkin, N. K. Waheed, J. Hornegger, J. G. Fujimoto, and J. S. Duker. "Choroidal analysis in healthy eyes using swept-source optical coherence tomography compared to spectral domain optical coherence tomography". *Am J Ophthalmol*, Vol. 157, No. 6, pp. 1272–1281 e1, Jun 2014.

[Adle 04] D. C. Adler, T. H. Ko, and J. G. Fujimoto. "Speckle reduction in optical coherence tomography images by use ofa spatially adaptive wavelet filter". *Opt. Lett.*, Vol. 29, No. 24, pp. 2878–2880, Dec 2004.

[Ahse 13] O. O. Ahsen, Y. K. Tao, B. M. Potsaid, Y. Sheikine, J. Jiang, I. Grulkowski, T.-H. Tsai, V. Jayaraman, M. F. Kraus, J. L. Connolly, J. Hornegger, A. Cable, and J. G. Fujimoto. "Swept source optical coherence microscopy using a 1310 nm VCSEL light source". *Opt. Express*, Vol. 21, No. 15, pp. 18021–18033, Jul 2013.

[Alas 15] T. Alasil, D. Ferrara, M. Adhi, E. Brewer, M. F. Kraus, C. R. Baumal, J. Hornegger, J. G. Fujimoto, A. J. Witkin, E. Reichel, J. S. Duker, and N. K. Waheed. "En Face Imaging of the Choroid in Polypoidal Choroidal Vasculopathy Using Swept-Source Optical Coherence Tomography". *American Journal of Ophthalmology*, Vol. 159, No. 4, pp. 634–643.e2, Apr 2015.

[ANSI 07] ANSI. *American National Standard for Safe Use of Lasers*. Laser Institute of America, Orlando, 2007.

[Anto 11] B. Antony, M. D. Abramoff, 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". *Biomed Opt Express*, Vol. 2, No. 8, pp. 2403–16, Aug 2011.

[Bash 00] M. Bashkansky and J. Reintjes. "Statistics and reduction of speckle in optical coherence tomography". *Opt. Lett.*, Vol. 25, No. 8, pp. 545–547, Apr 2000.

[Baum 11a] B. Baumann, B. Potsaid, M. F. Kraus, J. J. Liu, D. Huang, J. Hornegger, A. E. Cable, J. S. Duker, and J. G. Fujimoto. "Total retinal blood flow measurement with ultrahigh speed swept source/Fourier domain OCT". *Biomed Opt Express*, Vol. 2, No. 6, pp. 1539–52, Jun 2011.

[Baum 11b] B. Baumann, B. Potsaid, J. J. Liu, M. F. Kraus, D. Huang, J. Hornegger, J. S. Duker, and J. G. Fujimoto. "Retinal blood flow measurement with ultrahigh-speed swept-source / Fourier domain optical coherence tomography". In: *Proc. SPIE*, pp. 78850H–78850H–9, Feb 2011.

[Baum 12] B. Baumann, W. Choi, B. Potsaid, D. Huang, J. S. Duker, and J. G. Fujimoto. "Swept source / Fourier domain polarization sensitive optical coherence tomography with a passive polarization delay unit". *Opt. Express*, Vol. 20, No. 9, pp. 10229–10241, Apr 2012.

[Boer 03] J. F. de Boer, B. Cense, B. H. Park, M. C. Pierce, G. J. Tearney, and B. E. Bouma. "Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography". *Opt. Lett.*, Vol. 28, No. 21, pp. 2067–2069, Nov 2003.

[Bomm 06] U. Bommas-Ebert, P. Teubner, and R. Voss. *Kurzlehrbuch Anatomie und Embryologie: 47 Tabellen;[mit Muskeltrainer]*. Georg Thieme Verlag, 2006.

[Boni 10] T. Bonin, G. Franke, M. Hagen-Eggert, P. Koch, and G. Huttmann. "In vivo Fourier-domain full-field OCT of the human retina with 1.5 million A-lines/s". *Opt Lett*, Vol. 35, No. 20, pp. 3432–4, Oct 2010.

[Capp 11] A. G. Capps, R. J. Zawadzki, Q. Yang, D. W. Arathorn, C. R. Vogel, B. Hamann, and J. S. Werner. "Correction of eye-motion artifacts in AO-OCT data sets". In: *Proc. SPIE*, pp. 78850D–78850D–7, Feb 2011.

[Chin 97] S. R. Chinn, E. A. Swanson, and J. G. Fujimoto. "Optical coherence tomography using a frequency-tunable optical source". *Opt. Lett.*, Vol. 22, No. 5, pp. 340–342, Mar 1997.

[Chit 12] S. Chitchian, M. A. Mayer, A. R. Boretsky, F. J. van Kuijk, and M. Motamedi. "Retinal optical coherence tomography image enhancement via shrinkage denoising using double-density dual-tree complex wavelet transform". *Journal of Biomedical Optics*, Vol. 17, No. 11, pp. 116009–116009, Nov 2012.

[Chiu 10] S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu. "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation". *Opt. Express*, Vol. 18, No. 18, pp. 19413–19428, Aug 2010.

[Choi 12] W. Choi, B. Baumann, J. J. Liu, A. C. Clermont, E. P. Feener, J. S. Duker, and J. G. Fujimoto. "Measurement of pulsatile total blood flow in the human and rat retina with ultrahigh speed spectral/Fourier domain OCT". *Biomedical Optics Express*, Vol. 3, No. 5, pp. 1047–1061, May 2012.

[Chom 03] M. Choma, M. Sarunic, C. Yang, and J. Izatt. "Sensitivity advantage of swept source and Fourier domain optical coherence tomography". *Opt. Express*, Vol. 11, No. 18, pp. 2183–2189, Sep 2003.

[Clin 80] D. Cline, H. Hofstetter, J. Griffin, and M. Schapero. *Dictionary of visual science*. Chilton Book Co., 1980.

[Drex 08] W. Drexler and J. G. Fujimoto. "State-of-the-art retinal optical coherence tomography". *Prog Retin Eye Res*, Vol. 27, No. 1, pp. 45–88, 2008.

[Drex 15] W. Drexler and J. G. Fujimoto. *Optical coherence tomography: technology and applications. Second edition*. Springer, 2015.

[Engb 03] R. Engbert and R. Kliegl. "Microsaccades uncover the orientation of covert attention". *Vision Research*, Vol. 43, No. 9, pp. 1035–1045, Apr 2003.

[Esma 14] M. Esmaeelpour, V. Kajic, B. Zabihian, R. Othara, S. Ansari-Shahrezaei, L. Kellner, I. Krebs, S. Nemetz, M. F. Kraus, J. Hornegger, J. G. Fujimoto, W. Drexler, and S. Binder. "Choroidal Haller's and Sattler's Layer Thickness Measurement Using 3-Dimensional 1060-nm Optical Coherence Tomography". *PLoS One*, Vol. 9, No. 6, p. e99690, Jun 2014.

[Ferc 03] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser. "Optical coherence tomography - principles and applications". *Reports on Progress in Physics*, Vol. 66, No. 2, p. 239, Jan 2003.

[Ferc 86] A. F. Fercher and E. Roth. "Ophthalmic Laser Interferometry". In: *Proc. SPIE*, pp. 48–51, Sep 1986.

[Ferc 88] A. F. Fercher, K. Mengedoht, and W. Werner. "Eye-length measurement by interferometry with partially coherent light". *Opt. Lett.*, Vol. 13, No. 3, pp. 186–188, Mar 1988.

[Ferc 95] A. Fercher, C. Hitzenberger, G. Kamp, and S. El-Zaiat. "Measurement of intraocular distances by backscattering spectral interferometry". *Optics Communications*, Vol. 117, No. 1 - 2, pp. 43 – 48, May 1995.

[Ferg 04] R. D. Ferguson, D. Hammer, L. A. Paunescu, S. Beaton, and J. S. Schuman. "Tracking optical coherence tomography". *Opt. Lett.*, Vol. 29, No. 18, pp. 2139–2141, Sep 2004.

[Ferr 14] D. Ferrara, K. J. Mohler, N. Waheed, M. Adhi, J. J. Liu, I. Grulkowski, M. F. Kraus, C. Baumal, J. Hornegger, J. G. Fujimoto, and J. S. Duker. "En face enhanced-depth swept-source optical coherence tomography features of chronic central serous chorioretinopathy". *Ophthalmology*, Vol. 121, No. 3, pp. 719–26, Mar 2014.

[Fish 95] N. I. Fisher. *Statistical analysis of circular data*. Cambridge University Press, 1995.

[Flam 02] J. Flammer, S. Orgul, V. P. Costa, N. Orzalesi, G. K. Krieglstein, L. M. Serra, J. P. Renard, and E. Stefansson. "The impact of ocular blood flow in glaucoma". *Prog Retin Eye Res*, Vol. 21, No. 4, pp. 359–93, Jul 2002.

[Fran 98] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. *Multiscale vessel enhancement filtering*, pp. 130–137. Vol. 1496, Springer Berlin Heidelberg, 1998.

[Fuji 86] J. G. Fujimoto, C. A. Puliafito, R. Margolis, A. Oseroff, S. D. Silvestri, and E. P. Ippen. "Femtosecond optical ranging in biological systems". *Opt. Lett.*, Vol. 11, No. 3, pp. 150–152, Mar 1986.

[Gabr 07] M. L. Gabriele, H. Ishikawa, G. Wollstein, R. A. Bilonick, L. Kagemann, M. Wojtkowski, V. J. Srinivasan, J. G. Fujimoto, J. S. Duker, and J. S. Schuman. "Peripapillary nerve fiber layer thickness profile determined with high speed, ultrahigh resolution optical coherence tomography high-density scanning". *Invest Ophthalmol Vis Sci*, Vol. 48, No. 7, pp. 3154–3160, Jul 2007.

[Good 76] J. W. Goodman. "Some fundamental properties of speckle". *J. Opt. Soc. Am.*, Vol. 66, No. 11, pp. 1145–1150, Nov 1976.

[Grou 04] E. D. P. R. Group *et al.* "Prevalence of age-related macular degeneration in the United States". *Archives of ophthalmology*, Vol. 122, No. 4, pp. 564–572, Apr 2004.

[Habe 97] U. Haberland, P. Jansen, V. Blazek, and H. J. Schmitt. "Optical coherence tomography of scattering media using frequency-modulated continuous-wave techniques with tunable near-infrared laser". *Proc. SPIE*, Vol. 2981, No. , pp. 20–28, May 1997.

[Hamm 05] D. X. Hammer, R. D. Ferguson, J. C. Magill, L. A. Paunescu, S. Beaton, H. Ishikawa, G. Wollstein, and J. S. Schuman. "Active retinal tracker for clinical optical coherence tomography systems". *Journal of Biomedical Optics*, Vol. 10, No. 2, pp. 024038–024038, Mar-Apr 2005.

[Haus 98] G. Häusler and M. W. Lindner. ""Coherence Radar" and "Spectral Radar" – New Tools for Dermatological Diagnosis". *Journal of Biomedical Optics*, Vol. 3, No. 1, pp. 21–31, Jan 1998.

[Hend 13] H. C. Hendargo, R. Estrada, S. J. Chiu, C. Tomasi, S. Farsiu, and J. A. Izatt. "Automated non-rigid registration and mosaicing for robust imaging of distinct retinal capillary beds using speckle variance optical coherence tomography". *Biomed. Opt. Express*, Vol. 4, No. 6, pp. 803–821, Jun 2013.

[Hitz 03] C. Hitzenberger, P. Trost, P.-W. Lo, and Q. Zhou. "Three-dimensional imaging of the human retina by high-speed optical coherence tomography". *Optics Express*, Vol. 11, No. 21, pp. 2753–2761, Oct 2003.

[Ho 09] J. Ho, A. C. Sull, L. N. Vuong, Y. Chen, J. Liu, J. G. Fujimoto, J. S. Schuman, and J. S. Duker. "Assessment of Artifacts and Reproducibility across Spectral- and Time-Domain Optical Coherence Tomography Devices". *Ophthalmology*, Vol. 116, No. 10, pp. 1960 – 1970, Oct 2009.

[Hou 06] Z. Hou. "A Review on MR Image Intensity Inhomogeneity Correction". *Int J Biomed Imaging*, Vol. 2006, p. 49515, 2006.

[Huan 91] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto. "Optical Coherence Tomography". *Science*, Vol. 254, No. 5035, pp. 1178–1181, Nov 1991.

[Hube 06] R. Huber, M. Wojtkowski, and J. G. Fujimoto. "Fourier Domain Mode Locking (FDML): A new laser operating regime and applications for optical coherence tomography". *Opt. Express*, Vol. 14, No. 8, pp. 3225–3237, Apr 2006.

[Hube 64] P. J. Huber. "Robust estimation of a location parameter". *The Annals of Mathematical Statistics*, Vol. 35, No. 1, pp. 73–101, Mar 1964.

[Ishi 06] H. Ishikawa, M. L. Gabriele, G. Wollstein, R. D. Ferguson, D. X. Hammer, L. A. Paunescu, S. A. Beaton, and J. S. Schuman. "Retinal nerve fiber layer assessment using optical coherence tomography with active optic nerve head tracking". *Invest Ophthalmol Vis Sci*, Vol. 47, No. 3, pp. 964–967, Mar 2006.

[Jia 12] Y. Jia, O. Tan, J. Tokayer, B. Potsaid, Y. Wang, J. J. Liu, M. F. Kraus, H. Subhash, J. G. Fujimoto, J. Hornegger, and D. Huang. "Split-spectrum amplitude-decorrelation angiography with optical coherence tomography". *Opt. Express*, Vol. 20, No. 4, pp. 4710–4725, Feb 2012.

[Jia 14a] Y. Jia, E. Wei, X. Wang, X. Zhang, J. C. Morrison, M. Parikh, L. H. Lombardi, D. M. Gattey, R. L. Armour, B. Edmunds, M. F. Kraus, J. G. Fujimoto, and D. Huang. "Optical Coherence Tomography Angiography of Optic Disc Perfusion in Glaucoma". *Ophthalmology*, Jul 2014.

[Jia 14b] Y. Jia, S. T. Bailey, D. J. Wilson, O. Tan, M. L. Klein, C. J. Flaxel, B. Potsaid, J. J. Liu, C. D. Lu, and M. F. Kraus. "Quantitative optical coherence tomography angiography of choroidal neovascularization in age-related macular degeneration". *Ophthalmology*, Vol. 121, No. 7, pp. 1435–1444, Jul 2014.

[Kaji 13] V. Kajić, M. Esmaeelpour, C. Glittenberg, M. F. Kraus, J. Honegger, R. Othara, S. Binder, J. G. Fujimoto, and W. Drexler. "Automated three-dimensional choroidal vessel segmentation of 3D 1060 nm OCT retinal data". *Biomed. Opt. Express*, Vol. 4, No. 1, pp. 134–150, Jan 2013.

[Kara 05] B. Karamata, K. Hassler, M. Laubscher, and T. Lasser. "Speckle statistics in optical coherence tomography". *J. Opt. Soc. Am. A*, Vol. 22, No. 4, pp. 593–596, Apr 2005.

[Keys 81] R. Keys. "Cubic convolution interpolation for digital image processing". *Acoustics, Speech and Signal Processing, IEEE Transactions on*, Vol. 29, No. 6, pp. 1153–1160, Dec 1981.

[Klei 11] T. Klein, W. Wieser, C. M. Eigenwillig, B. R. Biedermann, and R. Huber. "Megahertz OCT for ultrawide-field retinal imaging with a 1050nm Fourier domain mode-locked laser". *Opt. Express*, Vol. 19, No. 4, pp. 3044–3062, Feb 2011.

[Klei 12] T. Klein, W. Wieser, R. André, T. Pfeiffer, C. M. Eigenwillig, and R. Huber. "Multi-MHz FDML OCT: snapshot retinal imaging at 6.7 million axial-scans per second". *Proc. SPIE 8213, Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XVI, (February 9, 2012)*, p. 82131E, 2012.

[Krau 09] M. Kraus. *Improving neighborhood consistency in OCT volume scans*. Master's thesis, University Erlangen-Nuremberg, 2009.

[Krau 11] M. Kraus, B. Potsaid, J. Fujimoto, M. Mayer, R. Bock, and J. Hornegger. "Method and apparatus for motion correction and image enhancement for optical coherence tomography". Nov. 3 2011. US Patent App. 13/097,785.

[Krau 12] M. F. Kraus, B. Potsaid, M. A. Mayer, R. Bock, B. Baumann, J. J. Liu, J. Hornegger, and J. G. Fujimoto. "Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns". *Biomed. Opt. Express*, Vol. 3, No. 6, pp. 1182–1199, Jun 2012.

[Krau 14] M. F. Kraus, J. J. Liu, J. Schottenhamml, C.-L. Chen, A. Budai, L. Branchini, T. Ko, H. Ishikawa, G. Wollstein, J. Schuman, J. S. Duker, J. G. Fujimoto, and J. Hornegger. "Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization". *Biomed. Opt. Express*, Vol. 5, No. 8, pp. 2591–2613, Aug 2014.

[Krau 16] M. Kraus, B. Potsaid, J. Fujimoto, M. Mayer, R. Bock, and J. Hornegger. "Method and apparatus for motion correction and image enhancement for optical coherence tomography". Feb. 23 2016. US Patent 9,269,144.

[Leit 03a] R. Leitgeb, C. Hitzenberger, and A. Fercher. "Performance of fourier domain vs. time domain optical coherence tomography". *Opt. Express*, Vol. 11, No. 8, pp. 889–894, Apr 2003.

[Leit 03b] R. Leitgeb, L. Schmetterer, W. Drexler, A. Fercher, R. Zawadzki, and T. Bajraszewski. "Real-time assessment of retinal blood flow with ultrafast acquisition by color Doppler Fourier domain optical coherence tomography". *Optics Express*, Vol. 11, No. 23, pp. 3116–21, Nov 2003.

[Leki 05] F. Lekien and J. Marsden. "Tricubic Interpolation in Three Dimensions". *Journal of Numerical Methods and Engineering*, Vol. 63, pp. 455–471, Mar 2005.

[Liu 13a] J. J. Liu, I. Grulkowski, B. Potsaid, V. Jayaraman, A. E. Cable, M. F. Kraus, J. Hornegger, J. S. Duker, and J. G. Fujimoto. "4D dynamic imaging of the eye using ultrahigh speed SS-OCT". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, March 2013.

[Liu 13b] J. J. Liu, I. Grulkowski, M. F. Kraus, B. Potsaid, C. D. Lu, B. Baumann, J. S. Duker, J. Hornegger, and J. G. Fujimoto. "In vivo imaging of the rodent eye with swept source/Fourier domain OCT". *Biomed. Opt. Express*, Vol. 4, No. 2, pp. 351–363, Feb 2013.

[Liu 14] J. J. Liu, A. J. Witkin, M. Adhi, I. Grulkowski, M. F. Kraus, A.-H. Dhalla, C. D. Lu, J. Hornegger, J. S. Duker, and J. G. Fujimoto. "Enhanced Vitreous Imaging in Healthy Eyes Using Swept Source Optical Coherence Tomography". *PLoS ONE*, Vol. 9, No. 7, p. e102950, Jul 2014.

[Liu 89] D. Liu and J. Nocedal. "On the limited memory BFGS method for large scale optimization". *Mathematical Programming*, Vol. 45, No. 1-3, pp. 503–528, Dec 1989.

[Lu 13] C. D. Lu, M. F. Kraus, B. Potsaid, J. J. Liu, W. Choi, V. Jayaraman, A. E. Cable, J. Hornegger, J. S. Duker, and J. G. Fujimoto. "Handheld ultra-high speed swept source optical coherence tomography instrument using a MEMS scanning mirror". *Biomedical Optics Express*, Vol. 5, No. 1, pp. 293–311, Jan 2013.

[Magu 07] G. Maguluri, M. Mujat, B. H. Park, K. H. Kim, W. Sun, N. V. Iftimia, R. D. Ferguson, D. X. Hammer, T. C. Chen, and J. F. de Boer. "Three dimensional tracking for volumetric spectral-domain optical coherence tomography". *Opt. Express*, Vol. 15, No. 25, pp. 16808–16817, Dec 2007.

[Mari 10] A. Mariampillai, M. K. Leung, M. Jarvi, B. Standish, K. Lee, B. Wilson, A. Vitkin, and V. D. Yang. "Optimized speckle variance OCT imaging of microvasculature". *Opt. Lett.*, Vol. 35, No. 8, pp. 1257–1259, Apr 2010.

[Mart 04] S. Martinez-Conde, S. L. Macknik, and D. H. Hubel. "The role of fixational eye movements in visual perception". *Nature Reviews Neuroscience*, Vol. 5, No. 3, pp. 229–240, Mar 2004.

[Mast 98] B. Masters. "Three-dimensional confocal microscopy of the human optic nerve in vivo". *Opt. Express*, Vol. 3, No. 10, pp. 356–359, Nov 1998.

[Maye 12] M. A. Mayer, A. Borsdorf, M. Wagner, J. Hornegger, C. Y. Mardin, and R. P. Tornow. "Wavelet denoising of multiframe optical coherence tomography data". *Biomed. Opt. Express*, Vol. 3, No. 3, pp. 572–589, Mar 2012.

[Nadl 13] Z. Nadler, B. Wang, G. Wollstein, J. E. Nevins, H. Ishikawa, L. Kagemann, I. A. Sigal, R. D. Ferguson, D. X. Hammer, I. Grulkowski, J. J. Liu, M. F. Kraus, C. D. Lu, J. Hornegger, J. G. Fujimoto, and J. S. Schuman. "Automated lamina cribrosa microstructural segmentation in optical coherence tomography scans of healthy and glaucomatous eyes". *Biomed. Opt. Express*, Vol. 4, No. 11, pp. 2596–2608, Nov 2013.

[Nick 08] J. Nickolls, I. Buck, M. Garland, and K. Skadron. "Scalable parallel programming with CUDA". *Queue*, Vol. 6, No. 2, pp. 40–53, Mar/Apr 2008.

[Noce 80] J. Nocedal. "Updating quasi-Newton matrices with limited storage". *Mathematics of computation*, Vol. 35, No. 151, pp. 773–782, Jul 1980.

[Noce 99] J. Nocedal and S. Wright. *Numerical optimization*. Springer, 1999.

[Papp 12] R. R. Pappuru, C. Briceno, Y. Ouyang, A. C. Walsh, and S. R. Sadda. "Clinical Significance of B-Scan Averaging With SD-OCT". *Ophthalmic surgery, lasers and imaging: the official journal of the International Society for Imaging in the Eye*, Vol. 43, No. 1, p. 63, Jan-Feb 2012.

[Parz 62] E. Parzen. "On Estimation of a Probability Density Function and Mode". *The Annals of Mathematical Statistics*, Vol. 33, No. 3, pp. 1065–1076, 09 1962.

[Pirc 07] M. Pircher, B. Baumann, E. Götzinger, H. Sattmann, and C. K. Hitzenberger. "Simultaneous SLO/OCT imaging of the humanretina with axial eye motion correction". *Opt. Express*, Vol. 15, No. 25, pp. 16922–16932, Dec 2007.

[Pots 08] B. Potsaid, I. Gorczynska, V. J. Srinivasan, Y. Chen, J. Jiang, A. Cable, and J. G. Fujimoto. "Ultrahigh speed spectral / Fourier domain OCT ophthalmic imaging at 70,000 to 312,500 axial scans per second". *Opt Express*, Vol. 16, No. 19, pp. 15149–69, Sep 2008.

[Pova 09] B. Povazay, B. Hofer, C. Torti, B. Hermann, A. R. Tumlinson, M. Esmaeelpour, C. A. Egan, A. C. Bird, and W. Drexler. "Impact of enhanced resolution, speed and penetration on three-dimensional retinal optical coherence tomography". *Opt Express*, Vol. 17, No. 5, pp. 4134–50, Mar 2009.

[Ricc 09] S. Ricco, M. Chen, H. Ishikawa, G. Wollstein, and J. Schuman. "Correcting Motion Artifacts in Retinal Spectral Domain Optical Coherence Tomography via Image Registration". *Medical Image Computing and Computer-Assisted Intervention - Miccai 2009, Pt I, Proceedings*, Vol. 5761, pp. 100–107, 2009.

[Rior 08] P. Riordan-Eva and J. Whitcher. *Vaughan & Asbury's general ophthalmology*. Wiley Online Library, 2008.

[Ryan 13] S. J. Ryan. *Retina*. Saunders/Elsevier, London, 5th Ed., 2013.

[Saka 08] A. Sakamoto, M. Hangai, and N. Yoshimura. "Spectral-Domain Optical Coherence Tomography with Multiple B-Scan Averaging for Enhanced Imaging of Retinal Diseases". *Ophthalmology*, Vol. 115, No. 6, pp. 1071 – 1078.e7, Jun 2008.

[Schm 97] J. M. Schmitt. "Array detection for speckle reduction in optical coherence microscopy". *Physics in Medicine and Biology*, Vol. 42, No. 7, p. 1427, Jul 1997.

[Schm 99a] L. Schmetterer and M. Wolzt. "Ocular blood flow and associated functional deviations in diabetic retinopathy". *Diabetologia*, Vol. 42, No. 4, pp. 387–405, Apr 1999.

[Schm 99b] J. M. Schmitt, S. H. Xiang, and K. M. Yung. "Speckle in Optical Coherence Tomography". *Journal of Biomedical Optics*, Vol. 4, No. 1, pp. 95–105, Jul 1999.

[Schu 95] J. S. Schuman, M. R. Hee, C. A. Puliafito, C. Wong, T. Pedut-Kloizman, C. P. Lin, E. Hertzmark, J. A. Izatt, E. A. Swanson, and J. G. Fujimoto. "Quantification of nerve fiber layer thickness in normal and glaucomatous eyes using optical coherence tomography: a pilot study". *Archives of Ophthalmology*, Vol. 113, No. 5, pp. 586–596, May 1995.

[Seba 12] V. Sebastián, S.-K. Lee, C. Zhou, M. F. Kraus, J. G. Fujimoto, and K. F. Jensen. "One-step continuous synthesis of biocompatible gold nanorods for optical coherence tomography". *Chemical Communications*, Vol. 48, No. 53, pp. 6654–6656, May 2012.

[Sebe 00] N. Sebe, M. S. Lew, and D. P. Huijsmans. "Toward improved ranking metrics". *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 22, No. 10, pp. 1132–1143, Oct 2000.

[Swan 93] E. A. Swanson, J. A. Izatt, C. P. Lin, J. G. Fujimoto, J. S. Schuman, M. R. Hee, D. Huang, and C. A. Puliafito. "In vivo retinal imaging by optical coherence tomography". *Opt. Lett.*, Vol. 18, No. 21, pp. 1864–1866, Nov 1993.

[Szku 12] M. Szkulmowski, I. Gorczynska, D. Szlag, M. Sylwestrzak, A. Kowalczyk, and M. Wojtkowski. "Efficient reduction of speckle noise in optical coherence tomography". *Opt. Express*, Vol. 20, No. 2, pp. 1337–1359, Jan 2012.

[Tan 12] O. Tan, Y. Li, Y. Wang, M. F. Kraus, J. J. Liu, B. Potsaid, B. Baumann, J. G. Fujimoto, and D. Huang. "Speckle reduction in swept source optical coherence tomography images with slow-axis averaging". In: *Proc. SPIE*, pp. 82132Z–82132Z–8, 2012.

[Thev 00] P. Thevenaz, T. Blu, and M. Unser. "Interpolation revisited [medical images application]". *Medical Imaging, IEEE Transactions on*, Vol. 19, No. 7, pp. 739–758, July 2000.

[Toll 09] D. A. Tolliver, H. Ishikawa, J. S. Schuman, and G. L. Miller. "An inpainting method for combining multiple SD-OCT scans with applications to Z-motion recovery, noise reduction and longitudinal studies". In: *ARVO 2009*, p. 1100, May 2009.

[Tsai 11a] T.-H. Tsai, B. Potsaid, M. F. Kraus, C. Zhou, Y. K. Tao, J. Hornegger, and J. G. Fujimoto. "Piezoelectric-transducer-based miniature catheter for ultrahigh-speed endoscopic optical coherence tomography". *Biomed. Opt. Express*, Vol. 2, No. 8, pp. 2438–2448, Aug 2011.

[Tsai 11b] T.-H. Tsai, B. M. Potsaid, M. Kraus, J. J. Liu, C. Zhou, J. Hornegger, and J. G. Fujimoto. "Piezoelectric transducer based miniature catheter for ultrahigh speed endoscopic optical coherence tomography". In: *Proc. SPIE*, pp. 788919–788919–6, 2011.

[Tsai 13a] T.-H. Tsai, B. Potsaid, Y. K. Tao, V. Jayaraman, J. Jiang, P. J. S. Heim, M. F. Kraus, C. Zhou, J. Hornegger, H. Mashimo, A. E. Cable, and J. G. Fujimoto. "Ultrahigh speed endoscopic optical coherence tomography using micromotor imaging catheter and VCSEL technology". *Biomed. Opt. Express*, Vol. 4, No. 7, pp. 1119–1132, Jul 2013.

[Tsai 13b] T.-H. Tsai, Y. K. Tao, B. M. Potsaid, V. Jayaraman, M. F. Kraus, P. J. S. Heim, J. Hornegger, H. Mashimo, A. E. Cable, and J. G. Fujimoto. "Ultrahigh speed endoscopic optical coherence tomography using micromotor imaging catheter and VCSEL technology". In: *Proc. SPIE*, pp. 85710N–85710N–6, 2013.

[Velt 06] M. E. van Velthoven, M. H. van der Linden, M. D. de Smet, D. J. Faber, and F. D. Verbraak. "Influence of cataract on optical coherence tomography image quality and retinal thickness". *Br J Ophthalmol*, Vol. 90, No. 10, pp. 1259–62, Oct 2006.

[Wang 09] Y. Wang, A. Lu, J. Gil-Flamer, O. Tan, J. A. Izatt, and D. Huang. "Measurement of total blood flow in the normal human retina using Doppler Fourier-domain optical coherence tomography". *Br J Ophthalmol*, Vol. 93, No. 5, pp. 634–7, May 2009.

[Wang 13] B. Wang, J. E. Nevins, Z. Nadler, G. Wollstein, H. Ishikawa, R. A. Bilonick, L. Kagemann, I. A. Sigal, I. Grulkowski, J. J. Liu, M. Kraus, C. D. Lu, J. Hornegger, J. G. Fujimoto, and J. S. Schuman. "In vivo lamina cribrosa micro-architecture in healthy and glaucomatous eyes as assessed by optical coherence tomography". *Invest Ophthalmol Vis Sci*, Vol. 54, No. 13, pp. 8270–4, Dec 2013.

[Wang 14a] B. Wang, J. E. Nevins, Z. Nadler, G. Wollstein, H. Ishikawa, R. A. Bilonick, L. Kagemann, I. A. Sigal, I. Grulkowski, J. J. Liu, M. Kraus, C. D. Lu, J. Hornegger, J. G. Fujimoto, and J. S. Schuman. "Reproducibility of In-Vivo OCT Measured Three-Dimensional Human Lamina Cribrosa Microarchitecture". *PLoS One*, Vol. 9, No. 4, p. e95526, Apr 2014.

[Wang 14b] Z. Wang, H.-C. Lee, O. O. Ahsen, B. Lee, W. Choi, B. Potsaid, J. Liu, V. Jayaraman, A. Cable, M. F. Kraus, K. Liang, J. Hornegger, and J. G. Fujimoto. "Depth-encoded all-fiber swept source polarization sensitive OCT". *Biomedical Optics Express*, Vol. 5, No. 9, pp. 2931–2949, Sep 2014.

[Webb 87] R. H. Webb, G. W. Hughes, and F. C. Delori. "Confocal scanning laser ophthalmoscope". *Appl Opt*, Vol. 26, No. 8, pp. 1492–1499, Apr 1987.

[Webb 90] R. Webb. "Scanning Laser Ophthalmoscope". In: B. Masters, Ed., *Noninvasive Diagnostic Techniques in Ophthalmology*, pp. 438–450, Springer New York, 1990.

[Whit 03] B. White, M. Pierce, N. Nassif, B. Cense, B. Park, G. Tearney, B. Bouma, T. Chen, and J. de Boer. "In vivo dynamic human retinal blood flow imaging using ultra-high-speed spectral domain optical coherence tomography". *Optics Express*, Vol. 11, No. 25, pp. 3490–7, Dec 2003.

[Wies 10] W. Wieser, B. R. Biedermann, T. Klein, C. M. Eigenwillig, and R. Huber. "Multi-Megahertz OCT: High quality 3D imaging at 20 million A-scans and 4.5 GVoxels per second". *Opt. Express*, Vol. 18, No. 14, pp. 14685–14704, Jul 2010.

[Wolf 69] P. Wolfe. "Convergence Conditions for Ascent Methods". *SIAM Review*, Vol. 11, No. 2, pp. 226–235, Apr 1969.

[Wong 10] A. Wong, A. Mishra, K. Bizheva, and D. A. Clausi. "General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery". *Opt. Express*, Vol. 18, No. 8, pp. 8338–8352, Apr 2010.

[Xian 10] Z. Xiang and P. Milanfar. "Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content". *Image Processing, IEEE Transactions on*, Vol. 19, No. 12, pp. 3116–3132, Dec 2010.

[Xian 98] S. H. Xiang, L. Zhou, and J. M. Schmitt. "Speckle noise reduction for optical coherence tomography". In: *Proc. SPIE*, pp. 79–88, Jan 1998.

[Zawa 07] R. J. Zawadzki, A. R. Fuller, S. S. Choi, D. F. Wiley, B. Hamann, and J. S. Werner. "Correction of motion artifacts and scanning beam distortions in 3D ophthalmic optical coherence tomography imaging". In: *Proc. SPIE*, pp. 642607–642607–11, Mar 2007.

[Zito 03] B. Zitová and J. Flusser. "Image registration methods: a survey". *Image and Vision Computing*, Vol. 21, No. 11, pp. 977–1000, Oct 2003.