Knowledge Mining
In The Clinical Assessment
Of Glaucoma

Haogang Zhu

A thesis submitted
for the degree of
Doctor of Philosophy

City University London
School of Community and Health Sciences

December 2010

Table of Contents

List of Figures and Tables

Acknowledgements

Declaration

This thesis has been completed solely by the candidate, Haogang Zhu. The work contained in it was done by the candidate.

It has not been submitted for any other degrees, either now or in the past. Where work contained within it has been previously published, this has been stated in the text.

All sources of information have been acknowledged and references have been given.

The University Librarian of City University London has the powers of discretion to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

Glaucoma is a leading cause of irreversible blindness and visual impairment. In the clinic, glaucomatous damage can be characterized by structural changes in the optic nerve head (ONH) and retinal nerve fibre layer (RNFL) that can be evaluated by various retinal-imaging techniques such as scanning laser polarimetry and optical coherence tomography (OCT). The structural damage can lead to functional damage in the visual field (VF), normally assessed with standard automated perimetry, which assesses the differential light sensitivity in the field of view. The clinical measurements of retinal structure and visual function play an important role in the detection and management of glaucoma, but the data generated is often complex and highly variable, thus making it hard to clinically interpret. The purpose of this thesis was to investigate *knowledge mining* procedures for extracting clinically useful information from these measurements. *Knowledge mining* describes iterative divide-and-conquer type analyses of large-scale questions: solutions to individual smaller problems are used to generate better quality knowledge, which in the case of work reported in this thesis can be translated into clinically useful analysis tools. This thesis describes five *knowledge mining* procedures specifically developed and applied to structural and functional measurements in glaucoma: (1) probabilistic inference to aid image acquisition of OCT images; (2) a robust and efficient segmentation algorithm to extract features of retina tissue layer structures in large-scale 3-dimensional image volumes acquired by OCT; (3) a predictive structure-function relationship model to bridge the retinal structure and visual function measurements in glaucoma; (4) quantification and visualization of structure-function discordance using the model about structure-function relationship; (5) integration of structural and functional measurements to improve the reproducibility of the measurements. In conclusion the *knowledge mining* approaches improved the acquisition and/or accuracy of the measurements and provide new clinical analysis tools to detect and manage glaucoma.

List of Abbreviations

ANN	Artificial Neural Networks
BBS	Broadband Source
BRBF	Bayesian Radial Basis Function
CAG	Closed-Angle Glaucoma
c-SLO	Confocal Scanning Laser Ophthalmoscope
CV	Coefficient of Variation
ECC	Enhanced Corneal Compensation
EM	Expectation-Maximization
GHT	Glaucoma Hemifield Test
GP	Gaussian Process
HFA	Humphrey Field Analyzer
HRT	Heidelberg Retina Tomograph
IOP	Intraocular Pressure
KDD	Knowledge Discovery in Databases
MAD	Mean Absolute Difference
MAAD	Mean Absolute Angular Difference
MD	Mean Deviation
MLP	Multilayer Perceptron
NFI	Nerve Fibre Indicator
NTG	Normal Tension Glaucoma
NRR	Neuroretinal Rim
OAG	Open-Angle Glaucoma
OCT	Optical Coherence Tomography
OHT	Ocular Hypertension
ONH	Optic Nerve Head
PCA	Principal Component Analysis
PDM	pattern discordance map
POAG	Primary Open-Angle Glaucoma
PSD	Pattern Dtandard Deviation
TD-OCT	Time Domain Optical Coherence Tomography
VF	Visual Field
RANSAC	Random Sample Consensus
RBF	Radial Basis Functions
RGCs	Retinal Ganglion Cells
RPE	Retinal Pigment Epithelium
RNFL	Retinal Nerve Fibre Layer
RNFLT	Retinal Nerve Fibre Layer Thickness
SAP	Standard Automated Perimetry
SD	Standard Deviation
SD-OCT	Spectral Domain Optical Coherence Tomography
SFDI	Structure-Function Discordance Index
SITA	Swedish Interactive Threshold Algorithm
SLP	Scanning Laser Polarimetry
VCC	Variable Corneal Compensator
VFI	Visual field Index

Chapter 1 : Introduction

1.1. A brief introduction to glaucoma

Glaucoma is a leading cause of irreversible blindness in the world (Thylefors et al., 1995, Quigley and Broman, 2006). This condition is defined as a group of irreversible chronic, progressive diseases of the optic nerve head characterized by the degeneration of retinal ganglion cells and their axons (Quigley et al., 1982). In the clinic, glaucomatous damage can be characterized by structural change in the optic nerve head and retinal nerve fibre layer (Quigley et al., 1982). This structural damage can lead to functional damage, which manifests as partial to full loss of the visual field, which is the portion of space where objects are simultaneously visible in a steadily fixating eye (Harrington, 1976). Visual field loss can be asymptomatic in the early stages of glaucoma and the patient is not aware of this loss. However, if left untreated, visual field damage can worsen and lead to complete loss of vision.

1.1.1. Risk factors and classification of glaucoma

The cause of glaucomatous damage is closely associated with aqueous humour, which is the clear watery fluid circulating through the anterior chamber. Simply described, aqueous humour 'nourishes' the area around the iris and the cornea, and it maintains the shape of the eye by exerting appropriate amount of pressure, which is known as the intraocular pressure (IOP). IOP is maintained under a dynamic equilibrium of the aqueous humour flow: the fluid is continuously produced, while this inflow is offset by an outflow via the drainage between the iris and cornea. About 80 to 90% of the outflow is through a sponge-like substance known as the trabecular meshwork, and the remaining fluid drains independently through the uveoscleral pathway (Hitchings, 2000).

Elevated IOP has been evidenced as a key risk factor for glaucomatous damage, (Pohjanpelto and Palva, 1974, Leske, 1983) although the underlying mechanisms have not yet been fully elucidated. A complete causality between raised IOP and glaucoma has not yet been found. Although an increased prevalence of glaucoma should be expected in eyes with elevated IOP, only 10% of patients with ocular hypertension (OHT; IOP > 21mmHg) develop glaucomatous visual field damage, and glaucoma can also occur in eyes with normal IOP (IOP<21mmHg; normal tension glaucoma (NTG)) (Sommer et al., 1991). This finding could be explained by the variability in susceptibility to IOP of different eyes: some may require a much higher IOP level to develop damage, while others are more vulnerable to IOP variation and therefore develop significant damage at a much lower IOP level.

Glaucoma can be classified as primary open-angle glaucoma (POAG), closed-angle glaucoma (CAG) or congenital glaucoma. It can also be differentiated into primary and secondary glaucoma, where the glaucomatous damage results from other conditions such as other ocular or orbital diseases. Normal tension glaucoma (NTG: IOP<21mmHg) is a subdivision of POAG. In CAG, the iris is pushed against the trabecular meshwork, which then closes off the drainage angle and increases the IOP. CAG is more common in Asian populations, and accounts for up to 50% of glaucoma worldwide (Thylefors et al., 1995, Quigley and Broman, 2006). Congenital glaucoma is relatively rare and is characterized by malformation of the aqueous drainage route. POAG is the most common type of glaucoma, especially in European and North American populations. The measurements of glaucomatous damage used in this thesis were all conducted in patients with POAG.

Other important risk factors for POAG include age (Tielsch et al., 1991, Klein et al., 1992, Coffey et al., 1993, Dielemans et al., 1994, Leske et al., 1994, Mitchell et al., 1996), race (Tielsch et al., 1991, Klein et al., 1992, Coffey et al., 1993, Dielemans et al., 1994, Leske et al., 1994, Mitchell et al., 1996), genes (Libby et al., 2005), diabetes (Leske, 1983) and elevated blood pressure (Tielsch et al., 1995). In addition to these risk factors, those for NTG also include being of Japanese descent (Shiose et al., 1991), and those for CAG also include sex and being of Asian or African descent (Foster and Johnson, 2000). All the risk factors for glaucoma have been comprehensively reviewed elsewhere (Hitchings, 2000).

1.1.2. Structural changes relevant to glaucoma

In order to understand glaucoma, it is worthwhile briefly considering how the eye functions, especially within the retina. A photograph of a glaucomatous retina (left eye) and an illustration of retinal histology are given in Figure 1-1. For ease of description, the retina is divided into nasal and temporal regions horizontally, and into superior and inferior regions vertically.

Upon entering the eye, light travels through the anterior parts of the eye until it reaches the retina. The retina consists of several layers (Figure 1-1b), and light passes through these layers (from top to bottom; Figure 1-1b) until it reaches the photoreceptor layer containing cells called cones and rods. Rods become active at low levels, while cones are active at high levels, of luminance. At the photoreceptor layer, light elicits a chemical transformation mediating a propagation of signal upwards, back to the retinal ganglion cells (RGCs) which ultimately produce action potentials on their axons. These axons form the retinal nerve fibre layer (RNFL); bundles of retinal nerve fibres can be seen at the top of Figure 1-1a. Nerve fibre bundles spread across the retina and exit the eye to enter the brain via the optic nerve head (ONH), also called the optic disc. They carry signals from different locations of the retina in a spatiotemporal

pattern. The ONH is usually known as the 'blind spot' because it does not contain any photoreceptor. Central vision resolution is high in the macula, which is the centre of the retina (Figure 1-1a) and has a high concentration of RGCs (Bennett and Rabbetts, 1998).
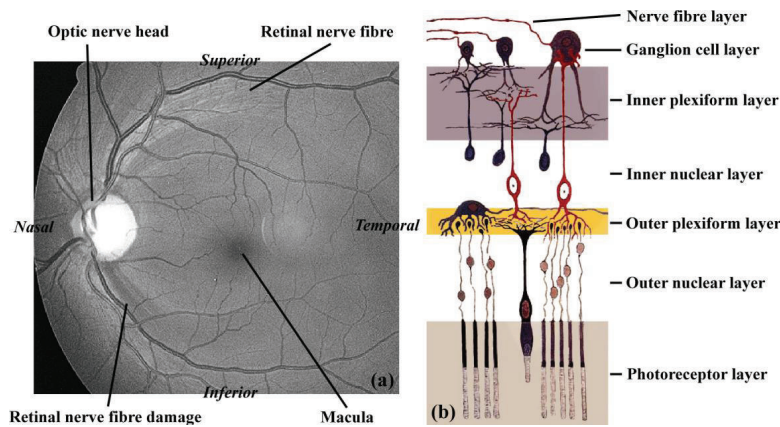
Figure 1-1. (a) Photograph of a retina with glaucomatous damage; (b) illustration of retinal histology. Retina histology figure was reproduced from http://en.wikipedia.org/wiki/File:Fig_retine.png (accessed 17th December 2010).

In general, the retinal nerve fibres coming from RGCs located closer to the ONH lie underneath those from RGCs located more peripherally (Minckler, 1980). Therefore, RNFL thickness (RNFLT) increases from the periphery of the retina to the area between the ONH and the macula. In healthy eyes, the RNFL is relatively thicker and more visible (Figure 1-1a) in the superior and inferior areas compared with nasal and temporal areas. Moreover, nerve fibres converge at the ONH with regional variation: those coming from the nasal, superior and inferior retina have a relatively straight path towards the ONH, while nerve fibres coming from the temporal regions of the retina describe an arcuate path around the macula (Radius and Anderson, 1979). As shown later, this bundle of nerve fibres across the retina becomes important when attempts are made to correlate the visual function at a particular location of the retina with the RNFL defect.

RNFLT is currently a key parameter in the clinical examination of glaucomatous eyes. This is because retinal nerve fibres carry the information from the eye to the brain, and any damage to the RNFL due to glaucoma will cause a loss of this information. In the eye with glaucoma, the RNFL becomes thinner and this thinning can be focal or diffuse (Tuulonen and Airaksinen, 1991). Focal defects appear as dark bands, which are broader towards the periphery (Figure

1-1a), and are usually associated with localized visual field damage. Diffused defects are characterized by a general thinning of the RNFL, which occurs in a wider range of retinas. RNFL defects can be quantified by imaging techniques that measure certain surrogates of RNFLT (see later). An initial glaucomatous abnormality can be focal and/or diffused RNFL thinning. As RNFL loss progresses, the defects become wider and deeper.

All the nerve fibre bundles exit the eye at the ONH, which is usually slightly tilted so that the temporal nerve fibres follow an obtuse angle, and the nasal ones turn acutely into the ONH. The nerve fibre bundles only occupy the outer portion of the ONH, which forms the neuroretinal rim (NRR). The centre of the ONH does not contain nerve fibres, and is occupied by the central retinal vessels and their connective tissues. This central part of the ONH is clinically defined as 'optic disc cup' (Weinreb and Khaw, 2004). Because of the topography of RNFLT, the rim area is usually the widest in the inferior (I) areas, followed by the superior (S), nasal (N) and temporal (T) areas. This is also referred as the 'ISNT rule'. Generally, the ONH cup area is directly influenced by the size of the ONH: a larger ONH gives rise to larger cup (Jonas et al., 1988).

Because the NRR is mainly composed of retinal nerve fibres, the rim area should be a good indicator of the number of nerve fibres and the level of the damage. Nerve fibre damage in glaucoma may affect both the shape and the area of the rim. The ISNT rule is clinically helpful because any deviation from this rule may indicate the suspicion of glaucomatous damage. Moreover, rim loss is usually accompanied by cup size increase. There are various patterns of rim shape changes caused by glaucoma and these have been reviewed elsewhere (Tuulonen and Airaksinen, 1991).

1.1.3. Visual field in glaucoma

Structural damage of the retina is usually associated with visual field defect. Visual field is defined as the area from which light can enter the eye and reach the retina (Henson, 2000, Tate and Lynn, 1977, Werner, 1991). The central visual field corresponds to the macula, and this is the point where people look directly during fixation. With a bright stimulus, a healthy visual field can extend to 60° superior, 75° inferior, 100° temporal and 60° nasal from the line of sight (central vision) (Henson, 2000). Visual field damage is described as the reduced or non-existent vision in a certain direction when the patient's central vision is fixated on a given point (fixation point). With a damaged visual field, the patient is not able to see an object presented in the direction of damage.

Visual field assessment has been used for a long time as a clinical tool for glaucoma management. Visual fields are typically measured by perimetry (Lachenmayr and Vivell, 1993), which will be introduced in detail in the next section.

1.1.4 Clinical management of glaucoma

Glaucoma is characterized by a chronic, progressive degeneration of the optic nerve, which means that the disease typically continues to progress without treatment or without adequate treatment. Structural assessment of ONH changes and visual function measurements (see next section), are clinically useful tools for glaucoma diagnosis and management.

In addition to detecting and confirming glaucomatous damage, it is essential to determine the rate at which the damage is worsening so that decisions on clinical management can be made and the effectiveness of treatment can be determined. Clinical evidence from randomized clinical trials in this area, however, has been scarce. For ethical reasons, apart from findings in patients at very early stages of glaucoma (Heijl et al., 2002, Hyman et al., 2010), or in those with NTG (McNaught et al., 1995), only limited data are available on the natural history of glaucoma. The best evidence to date for the beneficial effect of medical IOP-lowering treatment in open-angle glaucoma (OAG) with established VF loss comes from the Early Manifest Glaucoma Trial (EMGT) (Heijl et al., 2002, Leske et al., 2003, Hyman et al., 2010): a randomised, controlled clinical trial designed to evaluate the efficacy of IOP-lowering treatment on the progression of OAG. Results indicated that on average, treatment reduced IOP by 5.1mmHg or 25%. Furthermore, progression was less frequent in the treatment group than in controls; 45% and 62% respectively, and occurred significantly later in treated patients. Interestingly, in the context of the aim of this thesis, although both structural and functional measurements were frequently taken throughout the trial, little was done to combine the evidence from this data. It is important to note that no universally recognised methodology exists for monitoring glaucoma progression over time.

Clinical observations suggest that change due to glaucoma is typically a slow process, which happens over several years (Hitchings, 2000). On the whole, treatment is effective for managing the majority of patients; a discussion of different forms of medical and surgical treatment is beyond the scope of this work and is described elsewhere (Hitchings, 2000). Yet in the case of a fast-progressing eye, more intensified treatment needs to be prescribed to attempt to arrest deterioration to preserve the patients' visual function and quality of life. It also might be that a patient requires less treatment in cases where measurable stability of functional loss can be established. Therefore, a real challenge, and the motivation underpinning this thesis, is to provide clinicians with tools that allow them to confirm true progression or stability under

treatment. These tools should be both sensitive and specific, to avoid over-treating patients due to false-positive diagnosis, or under-treating patients due to false-negative diagnosis. Energy devoted to the development of new tools for measuring the structural and functional impact of glaucoma in an individual is also worthwhile because the measurements are mandatory for clinical trials as outcome measures and evidence for effectiveness of treatment.

1.2. Structural and functional measurements of glaucoma

This section provides an introduction to some of the techniques that are used for structural and functional measurements in the clinical management of glaucoma. The studies reported in this thesis uses data acquired from these techniques.

1.2.1. Optical coherence tomography (OCT)

Optical coherence tomography (OCT) (Huang et al., 1991, Sampson and Hillman, 2004) is an imaging technique that can acquire high-resolution cross-sectional retinal images of two or three dimensions (Bouma and Tearney, 2001). It combines lateral-point beam-scanning with the depth-sectioning ability of low-coherence interferometry. Because the resolution in depth (axial) depends almost entirely on the optical bandwidth of the light source (instead of an aperture of the optical system), OCT can combine the benefits of high axial resolution and large axial depths. Because of this advantage, the primary applications of OCT were for *in vivo* imaging through thick layers of biological tissues. It was not until the early 1990s that OCT with micrometre-resolution was used for the first time in the biomedical field to examine the structure of the retina cross-sectionally (Huang et al., 1991).

Figure 1-2 shows a typical setup of a time-domain OCT (TD-OCT) system, the name of which will be explained when the spectral-domain OCT is introduced later. The system consists of a low-coherence, broadband (large range of optical wavelengths) light source and a Michelson interferometer (Bouma and Tearney, 2001). The broadband source (BBS) is typically characterized by the coherence length and is usually linked to a series of very short light pulses defined by this coherence length. The Michelson interferometer splits the light from the BBS into two arms: the reference arm implemented by a mirror that translates back and forth and the sample arm where the light is weakly focussed into the sample to be imaged. The reflected light from both arms is re-combined at the interferometer where constructive interference occurs between the reflected reference light and the backscattered sample light only if the optical path lengths of them differ by less than the coherence length of the BBS. An axial scan (A-scan; scan in depth) is acquired by moving the reference mirror while detecting the interference signal, and a transverse scan (B-scan) on the sample is achieved via the rotation of a sample arm galvanometer mirror. The B-scans can then be used to compose two- or three-dimensional

images. More comprehensive details about OCT system can be found in Bouma et al (Bouma and Tearney, 2001).
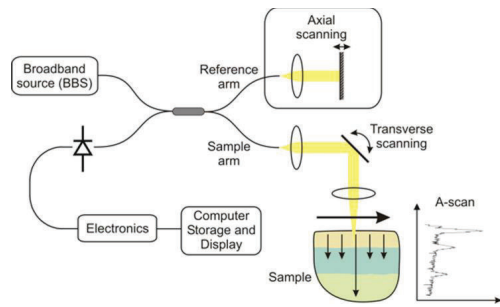
Figure 1-2. Schematic representation of an OCT system setup. The figure was reproduced from http://obel.ee.uwa.edu.au/research/oct/intro/ **(accessed 10th August 2009)**

When used in retinal imaging, OCT forms an intensity map of light backscattered or reflected from the retina at different locations and depths. The retinal layers in the tomography image can be differentiated because different tissues in the layers of the retina reflect light with different intensities. A widely used TD-OCT system for retinal imaging in clinical applications is the StratusOCT (Figure 1-3; Carl Zeiss Meditec Inc., Dublin, CA, USA). The axial resolution of the StratusOCT has been reported to be $\leq 10\mu m$, and the theoretically highest available lateral resolution is about $20\mu m$ determined by the imaging beam size at the retina. It can acquire 400 A-scans per second.

Figure 1-3. The StratusOCT system. The figure was reproduced from http://www.meditec.zeiss.com/C125679E0051C774/Contents-Frame/2BF7095D5578B41D882572430063F9AD **(accessed 30th Nov 2010).**

For image acquisition, the StratusOCT uses various scan patterns (protocols), each of which has been optimized for the detection of certain diseases. Among them, a circular scanning protocol referred to as 'RNFLT (3.4)' is the most widely used protocol for glaucoma. Each RNFLT (3.4) circular scan is carried out under a scan circle that is 3.4mm in diameter, and is centred on the ONH. It images the peripapillary retina with 512 consecutive A-scans around the ONH. This

process is repeated three times and the three images are averaged in order to reduce noise and other possible variability. Because each circular scan takes more than 1 second (the speed of the StratusOCT is 400 A-scans per second), the three scans have to be taken separately to allow for blinks between two scans. A slight variation of RNFLT (3.4) is Fast RNFLT (3.4), for which each circular scan requires 256 A-scans around the ONH, and is thus twice as fast as RNFLT (3.4) therefore allowing the three scans to be repeated successively within 1.92 seconds. Each Fast RNFLT (3.4)/RNFLT (3.4) scan provides an image on a cylinder surface around the ONH. To visualize and analyse this image, the 'cylinder' surface is 'cut' at the mid-temporal location and is straightened into a 2D image as shown in Figure 1-4.
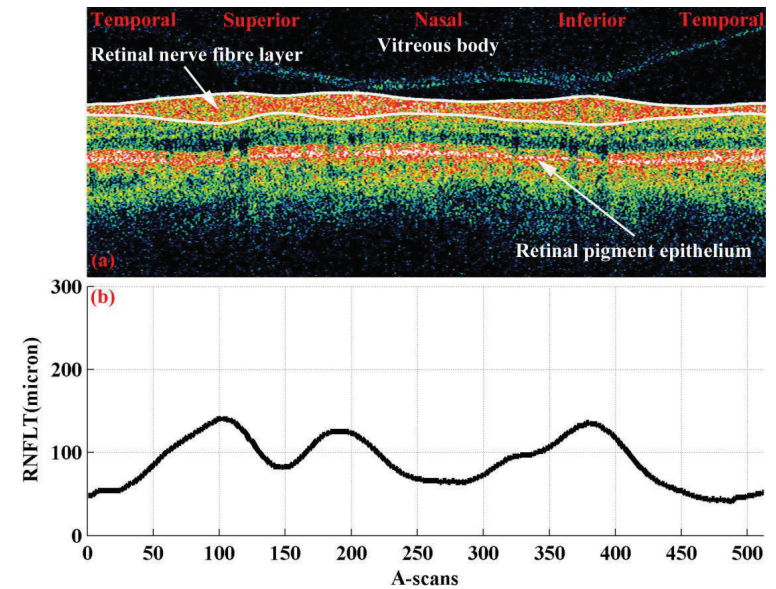
Figure 1-4. A tomography image obtained with StratusOCT from a subject (see chapter 2). (a) intensity map acquired with the RNFLT (3.4) protocol (the RNFL anterior and posterior boundaries segmented by the StratusOCT software are indicated by two white lines); (b) the corresponding RNFLT in (a).

In Figure 1-4a, the image starts from the middle point of the temporal (T) area and, from left to right, extends to the superior (S), nasal (N), inferior (I) and lower temporal (T) areas. This 'T-S-N-I-T' type of image acquisition is adopted by many types of instruments used in glaucoma, and the measurement results (e.g. RNFLT) are usually referred as 'TSNIT profile'. The reflection intensity is plotted as logarithmic values on a false colour-coded scale: hotter colour (more red

and orange) means higher reflection intensity; colder colour (more blue and green) indicates lower reflection intensity. There are two sections with significant reflection: RNFL on the top of the retina, and retinal pigment epithelium (RPE) at the back of the retina. The RNFL can be segmented by algorithms separating the highly reflective area on the top of the retina from the area with weak reflection. The segmentation result (anterior and posterior boundaries of RNFL) from the StratusOCT software is shown by the two white lines in Figure 1-4a. The RNFLT, as the target measured variable, is calculated as the distance between the two boundaries of the RNFL and is shown in Figure 1-4b. Note that the TSNIT profile of RNFLT features a 'double-hump' shape (relatively thicker RNFLT in the superior and inferior areas) according to the ISNT rule. However, because OCT only measures the intensity surrogate of the RNFL, the RNFLT measurement from it depends on various factors including optical properties of tissues, image resolution, pre-retinal factors such as media opacities, and other factors (Sung et al., 2010). Therefore, in short, RNFLT measurements from OCT may not be an exact measurement of the RNFL.

Before the acquisition of the OCT image, the scan circle is manually adjusted to be centred on the ONH as guided by a 'live' image of the retinal fundus. The location of the scan circle, however, is unknown during the image acquisition. Displacement of the scan circle position may happen due to the operator's subjective placement or because of eye movement after the manual adjustment. The inevitable displacement between tests means that the RNFLT is not necessarily sampled at the same location and thus contributes to the variability and error in the measurement of RNFLT (Gabriele et al., 2008), restricting the use of the technique especially in determining the deterioration of the RNFLT in the longitudinal assessment or follow-up of glaucoma. Therefore a method for identifying and aligning the location of the scan circle on the retina would be clinically useful. A *knowledge mining* solution for this problem is described in chapter 2.

As already described, TD-OCT acquires an A-scan at different depths by moving the reference arm. The mechanical movement of the reference mirror slows down the imaging speed, and only a few hundred A-scans can be taken within an acceptable period of time to avoid the patient's blink or fixation loss during the image acquisition process.

In contrast, another recently developed OCT technique, spectral domain OCT (SD-OCT) (Fercher et al., 1995, Wojtkowski et al., 2002, Schuman et al., 2007), does not involve mechanical movement for A-scan acquisition. In SD-OCT, the broadband interference is acquired by a group of spectrally separated detectors (e.g. a grating with a linear detector array). An A-scan can, therefore, be immediately calculated by a Fourier-transform from the acquired spectra, without moving the reference arm. The imaging speed is dramatically improved (Nassif

et al., 2004) by this feature: ~26000 A-scans per second compared with 400 A-scans per second with TD-OCT. Moreover, compared with SD-OCT, in which light is generated and detected as a 'function' of spectra, light in TD-OCT is a 'function' of time and, therefore, this type of technique is referred to as time 'domain'. More details about SD-OCT can be found elsewhere (Fercher et al., 1995, Wojtkowski et al., 2002, Schuman et al., 2007).

The improved speed in SD-OCT makes it possible to acquire a large number of A-scans in a short period of time. These A-scans form circular scans with significantly higher sample rate. They can also cover a large area on the retina, and can be used to build a map of RNFLT. For example, a SD-OCT system for clinical use, the RTVue-100 system (Figure 1-5; Optovue, Fremont, CA, USA), can acquire a three-dimensional volume of the retina with 513-by-101 A-scans (axial resolution of $5\mu m$) around the ONH and each A-scan has 768 axial sample points. This results in high 513-by-101-by-768 volumetric data with abundant information about ONH and RNFL. However, the large quantity data from SD-OCT poses great challenges on the computing power and image processing algorithms to quantify useful information in the volumetric data.

An extensive study of SD-OCT images is reported in chapter 3. A *knowledge mining* approach is used to provide a robust and efficient segmentation algorithm to process the 3D image volumes in a novel fashion, providing clinically useful and reproducible measurements of the RNFLT with this technology.

1.2.2. Scanning laser polarimetry (SLP)

The nerve fibre bundles appear as closely spaced cylindrical 'tubes' (Zhou and Knighton, 1997). RNFL exhibits substantial linear birefringence with its slow axis parallel to the direction of these tubes (Weinreb et al., 1990, Dreher et al., 1992, Zhou and Knighton, 1997): the light wave propagates more 'slowly' if its polarization direction is the same as the direction of these tubular structures (Figure 1-6). This forms the basic principle of scanning laser Polarimetry (SLP), which is a technique used to evaluate RNFLT by measuring the total phase retardation in light reflected from the retina (Weinreb et al., 1990, Dreher et al., 1992). This retardation is proportional to RNFLT, which can be calculated from it. The SLP models used on human eyes were described in detail elsewhere (Knighton and Huang, 2002a, Zhou et al., 2003).
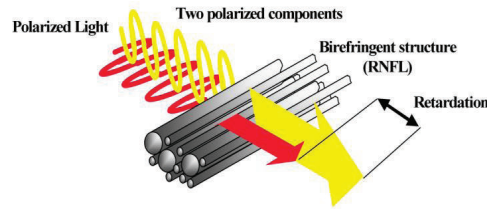
Figure 1-6. Illustration of birefringence of nerve fibre bundles (GDxVCC Instrument Manual 2004). The light with polarization direction parallel (red) to the nerve fibre tubes has a slower phase speed than the light polarized in perpendicular direction (yellow) to the tubes. The phase retardation in the reflected light is detected and is proportional to RNFLT.

Figure 1-7. The GDxVCC system. The figure was reproduced from http://www.adlereyeassociates.com/pages/services/optometric-instruments.shtml (accessed 30th Nov 2010).

A widely used clinical SLP system is the GDxVCC (Figure 1-7; Carl Zeiss Meditec Inc. Dublin, CA, USA), which is a confocal scanning laser ophthalmoscope integrated with a polarimeter. It acquires an ocular fundus image by raster-scanning the retina with a near-infrared laser beam (wavelength of 780nm). The image from GDxVCC covers 40 degree horizontally and 20 degree vertically, and this area includes both peripapillary RNFL around the ONH and the macular. In the GDxVCC system, a variable corneal compensator (VCC) (Zhou et al., 2003) was used to reduce and compensate the anterior segment birefringence (mainly corneal birefringence) (Knighton and Huang, 2002b), which is a confounder causing an overestimation of RNFLT by SLP. After the VCC compensation, an additional SLP anomaly, which is likely unrelated to corneal compensation, appears in a small proportion of scanned individuals (Bagga et al., 2005). This artefact significantly overestimates the RNFLT and became known as an atypical birefringence pattern, which is quantified with a typical scan score (TSS) in GDxVCC software.

A newer compensation scheme, called enhanced corneal compensation (ECC), was recently developed to eliminate the artefacts that are associated with the atypical birefringence pattern (Tóth and Holló, 2005). The GDxVCC technique is, however, still being used in both glaucoma clinics and research studies and data from this device is used in this thesis. This 'updating' of instrumentation and technology reflects a common limitation of newer imaging techniques that ought to be noted at this point: the data collected using old techniques are generally not transferable to the new technique, causing the lack of consistent longitudinal data. This problem is particularly significant when instituting longitudinal research studies on progression because instruments, or elements with them, might change during the follow-up.
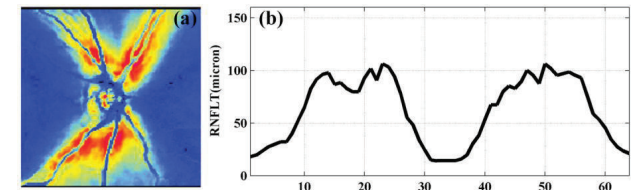
Figure 1-8. RNFLT measured by GDxVCC from a patient described in chapter 4. (a) RNFLT map (128-by-128 pixels) acquired by GDxVCC. The image is shown in the right eye direction (temporal area facing left). The hotter (more red and orange) colour denotes a thicker RNFL. (b) Peripapillary RNFLT around the ONH in a similar way as TD-OCT using the RNFLT (3.4) scan protocol. The GDxVCC RNFLT profile has 64 sectors.

An example of RNFLT measured by GDxVCC is shown in Figure 1-8. The ONH can be seen at the centre of the image. In Figure 1-8(a), warmer (more red and orange) colour stands for thicker RNFLT. GDxVCC summarises peripapillary RNFLT in a similar way as TD-OCT using

the RNFLT (3.4) protocol: a calculation annulus (inner diameter 2.4mm and outer diameter 3.2mm) is placed at the centre of the ONH which is identified manually with the help of the GDxVCC software; the annulus is then divided into 64 sectors from the mid-temporal area to the superior, nasal, inferior and lower temporal areas; the RNFLT in each sector is calculated as the mean of pixel values within this sector. The 64-sector TSNIT profile of RNFLT is straightened to a 2D plot (Figure 1-8b) to form a similar output as OCT. The significantly higher RNFLT in the superior and inferior areas (ISNT rule) is consistent with RNFL topography.

Compared with OCT technique, GDxVCC can take a retinal image in about 0.8 seconds, and no complex post-processing (e.g. segmentation in OCT) is required to form the RNFLT map. However, RNFL boundaries are not visible with GDxVCC, and this 'black box' procedure sometimes causes difficulties in understanding the images. Two quality scores come with each image from GDxVCC system: image quality score (between 0 and 10) and TSS (between 0 and 100). The images with either score lower than certain threshold should be considered unreliable. Thresholds can be determined according to the purpose and the method of the study.

More detail on GDxVCC imaging is reported in the studies linking structure and function measurement in glaucoma as presented in chapters 4, 5 and 6. In particular, *knowledge mining* techniques are developed with the aim of providing an analysis tool that uses the structure information gleaned from SLP to 'improve' the measurement of a patient's visual function. It also provides a clinically useful method for highlighting poor structural and functional measurements.

1.2.3. Confocal scanning laser ophthalmoscope (c-SLO)

Confocal scanning laser ophthalmoscope (c-SLO), as indicated by its name, is a technique to acquire images of target sample using a raster-scanning laser. Data from c-SLO will not be reported in the later chapters, although most of the subjects whose data were used in this thesis were examined using c-SLO, together with other techniques. For this reason, c-SLO is only described briefly here. More details on this technique and its quantification methods can be found elsewhere (Strouthidis and Garway-Heath, 2008).

The Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Heidelberg, Germany) has been the most successful commercial realisation of c-SLO technology. The HRT was designed to acquire three-dimensional images of the retina. In short, a c-SLO system, uses a focused laser beam to scan the retina at different locations and depths, and the amount of light reflected from each scanned point is recorded. A confocal pinhole, which is a diaphragm sitting in front of the

light detector, ensures that only light from the corresponding focal plane reaches the detector and thus contributes to the image.

In the clinic, the HRT is widely used as an instrument for the quantitative analysis of the ONH. A topography surface of the retina around the ONH is derived from the three-dimensional HRT image. This surface contains the height measurement of the retina and is the main output on which most post-analytical methods are based. The topography surface contains important information about the three-dimensional structure of the ONH and has been successfully used to evaluate the cup and rim area (Strouthidis and Garway-Heath, 2008).

1.2.4. Perimetry

Although the imaging techniques described have become popular clinical tools, visual field measurement remains as the cornerstone of clinical management of glaucoma because it is still the only way of estimating a patient's visual function. In turn, the visual field is most relevant to visual impairment, visual disability and quality of life. This makes it the most important outcome, for example, in clinical trials where a standard is required to determine if, for example, a certain treatment is working in establishing disease stability. Visual field measurement is performed by perimetry (Lachenmayr and Vivell, 1993), which is a technique measuring light difference sensitivity in the visual field. During a static perimetry test, only one eye is typically tested while another eye occluded. The patient is instructed to look directly forward at a pre-determined fixation point. A stimulus is presented at a location in the visual field at one time, and the patient is asked to indicate whether the stimulus is observed.

Perimetry tests can be conducted manually or by an automated computer program. Manual perimetry (Haley, 1987) was used before the advent of computers. Because it is very time-consuming and requires attendance of well-trained doctors or technicians, this technique is now rarely used. It is established that computerised automated perimetry has improved the evaluation of visual fields in glaucoma patients (Katz et al., 1995). Automated perimetry is 'instructed' by computer algorithms, requiring less supervision from technicians and can automatically alter stimuli properties such as duration, size and intensity. The efficiency of automated perimetry makes itself a widely used clinical tool in glaucoma diagnosis and follow-up.

Automated perimetry can be further divided into different categories according to the properties of the stimuli and background. This includes white-on-white perimetry (white stimuli on white background) and, for example, blue-on-yellow perimetry (blue stimuli on yellow background) (Wild et al., 1995), kinetic perimetry (moving stimuli) and frequency-doubling perimetry (Johnson and Samuels, 1997). Of these, white-on-white perimetry is the most widely used one in glaucoma management, and is referred to as 'standard automated perimetry' (SAP). In this

thesis, only SAP is considered and will be described in detail below. Comprehensive reviews of various types of perimetry are provided elsewhere (Lachenmayr and Vivell, 1993, Henson, 2000).

The SAP data used in the studies described in this thesis were acquired by the Humphrey Field Analyzer (HFA; Figure 1-9; Carl Zeiss Meditec Inc, Dublin, CA, USA) (Haley, 1987, Werner, 1991). It is the most commonly used instrument for perimetric tests. In this technique, a white background with a luminance of 31.5 apostilbs (A) luminance is used and the testing distance is 33 cm. Stimuli are presented against this background at varied intensity from 0.08 to 10,000 A. The differential threshold is recorded in decibel (dB) and is inversely related to the intensity of the stimulus (dB $= 10\log_{10}(10000/A)$. The aim of the test is to determine the 'weakest' stimulus intensity that one can detect at all test locations in the visual field. The weakest intensity is also known as 'contrast sensitivity threshold' or 'sensitivity'. Thus, for the brightest stimulus, a sensitivity of 0dB is recorded, and for the weakest stimulus a sensitivity of 40dB is recorded.

Figure 1-9. The HFA system. The figure was reproduced from http://www.baycollege.ca/services.html (accessed 30th Nov 2010).

Within the range from 0dB to 40dB, the sensitivity for healthy visual field is affected by multiple factors such as eccentricity and age. Sensitivity in healthy eyes decreases with eccentricity, which results in the well-known 'hill of vision' (Traquair, 1931). Sensitivity also declines, albeit very slowly (reported to be about 0.1dB/year (Heijl et al., 1991)), over time as a result of aging.

SAP tests in HFA are carried out at static predefined locations giving rise to test patterns. The most commonly used test patterns in glaucoma management are 24-2 and 30-2 (Figure 1-10). The 30-2 test pattern covers 30° of the visual field from the central vision. The test locations are 6° apart both horizontally and vertically. No test point falls on the horizontal or vertical

meridians. The 30-2 test pattern contains 74 test locations excluding the two points under the blind spot (Figure 1-10). The 24-2 test pattern is formed by a subset of test locations in 30-2 pattern by excluding some peripheral ones (Figure 1-10). It includes 52 test locations excluding the two corresponding to the area of the blind spot. The 24-2 test pattern is widely used for glaucoma management because the measurements in the peripheral locations in 30-2 is are even more variable than more central located points (Heijl et al., 1989) and for the simple reason that the test is shorter in duration. SAP tests reported in this thesis were carried out using 24-2 test pattern.
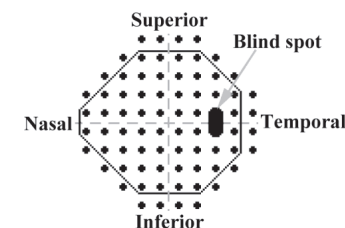
Figure 1-10. Standard 30-2 and 24-2 HFA visual field test patterns in the right eye. The black dots indicate the test locations. The 24-2 pattern contains a subset of test locations in 30-2 and is indicated by the black box. The cross point between the two dashed lines indicates central vision that corresponds to the macular on the retina. There are 74 and 52 test points in 30-2 and 24-2 patterns respectively, excluding the two locations corresponding to the area of the blind spot.

In the HFA test, the patient is instructed to look at the fixation point. An automated program instructed by a test strategy flashes one stimulus with certain brightness at one test location at a time. The patient operates a button to indicate whether the stimulus is observed or not. This procedure is repeated until the sensitivity thresholds at all test locations are found.

A test strategy computes the time and location to display a stimulus at given intensity in order to search the sensitivity thresholds quickly and reliably. Several test strategies are available in HFA. The full threshold test and Swedish interactive threshold algorithm (SITA) (Bengtsson et al., 1997) are the most widely used ones in glaucoma testing. The visual field reported in the thesis were all tested by either of these two test strategies.

The full threshold test was the previous standard and is still the benchmark to which other strategies are compared. In short, the full threshold algorithm measures the sensitivity threshold at each test location by a "repetitive staircase" procedure. At the point being measured, an initial stimulus that is brighter than the expected sensitivity is presented. If the stimulus is seen, the stimulus intensity at the same test location will be 4dB higher (dimmer). This is repeated until

the stimulus is not observed. The stimulus intensity is then increased by 2dB each time until it is observed again. On the other hand, if the initial stimulus is not seen, the process runs in reverse. The stimulus intensity decreases first in steps of 4dB (brighter) until being seen, and then it is increased by 2dB a time until not visible. The full threshold algorithm randomises the order of stimuli presentation instead of testing one location constantly. Spatially, full threshold algorithm begins with four initial locations, one in each quadrant of the visual field (Figure 1-10). Each initial point is thresholded twice. The algorithm then subsequently tests other locations. To achieve reliable measurement of sensitivity, some points (e.g. locations where sensitivity thresholds differ from the expectation by more than 4dB according to the sensitivities of the neighbouring points) are thresholded twice.

The principle of full threshold test is relatively well understood, but the repetitive staircase slows down the test procedure. SITA (Bengtsson et al., 1997) was designed to provide similar measurement reliability as full threshold test but with a lower cost of time. This is achieved, in part, by using the prior knowledge of the neighbouring test locations and the sensitivity distribution from other patients under a Bayesian framework. On average a test using SITA can be performed in about half the time it takes to perform a full threshold test (Bengtsson and Heijl, 1998, Shirato et al., 1999); this is achieved with no real increase in measurement variability compared with the full threshold test, although this has been debated by some reports (Shirato et al., 1999, Wild et al., 1999, Bengtsson and Heijl, 1999a, Sharma et al., 2000, Artes et al., 2002).

As a perimetry test depends on the subjective input from the patient, it is necessary to evaluate the reliability of the responses. Three quality parameters are examined during a perimetry test: false positive rate, false negative rate and fixation loss rate: 1) The full threshold test randomly 'cheats' the patient by making the noise as if it is presenting a stimulus without presenting a stimulus. A false positive response is recorded if the patient indicates an observable stimulus when HFA 'cheats' the patient (catch trial). 2) A stimulus that is much brighter than the sensitivity is randomly presented at a location where the threshold is already determined. A false negative response is recorded if the patient fails to respond to this stimulus presentation. 3) Stimuli are presented at the location of the physiological blind spot. A fixation loss is recorded if the patient responds to the stimuli in blind spot. SITA estimates these reliability 'measures' differently and this contributes to the saving in test time. For example, false positive rate is estimated using response time to the presentation of a stimulus (Bengtsson et al., 1997) and, therefore, the method does not require any additional catch trials. Perimetry tests with high false positive rate, false negative rate or fixation loss rate should be considered to be unreliable. The cut-off criteria of these three quality parameters depend on the requirement of the study.
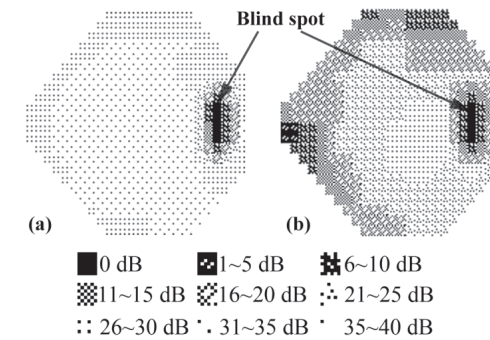
Figure 1-11. Gray scale sensitivity of a healthy (a) and glaucomatous (b) visual field measured by HFA with 24-2 pattern. Darker grey scale represents lower sensitivity.

SAP results are normally represented as a grey scale image (Figure 1-11) and, in the HFA, results are further presented via analysis software (Statpac) provided by the instrument. This software has been developed and improved over many years (Heijl, 1991, Bengtsson and Heijl, 1999b). It computes maps of the total deviation, pattern deviation and global indices such as mean deviation. The total deviation map presents decibel deviation from age-corrected normal threshold sensitivity (Heijl et al., 1987). The pattern deviation is derived in the same manner but is adjusted to take account of overall elevation or depression of the field, such as that commonly caused by cataract. Each deviation map is evaluated by a probability map that displays the significance of numerical deviation relative to the range of values found in a normative database of age related healthy people. The probability significance is displayed at various levels: <5%, <2%, <1%, <0.5%. So called global indices, expressed as single statistics, such as mean deviation (MD), pattern standard deviation (PSD), and more recently the Visual field Index (VFI) summarise departure of the measured visual field from a reference or normative database. The Glaucoma Hemifield Test (GHT) is another useful summary measure of a single visual field for the presence of a glaucomatous defect (Åsman and Heijl 1992). It was devised to detect field loss that is asymmetric about the horizontal meridian. Analysis is performed in five corresponding pairs of sectors that are based on the normal anatomy of the retinal nerve fibre layer. Fields are essentially classified as outside or within normal limits, borderline, or as having a general reduction in retinal sensitivity. All these measures provide statistical evidence about a visual field defect and are widely used in the clinical analysis of glaucoma. The strengths and weaknesses of these approaches have been debated over the years (Shaarawy et al., 2009), but it is fair to conclude that none of the methods represent perfect clinical analysis tools.

1.3 Difficulties in understanding measurements in glaucoma

The measurements of retinal structure and visual function introduced briefly above are used widely in the detection and follow up of glaucoma. However, these measurements are all surrogates of the target variable of real interest: retinal ganglion cells and their axons. These techniques are affected by multiple factors and thus produce measurement with noise and variability that introduce uncertainty into the decision making procedure in glaucoma management. Moreover, because these techniques measure the different 'perspectives' of the same process, it is expected that these measurements are correlated with each other. Nevertheless, these measurements are currently taken independently and are not sufficiently combined for better diagnosis and follow-up of glaucoma, beyond the way in which an experienced clinician might attempt to view the measurements in tandem.

1.3.1 Measurement variability

Variability is the tendency of the measurement process to produce different measurements on the same test item. It affects the precision of the instrument: even very precise instruments exhibit changes caused by random errors.

Despite advances in SAP the subjective nature of the test remains – after all, it is completely reliant on a subject reliably answering questions and pressing a button (Shaarawy et al., 2009). The variability in response or measurement noise is overwhelmingly multifactorial (Henson et al., 2000, Artes et al., 2002). These factors include: 1) *learning effects*, whereby a patient 'learns' to perform better over time after the first few perimetry tests, causing improved sensitivity after the first few tests (Wood et al., 1987); 2) *fatigue effects*, whereby patients tire and lose concentration during the test, and thus fail to respond to stimuli (Wall et al., 2001, Hudson et al., 1994); 3) *fixation loss*, whereby patients may lose their fixation due to eye or head movement during a perimetry test, so the visual stimuli are presented onto the wrong part of the visual field (Henson et al., 1995, Henson et al., 1996); 4) *variable reaction*, whereby patients may be 'trigger happy' to stimuli or be reluctant to respond to stimuli, causing overestimation and underestimation of the sensitivity thresholds; 5) *ocular media opacities: the* presence of cataract causes a general diffuse loss of sensitivity in visual field (Chen and Budenz, 1998, Smith et al., 1997). Some causes of variability are more controllable and more quantifiable than others, but not all variability can be explained by these factors (Blumenthal et al., 2003). Extensive discussion about variability in perimetry is given elsewhere (Henson, 2000, Lachenmayr and Vivell, 1993, Spry et al., 2002).
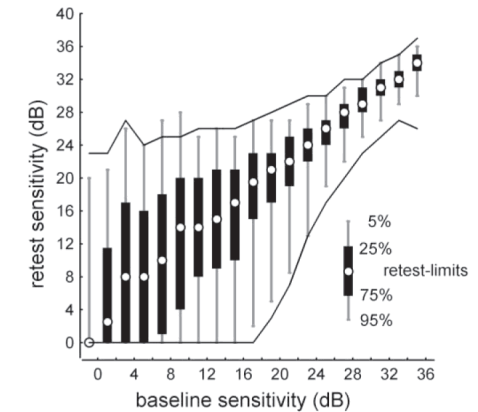
Figure 1-12. Retest limits of SAP tests using SITA standard algorithm (Artes et al., 2002). Thin vertical lines indicate the 90% retest intervals (5th and 95th percentiles of retest values); vertical bars indicate the interquartile ranges. The 90% retest limits of the full threshold strategy (solid lines) are shown for comparison.

Measurement variability for SAP can be estimated by so-called test-retest studies, where repeated tests are taken within a short period of time, under the assumption that any change that occurs between them cannot be due to the disease process itself. Figure 1-12 reproduces the test-retest limits of SAP tests using the SITA standard algorithm (Artes et al., 2002). In this figure the retest sensitivity is plotted against the baseline sensitivity for every test location and for every patient in the sample studied. This type of plot is a useful graphical descriptor of the variability: if there were no measurement variability, then the graph would simply give a diagonal line of unity with no spread on either side of the line. Note, the alarmingly large variability between the 5th and 95th percentiles of the retest limits. For instance, in Figure 1-12, if the sensitivity is estimated to be 15dB in a baseline test, a repeated test may estimate the sensitivity to be almost anywhere across the whole range of the measurement. Moreover, visual field test variability increases when the sensitivity decreases. In Figure 1-12, distance between 5% and 95% retest limits is narrower when the baseline sensitivity increases. This is also reported in previous studies (Heijl et al., 1989, Henson et al., 2000, Artes et al., 2002, Heijl et al., 1991). For example, Henson et al (Henson et al., 2000) indicated that this relationship between sensitivity and variability can be described by a log-linear model: $\ln(variability) = -0.081 \times sensitivity + 3.27$, where *sensitivity* is between 10dB and 39dB (Henson et al., 2000).

Variability of retinal structure measurement depends on the imaging technique used. For example, variability of measurements from the OCT may be caused by optical media opacity (Velthoven et al., 2006), fixation stability (Campbell et al., 2007), placement of scan circle (Gabriele et al., 2008), stages of disease (Budenz et al., 2005), operator skill (Villain and Greenfield, 2003), pupil dilation (Paunescu et al., 2004), magnification (axial length of eye and refractive error) (Wang et al., 2007), and post processing method (Patel et al., 2009, Ho et al., 2009). As with SAP, variability of retinal structure measurement can be evaluated by collecting multiple measurements in a short period of time. The results from such a study can be assessed by using, for example, the test-retest variability (two times standard deviation (SD)). Typical variability in RNFLT measured by StratusOCT using RNFLT (3.4) protocol has been evaluated (Budenz et al., 2005). Three repeated measurements were taken in this study. Test–retest variability in total RNFLT ranged from 3.5μm to 4.7μm in healthy eyes and 5.2μm to 6.6μm in glaucomatous eyes. Moreover, it was reported that measurements from RNFLT (3.4) has consistently less variability than those made with Fast RNFLT (3.4) (Budenz et al., 2005). This may be caused by the lower sample rate in Fast RNFLT (3.4).

In particular, scan circle displacement was suggested to be a source of variability for the measurement taken with circular scan protocols in TD-OCT and SD-OCT, and this was confirmed by other studies (Gabriele et al., 2008). This will be further discussed in chapter 2. Moreover, post processing methods also impact on the quality of measurement taken by OCT. For instance, the OCT can be quite misleading if the segmentation algorithms fail to quantify the tissue layers correctly (Patel et al., 2009). Even if SD-OCT demonstrates better technical specification (Kim et al., 2009, Forooghian et al., 2008), the technology may still be limited by the segmentation method employed to delineate the important retinal layers and clinically relevant features. Indeed, it has been reported that better segmentation methods are needed to improve the measurement from SD-OCT (Ho et al., 2009). This forms the main topic of chapter 3.

Variability of retinal structure and visual function measurements inevitably introduces uncertainty to the diagnosis and follow-up of the defect caused by glaucoma. Therefore, reduction of variability, or quantification of this variability, is a central task of the post-analysis of these measurements and is the motivation for much of the work in this thesis.

1.3.2 Structure-function relationship

It is assumed that visual function loss should follow from, and correlate to, structural loss caused by glaucoma (Garway-Heath et al., 2002, Bartz-Schmidt et al., 1999, Weber et al., 1990, Jonas and Grundler, 1997, Garway-Heath et al., 2000, Johnson et al., 2003, Caprioli and Miller,

1988). It would therefore be clinically useful to know the magnitude and location of structural loss that will result in visually important functional loss. Moreover, sources of variability in structural and functional measurements are different, so correlating and combining them should help to reduce the overall variability.

There has been no lack of interest in the relationship between structure and function in the glaucomatous process. This was, in part, precipitated by the initial work reported by Quigley et al (Quigley et al., 1982). The exhaustive scientific debate that has ensued on the subject will be discussed in chapter 4. In order to establish a relationship between functional measures and currently available *in-vivo* measures of retinal and ONH structure, it is essential to first establish a proper topographic relationship between them. That is to establish a relationship between the locations tested in clinical perimetry and the corresponding areas of RNFL and ONH, thus providing a plausible way of combining the measurements to provide better clinical information. This is the subject of much of the work in this thesis (chapters 4-6).
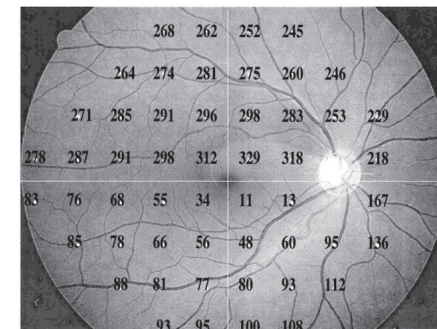
Figure 1-13. A 24-2 test pattern superimposed on a RNFL photograph. Each number indicates the angle (in degrees) at which the corresponding nerve fibre bundle enters the optic nerve head (ONH). The temporal (9-o'clock position, right eye) is designated 0°, and degrees are counted in a clockwise direction (Garway-Heath et al., 2000).

An important step in identifying the topographic relationship between glaucomatous RGC loss, damage to the RNFL and NRR and visual field damage was first reported by Garway-Heath et al (Garway-Heath et al., 2000). In this study, a grid of the 24-2 VF test pattern was superimposed on RNFL photographs (Figure 1-13). The relationship between a visual field test point and the ONH was determined by identifying points adjacent to the edge of a RNFL defect or prominent bundle and manually tracing the defect or bundle back to the ONH. A map was established to correlate the points in SAP test and the corresponding angles at which the nerve fibre bundles enter the ONH (Figure 1-13). The structure-function relationship from this study is

generally referred as an 'anatomical' map, because the relationship was derived using the anatomical defect observed in RNFL photographs instead of any surrogate imaging technique. This map will be used as a benchmark to evaluate the structure-function relationship model derived in chapter 4.

In Figure 1-14, the anatomical map divides the ONH into 6 sectors, each of which is related to a corresponding region in the visual field. This division is fundamental to many other recent studies about structure-function relationship that will be reviewed in detail in chapter 4.

(a)   (b)
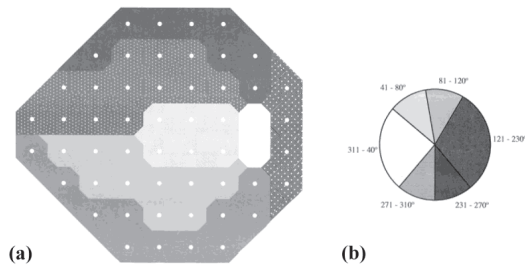
Figure 1-14. Division of (a) visual field and (b) optic nerve head (ONH). The correlated sectors in visual field and ONH are denoted in the same grey scale (Garway-Heath et al., 2000).

The anatomical map describes a topographic and qualitative relationship between retinal structure and visual function. However, this structure-function relationship is in a 'low resolution' format: each visual field location is only related to one discrete angle on the ONH. It is expected that, by using measurements with more information, a quantitative relationship can be derived in a 'higher resolution'. For example, with the 128-by-128-pixel RNFLT measurement acquired by GDxVCC, the relationship between each visual field point and individual RNFLT pixels can form a map at a 128-by-128 resolution for each visual field location. This forms another aim of the work reported in chapter 4.

However, quantifying the structure-function relationship using the surrogate measurements is not trivial. For instance, according to the anatomical map in Figure 1-13, the second point in the first row (3° on horizontal and 21° on vertical) of 24-2 visual field should be related to the nerve fibre bundle at 262°. To illustrate this, Figure 1-15 shows a plot the sensitivity at this visual field location against the corresponding RNFLT (measured by GDxVCC) at 262° from 464 healthy and glaucomatous eyes (this data is for illustration but will be described in more detail in chapter 4). The relationship between visual field sensitivity and RNFLT in this example is clearly ambiguous, with no clear relationship about how visual function changes with retinal structure.
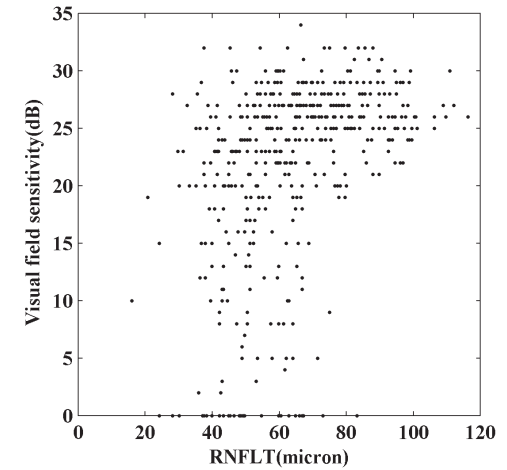
Figure 1-15. Plot of visual field sensitivity at the second point in the first row (3° on horizontal and 21° on vertical) of 24-2 test pattern against the corresponding GDxVCC RNFLT at 262°. The plot includes 464 healthy and glaucomatous eyes.

Confounders abound when attempting to evaluate the structure-function relationship. Variability in both structural and functional measurements may mask the true correlation. Moreover, the structure-function relationship may change at different stages of the disease (Gonzalez-Hernandez et al., 2009), which makes it unwise to plot the measurements with different defect severities on the same graph. Furthermore, artefacts in measurement may mislead the structure-function relationship. For instance, atypical scan in GDxVCC overestimates the RNFLT and moves the points towards the right in Figure 1-15. The floor effect of RNFLT measured by GDxVCC (Blumenthal et al., 2006) also limits the range of RNFLT measurements at the low end and skews the relationship in Figure 1-15. Other factors may include the possibly different structure-function relationship between the central and peripheral vision, the fact that the spacing of 24-2 visual field test does not accurately reflect the physiological distribution of RGCs, and the structural and functional measurements may interact as groups of points instead of individual points in a scatter plot. A quantitative structure-function relationship may be used to 'cross-validate' the structural and functional measurements. It may also help to quantify the agreement between the structural and functional measurements from the same patient and allow clinicians to consider the two types of measurements in tandem. These topics will be discussed in chapter 4 and 5.

1.4 Knowledge mining from clinical assessment of glaucoma

The variable and complex structure and function measurements pose great difficulty in understanding the 'intrinsic knowledge' in these measurements or how they relate to one another. Work reported in the thesis aims to extract the *knowledge* from the measurements taken from clinical assessment of glaucoma. This includes: 1) inference about the OCT scan circle displacement and investigation about its impact on RNFLT measurement; 2) robust and efficient quantification of retinal structure in large-scale 3D OCT image volumes; 3) quantification of structure-function relationship in glaucoma; 4) quantification of discordance between structural and functional measurements in glaucoma.

Despite the different solutions to these questions, they share the same underlying task: extract useful information, namely *knowledge*, from complex data. Several advances in this area have been brought together under the title of *data mining* (Han et al., 2005) and *knowledge discovery* (Cios et al., 2007) that developed out of the *data mining* domain. *Data mining* is a part of the overall process of *knowledge discovery* in databases (KDD) that has been defined in various ways (Fayyad et al., 1996a, Hand, 1998, Mackinnon and Glick, 1999). The central concept of KDD can be understood as a process of automatically searching large volumes of data for patterns that can be considered *knowledge* about the data. Data Mining can be considered as a central step of the overall process of the KDD process. The central techniques used in KDD and data mining range from simple pattern searching, discovery of association to advanced data visualization and machine learning methodologies.

Most data mining and KDD techniques are based on inductive learning (Mitchell, 1997) in which a model is constructed by generalizing from a sufficient number of training examples in order to 'fit' a generic model to the application data. These techniques have limited ability to incorporate the prior knowledge about the data and are thus not applicable to the highly specific prior information such as those reported in this thesis. Moreover, the aim of this thesis is not to apply existing data mining tools to extract information but to develop and integrate a wide range of data analysis methods that are able to derive directly or incrementally comprehensible and communicable scientific *knowledge* from large volumes of data using relevant prior knowledge. The *knowledge* obtained through the process may become additional data that can be used for further usage and discovery. Therefore, the studies to be presented in this thesis were characterised as *knowledge mining*.
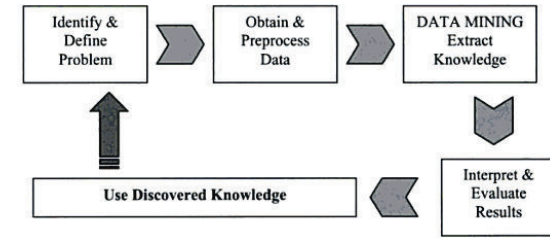
Figure 1-16. Illustration of the recurrent procedure of *knowledge mining* (Fayyad et al., 1996b). The components are iterated and new *knowledge* is discovered in each cycle of the process.

Instead of being defined as a specific type of technique, *knowledge mining* is better understood as an iterative procedure for extracting various types of *knowledge* such as data patterns, models and learning processes. This recurrent procedure is depicted in Figure 1-16 (Fayyad et al., 1996b). In each cycle, the application domain and the problem to be solved are formulated. Potential techniques are then investigated to form a representation of *knowledge* (e.g. patterns or models) that summarise the intrinsic information and structure underlying the data. The discovered *knowledge* is then evaluated and interpreted. The validated *knowledge* is then put into practical usage. The newly formed *knowledge*, once used in practice, may form a new application domain which starts the next cycle of *knowledge mining*.

So, this thesis includes five cycles of *knowledge mining* from the clinical assessment of glaucoma:

1) Scan circle displacement is not recorded in OCT circular scans. This information can be restored by a scan circle displacement inference algorithm proposed in chapter 2. The algorithm 'borrows' *knowledge* from retinal fundus image and aligns the vessels in OCT images to those in fundus image. The impact of scan circle displacement on RNFLT was also investigated using the algorithm.

2) Volumetric SD-OCT images pose real difficulties because they yield complex and large-scale data. In chapter 3, a newly developed technique called *FloatingCanvas*, provides a robust and efficient quantification of the *knowledge* about retinal structures in 3D SD-OCT image volumes. The measurement variability of extracted RNFLT was compared with that of TD-OCT.

3) The structure-function relationship is investigated in chapter 4 by using RNFLT and visual field measurements. A probabilistic model was developed to predict the visual field from the RNFLT measurements. The model was evaluated by prediction efficiency and the

derived spatial structure-function relationship compared with the anatomical map reviewed in this chapter.

4) Derived *knowledge* about the structure-function relationship was further used to quantify and visualize the discordance between the retinal structure and visual function in chapter 5. A structure-function discordance index and a pattern discordance map were proposed as clinical tools for this purpose.

5) The structure-function discordance index from the previous iteration of *knowledge mining* was then used to combine the structural and functional measurements in chapter 6. The combined measurements demonstrated significantly improved reproducibility, which was evaluated by examining the specificity in glaucoma change detection.

Chapter 2 : Aligning scan acquisition circles in optical coherence tomography images of the retinal nerve fibre layer

2.1 Summary

As introduced in chapter 1, optical coherence tomography (OCT) is widely used in the assessment of retinal nerve fibre layer thickness (RNFLT) in glaucoma (Huang et al., 1991, Velthoven et al., 2007). Images are typically acquired with a circular scan around the optic nerve head. Accurate registration of OCT scans is essential for measurement reproducibility and longitudinal examination. However, the *knowledge* about the scan circle location on the retina is 'missing' during the image acquisition. The work reported in this chapter developed and evaluated a novel algorithm to 'restore' the missing *knowledge* and to align the location of the OCT scan circles to the vessel features in the retina using probabilistic modelling that was optimised by an expectation-maximum algorithm. Evaluation of the method on 18 patients undergoing a large numbers of scans indicated improved data acquisition and better reproducibility of measured RNFLT when scanning circles were closely matched. The proposed method enables clinicians to consider the RNFLT measurement and its scan circle location on the retina in tandem, reducing RNFLT measurement variability and assisting detection of real change of RNFLT in the longitudinal assessment of glaucoma.

The work reported in this chapter has formed a paper in *IEEE Transactions on Medical Imaging* ((Zhu et al., 2011b); See List of supporting publications). The joint authors of this work are Haogang Zhu [HZ], David Crabb [DC], Patricio Schlottmann [PS], Gadi Wollstein [GW] and David Garway-Heath [DGH]. The data was collected by PS and GW, and the work was supervised by DC and DGH; everything else described in this chapter was completed by HZ. The results in this chapter have also been presented in part as a poster presentation at the Association for Research in Vision and Ophthalmology Meeting, Fort Lauderdale, Florida, USA in May 2007 (See List of supporting publications).

2.2. Introduction

Estimates of retinal ganglion cell axon (nerve fibre) loss can be made by the surrogate measurement of RNFLT using OCT (Huang et al., 1991, Velthoven et al., 2007). Similar to ultrasound technique, OCT detects the backscattered light from the retina and produces high resolution, cross-sectional images. This technique has formed time-domain OCT (TD-OCT) systems such as the StratusOCT (Carl Zeiss Meditec, Inc., Dublin, CA) that has been successfully used in the clinic as a clinical standard for measuring the RNFL thickness (RNFLT) in glaucoma in recent years (Schuman et al., 2005, Hee et al., 1995, Carpineto et al., 2003,

Bourne et al., 2005, Schuman et al., 2007). In the StratusOCT system measurement of the retinal layers is acquired by an axial-scan in depth (A-scan) and a cross-sectional scan (B-scan). The A-scans are sampled under a scan circle (typical diameter of 3.4 mm) which is manually centred on the optic nerve head (ONH; Figure 2-1I) as guided by a 'live' image of the fundus but the location of the scan circle is unknown during the image acquisition. The RNFLT is then analysed by segmentation algorithms provided by the software. One difficulty during image acquisition is the displacement of the circular scan due to the operator's subjective placement of the scan circle or the eye movement after the manual adjustment. This displacement means that the RNFLT is not necessarily sampled at the same location on the retina and this contributes to the variability and error in the measurement (Gabriele et al., 2008, Vizzeri et al., 2008, Vizzeri et al., 2009a), restricting the use of the technique especially in determining the deterioration of the RNFLT in the longitudinal assessment or follow-up of glaucoma. Moreover, the problem of RNFLT reproducibility due to image acquisition difficulties was also recently identified as a limiting factor for this technology in the diagnosis or management of multiple sclerosis (Bock et al., 2010). Therefore, a method for identifying and aligning the location of the scan circle would be clinically useful as it may offer better tracking of the same area of RNFLT over time.

Newer spectral-domain OCT (SD-OCT) (Fercher et al., 1995, Wojtkowski et al., 2002) operates with faster scans (Nassif et al., 2004) giving improved signal-to-noise ratio in the measurements (Leitgeb et al., 2003, Boer et al., 2003) compared to TD-OCT. Although some commercially available SD-OCT (e.g. Cirrus, Carl Zeiss Meditec, CA, USA) scan protocols extract the circular scan from 3D volume scan, most other SD-OCT systems (e.g. RTVue-100, Optovue, Fremont, CA, USA) still provide circular scan as one of the scan protocols or include circular scans in more complex protocols (e.g. RTVue-100 NHM4 protocol consisting of 6 circle and 12 line scans) so they may still be affected by displacement between scans (Vizzeri et al., 2009b, Gabriele et al., 2008). Therefore, improvements in this image acquisition protocol, or at least knowing the area of RNFLT that has been acquired, will still be beneficial for SD-OCT devices with such scan acquisition protocols. Moreover, TD-OCT has been used to follow up the progression of glaucoma long before the emergence of the SD-OCT, and it is still widely used by glaucoma services where clinicians (or research study co-ordinators) are reluctant to abandon series of data collected with TD-OCT over time since this provides important information about the longitudinal characteristic of glaucoma. The method proposed in this study may facilitate migration from TD-OCT to SD-OCT, for instance, by aligning TD-OCT scan circles on the volumetric images acquired by SD-OCT, so that longitudinal series are not wasted.

The effect of scan circle location on the RNFLT measurement has been previously investigated by simulating different scan circle locations on a volumetric image around the ONH taken by

ultrahigh-resolution OCT (Gabriele et al., 2008). The circular scans were simulated by sampling the A-scans under a scan circle (3.4 mm diameter) shifted with known displacements horizontally (x-shift), vertically (y-shift) and diagonally from the centre of the ONH. This method allowed for systematic investigation of the variable circle placement effect. The results from this study clearly demonstrated that location of the OCT scan circle adds substantial variability to the RNFLT measurements. Since registration of OCT scans is imperative for measurement reproducibility and longitudinal examination it would be very useful to have a method that could estimate the location of the scan circle on the retina. Kim et al. (Kim et al., 2010) proposed a method to align the circular scan image to a volumetric image around the ONH acquired by SD-OCT by using simulated cross-sectional images under scan circles at various locations sampled from the volumetric image. The circular scan was then aligned to the most similar sampled SD-OCT scans where the similarity was assessed by cross correlation between retinal structures in the A-scans from two images. One limitation is that the technique uses retinal structures that typically change during the worsening of glaucoma, giving it limited appeal in following up RNFLT changes if the circular scan and volumetric scans are acquired in different periods of time. This approach might help bridge the gap in RNFLT measurements between the TD-OCT circular scan and SD-OCT volumetric scan, providing longitudinal comparability. However this approach is only useful when both TD-OCT and SD-OCT are available and it will not be helpful in a common situation where a clinic might be following patients with TD-OCT technology alone for years.

This the work reported in this chapter proposes a new OCT scan circle alignment algorithm using blood vessel features which are considered to be relatively stable landmarks when considering longitudinal images in glaucoma. The algorithm can align multiple OCT circular scans to a retinal fundus image that is generally available in the glaucoma clinics from various imaging techniques, such as scanning laser polarimetry (SLP), scanning laser ophthalmoscope or even a retinal fundus camera, etc. The algorithm has been developed to have general applicability to any type of fundus and OCT images but is demonstrated in this study on StratusOCT images using fundus images acquired with SLP (GDxVCC; Carl Zeiss Meditec, Inc., Dublin, CA, USA). It has been successfully used in a recent study that assessed the axonal birefringence of RNFL by aligning the OCT scan circle onto the SLP image (Sehi et al., 2010).

2.3. Subjects and datasets

Eighteen patients (mean age of 65 (range 50 to 82) years) with a clinical diagnosis of glaucomatous optic neuropathy (primary open angle or normal tension glaucoma) with reproducible visual field defects were recruited. The study was approved by an ethics committee and informed consent, according to the tenets of the Declaration of Helsinki, was obtained prior

to examination from each subject. In the study protocol, a chosen eye from each subject was imaged 23 times with the StratusOCT system using the Fast RNFL Thickness (3.4) protocol. This protocol acquired three consecutive single scans in one image acquisition giving 69 single scans for each eye. Fundus images were acquired with the GDxVCC which covers an area of 5.9mm×5.9mm around the ONH. Patient identifiers were removed from the data before being transferred to a secure database at City University London.

2.4. Methods

The proposed method aligns an OCT scan circle on the retinal fundus image by a registration technique using the blood vessel features available in both types of images. The vessels in OCT images typically appear as shaded bands along the retinal pigment epithelium (RPE; Figure 2-1II). The RPE is detected as the tissue layer with the strongest intensity peaks in the OCT image. The 'shaded band' feature of vessels is then detected as the local minimums on the averaged pixel intensities around the RPE (Figure 2-1II). The vessel features in the retinal fundus images differ with the imaging techniques used, and the vessel segmentation in retinal images have been extensively studied previously (Soares et al., 2006, Staal et al., 2004, Zana and Klein, 2001, Hoover et al., 2000, Xiaoyi and Mojon, 2003, Lam et al., 2010). In the implementation of vessel detection for SLP fundus image, a measure of 'vesselness' serves as a pre-processing step for segmentation of vessels in the retinal fundus image. A technique using the multi-scale second order local structure of an image (Hessian) (Frangi et al., 1998) is used for this purpose. The vessels are then analytically reconstructed using cubic splines (Figure 2-1I) (Ahlberg et al., 1967).

A scan circle around the ONH and the detected vessels in the OCT (white circles superimposed on the blue scan circle) and fundus image (red lines) are shown in Figure 2-1. That the scan circle is displaced is indicated by the poorly aligned vessels. The method proposed in this work infers the scan circle location by aligning the vessels from both images.

2.4.1. Problem formalisation

In an acquisition of an OCT image (Figure 2-1), the circular scan starts from the mid-temporal area at 180° (blue arrow on scan circle in Figure 2-1I and traverses in a clock-wise direction to superior (90°), nasal (0°), inferior (-90°) and finally back to the mid-temporal area. The circular scan is then 'straightened' to a 'line' in two dimensions (Figure 2-1II). Each column in the OCT image is therefore associated with an angular value on the scan circle. The locations of detected vessels in the OCT image are converted to angular values (e.g. Figure 2-1III) and are denoted as $\{\mathbf{X}_i\}_{i=1}^{N}$ where N is the number of OCT vessels. The x- and y-coordinates of each vessel in the retinal fundus image are expressed as two cubic splines (Ahlberg et al., 1967) respectively, each

of which is essentially a piecewise function defined over a parameter u. The cubic spline has H-1 segments divided by H knots $(u_1,..., u_H)$ on u and the one having an intersection with the scan circle is assumed to be the hth segment, so the coordinate (x_{jh}, y_{jh}) of this segment of the jth fundus image vessel are:

$$\begin{cases} x_{jh} = a_{jxh}(u-u_h)^3 + b_{jxh}(u-u_h)^2 + c_{jxh}(u-u_h) + d_{jxh} \\ y_{jh} = a_{jyh}(u-u_h)^3 + b_{jyh}(u-u_h)^2 + c_{jyh}(u-u_h) + d_{jyh} \end{cases} \tag{2.1}$$

Given a scan circle with centre coordinate (s, t), radius r=3.4mm and the rotation θ around the scan circle centre, the intersections between this vessel and the scan circle can be calculated by solving the following polynomial equation with respect to u:

$$\begin{aligned} & \left[a_{jxh}(u-u_h)^3 + b_{jxh}(u-u_h)^2 + c_{jxh}(u-u_h) + d_{jxh} - s \right]^2 \\ & + \left[a_{jyh}(u-u_h)^3 + b_{jyh}(u-u_h)^2 + c_{jyh}(u-u_h) + d_{jyh} - t \right]^2 = r^2 \end{aligned} \tag{2.2}$$

with the constraint $u_h \le u < u_{h+1}$ and inserting the root $\hat{u}$ back to Equation (2.1) to calculate the solutions of $\hat{x}_{jh}$ and $\hat{y}_{jh}$. The angular value of the intersection is then calculated as:

$$f_j(s, t, \theta) = \begin{cases} \tan^{-1}\left(\dfrac{\hat{y}_{jh} - t}{\hat{x}_{jh} - s} \right) + \theta & \hat{x}_{jh} \ge s \\[3mm] \tan^{-1}\left(\dfrac{\hat{y}_{jh} - t}{\hat{x}_{jh} - s} \right) + \theta + 180 \quad \hat{x}_{jh} < s, \hat{y}_{jh} \ge t \\[3mm] \tan^{-1}\left(\dfrac{\hat{y}_{jh} - t}{\hat{x}_{jh} - s} \right) + \theta - 180 \quad \hat{x}_{jh} < s, \hat{y}_{jh} < t \end{cases} \tag{2.3}$$

for $\{ f_j(s, t, \theta) \}_{j=1}^{M}$ where M is the number of vessels in the fundus image. $f_j(s, t, \theta)$ is also used to represent the corresponding fundus image vessel.

The scan circle alignment can be decomposed into two tasks: vessel matching and displacement parameter inference. The vessel matching links an OCT vessel $\mathbf{X}_i$ and a fundus image vessel $f_j(s, t, \theta)$ if they belong to a same vessel. The displacement parameter inference infers the parameters (s, t) and θ to minimise the distances between the matched vessels in two images. These two steps are both non-trivial and affect each other in a complex way. For instance, the OCT and fundus image vessels in Figure 2-1 cannot be matched without knowing the location parameters. An OCT vessel cannot be simply matched to the nearest fundus image vessel. In the

example in Figure 2-1, the nearest-matching criterion results in obvious erroneous vessel pairs (c-C, f-J, i-M and j-N in Figure 2-1). On the other hand, the displacement parameters cannot be inferred without knowing how the vessels in two images are matched. Because of the complicated interaction between vessel matching and parameter inference, treating them independently would result in suboptimal solutions.
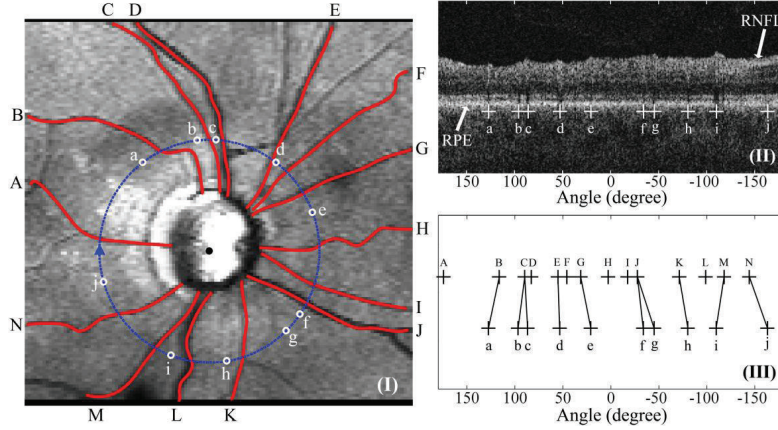
Figure 2-1. A retinal fundus image and OCT circular scan for the same eye. The detected vessels are modelled and delineated by cubic splines, shown here as red curves superimposed on the fundus image (I) and labelled with letters in uppercase (A to N). The image acquisition begins with a circle placed in an arbitrary position around the ONH, e.g. the blue line in (I). The scan starts from the mid-temporal area at 180° (blue arrow on scan circle) and traverses in a clock-wise direction to superior (90°), nasal (0°), inferior (-90°) and finally back to the mid-temporal area. The circular scan is 'straightened' to a line (in 2D) as shown in (II). The results of the OCT vessel detection technique are indicated in (II) as crosses and are superimposed as white circles in the fundus image (I). The OCT vessels are numbered by letters in lowercase (a to j). The angular values of the indicated position of the OCT vessels and the intersections between the scan circle and fundus image vessels are plotted in (III). The lines in (III) link each OCT vessel to the nearest fundus image vessel.

2.4.2. Probabilistic modelling

The complex relationship between the vessel matching and scan circle displacement parameter inference is modelled with a probabilistic unobserved variable model. The task of vessel matching can be divided into two processes. First, whether the OCT vessel $\mathbf{X}_i$ can be matched to a fundus image vessel is examined. If this can be done accurately then an inference about what M vessels in the fundus image it needs to be matched to. Two groups of unobserved

variables are introduced to model these two processes. The binary vector $\left\{ \mathbf{Y}_i = \left(\mathbf{Y}_{i0}, \; \mathbf{Y}_{i1} \right) \right\}_{i=1}^{N}$ is encoded with 1-out-of-2 notation in which only one of the two elements can be equal to 1 in $\mathbf{Y}_i$: $\mathbf{Y}_{i1} = 1$ if the ith OCT vessel $\mathbf{X}_i$ can be accurately aligned to a fundus image vessel, otherwise $\mathbf{Y}_{i0} = 1$. A prior probability over $\mathbf{Y}_i$ is introduced such that $P\left(\mathbf{Y}_{ik} = 1 \right) = \mathbf{u}_k$ for $k = 0$ or 1 so:

$$P\left(\mathbf{Y}_i \right) = \prod_{k=0}^{1} \left(\mathbf{u}_k \right)^{\mathbf{Y}_{ik}} \tag{2.4}$$

It is also required that probability values satisfy $\sum_{k=0}^{1} \mathbf{u}_k = 1$.

Another binary vector $\left\{ \mathbf{Z}_i = \left(\mathbf{Z}_{i1}, ..., \; \mathbf{Z}_{iM} \right) \right\}_{i=1}^{N}$ adopts 1-out-of-M notation so that only one of the M elements in $\mathbf{Z}_i$ can be equal to 1, and $\mathbf{Z}_{ij} = 1$ indicates that the ith OCT vessel is matched to the jth fundus image vessel. Similarly, a prior probability over $\mathbf{Z}_i$ is set as $P\left(\mathbf{Z}_{ij} = 1 \right) = \boldsymbol{\pi}_j$ for $k = 1$ to M so:

$$P\left(\mathbf{Z}_i \right) = \prod_{j=1}^{M} \left(\boldsymbol{\pi}_j \right)^{\mathbf{Z}_{ij}} \tag{2.5}$$

where $\left\{ \boldsymbol{\pi}_j \right\}_{j=1}^{M}$ are probability values satisfying $\sum_{j=1}^{M} \boldsymbol{\pi}_j = 1$.

The posterior probability of $\mathbf{X}_i$ given $\mathbf{Z}_{ij} = 1$ is defined as a mixture of two Gaussian distributions centred on the same mean of $f_j\left(s, \, t, \, \theta \right)$ but with different variance δ_k^2 that is decided by the value of $\mathbf{Y}_i$:

$$P(\mathbf{X}_i \mid \mathbf{Y}_i, \; \mathbf{Z}_{ij} = 1) = \prod_{k=0}^{1} \left(\mathcal{N}\left(\mathbf{X}_i \mid f_j\left(s, \, t, \, \theta \right), \; \delta_k^2 \right) \right)^{\mathbf{Y}_{ik}} \tag{2.6}$$

where the parameters $\left\{ \delta_k^2 \right\}_{k=0}^{1}$ are set to satisfy $\delta_1^2 << \delta_0^2$. This can be interpreted as if the OCT vessel $\mathbf{X}_i$ can be accurately aligned to the fundus image vessel $f_j\left(s, \, t, \, \theta \right)$ ($\mathbf{Y}_{i1} = 1$), then $\mathbf{X}_i$ needs to be close to $f_j\left(s, \, t, \, \theta \right)$ in order to 'score' a high probability, otherwise a small divergence (defined by small δ_1 such as 2°) from $f_j\left(s, \, t, \, \theta \right)$ would result in a probability near to zero. On the other hand, if the OCT vessel $\mathbf{X}_i$ cannot be accurately aligned to a fundus image

vessel ($\mathbf{Y}_{i0}=1$), $\mathbf{X}_i$ distributes more 'uniformly' (defined by large δ_0 such as 45°) with a small probability value, so the divergence from $f_j(s,\,t,\,\theta)$ has minimal effect on the probability.

From Equation (2.6), the posterior probability $P(\mathbf{X}_i\,|\,\mathbf{Y}_i,\,\mathbf{Z}_i)$ can be defined as:

$$P(\mathbf{X}_i\,|\,\mathbf{Y}_i,\,\mathbf{Z}_i) = \prod_{j=1}^{M}\left(\prod_{k=0}^{1}\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)^{\mathbf{Y}_{ik}}\right)^{\mathbf{Z}_{ij}} \qquad (2.7)$$

from which the joint probability $P(\mathbf{X}_i,\,\mathbf{Y}_i,\,\mathbf{Z}_i)$ can be calculated as the multiplication of Equation (2.4), (2.5) and (2.7):

$$P(\mathbf{X}_i,\,\mathbf{Y}_i,\,\mathbf{Z}_i) = \prod_{j=1}^{M}\left(\boldsymbol{\pi}_j\prod_{k=0}^{1}\left(\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)\right)^{\mathbf{Y}_{ik}}\right)^{\mathbf{Z}_{ij}} \qquad (2.8)$$

The joint probability $P(\mathbf{X}_i,\,\mathbf{Y}_i,\,\mathbf{Z}_i)$ defines a mixture of Gaussian mixture. This 'mixture of mixture' model structure was previously used for classification problems (Zio et al., 2007) and in other applications where the model was named after 'compound mixture model' (Qin and Leung, 2005).

Because $\{\mathbf{Y}_i\}_{i=1}^{N}$ and $\{\mathbf{Z}_i\}_{i=1}^{N}$ are all unobserved variables as opposed to directly observed variables, the likelihood can be calculated by marginalising the joint probability over these unobserved variables and assuming that $\{\mathbf{X}_i\}_{i=1}^{N}$ are independent and identically distributed:

$$\begin{aligned} P(\mathbf{X}) &= \prod_{i=1}^{N}\sum_{\mathbf{Y}_i,\,\mathbf{Z}_i}\prod_{j=1}^{M}\left(\boldsymbol{\pi}_j\prod_{k=0}^{1}\left(\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)\right)^{\mathbf{Y}_{ik}}\right)^{\mathbf{Z}_{ij}} \\ &= \prod_{i=1}^{N}\left[\sum_{j=1}^{M}\boldsymbol{\pi}_j\sum_{k=0}^{1}\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)\right] \end{aligned} \qquad (2.9)$$

2.4.3. Expectation-maximization algorithm

An *expectation-maximization* (EM) algorithm (Dempster et al., 1977) is used for finding maximum likelihood estimates of the parameters in Equation (2.9). In the *expectation* (E) step, the posterior probability of the unobserved variables $P(\mathbf{Y}_i,\,\mathbf{Z}_i\,|\,\mathbf{X}_i)$ is calculated as:

$$\begin{aligned} P(\mathbf{Y}_i,\,\mathbf{Z}_i\,|\,\mathbf{X}_i) &= \frac{P(\mathbf{X}_i,\,\mathbf{Y}_i,\,\mathbf{Z}_i)}{P(\mathbf{X}_i)} \\ &= \frac{\prod_{j=1}^{M}\left(\boldsymbol{\pi}_j\prod_{k=0}^{1}\left(\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)\right)^{\mathbf{Y}_{ik}}\right)^{\mathbf{Z}_{ij}}}{\sum_{j=1}^{M}\boldsymbol{\pi}_j\sum_{k=0}^{1}\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)} \end{aligned} \qquad (2.10)$$

by using Equation (2.8) and the *i*th component in Equation (2.9). The expectations of $\mathbf{Z}_{ij}$ and $\mathbf{Z}_{ij}\mathbf{Y}_{ik}$ w.r.t. the distribution $P(\mathbf{Y}_i,\,\mathbf{Z}_i\,|\,\mathbf{X}_i)$ are then calculated in the E-step and will be used in the following *maximization* (M) step, which computes parameters maximising the expected log likelihood found in the E step:

$$\boldsymbol{\alpha}_{ij} = \mathbb{E}_{P(\mathbf{Y},\,\mathbf{Z}|\mathbf{X})}\left(\mathbf{Z}_{ij}\right) = P(\mathbf{Z}_{ij}=1\,|\,\mathbf{X}_i) = \frac{\boldsymbol{\pi}_j\sum_{k=0}^{1}\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)}{\sum_{j=1}^{M}\boldsymbol{\pi}_j\sum_{k=0}^{1}\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)} \qquad (2.11)$$

and

$$\boldsymbol{\beta}_{ij}^k = \mathbb{E}_{P(\mathbf{Y},\,\mathbf{Z}|\mathbf{X})}\left(\mathbf{Z}_{ij}\mathbf{Y}_{ik}\right) = P(\mathbf{Y}_{ik}=1,\,\mathbf{Z}_{ij}=1\,|\,\mathbf{X}_i) = \frac{\boldsymbol{\pi}_j\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)}{\sum_{j=1}^{M}\boldsymbol{\pi}_j\sum_{k=0}^{1}\mathbf{u}_k\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)} \qquad (2.12)$$

In the M-step, the expectation of the complete likelihood $\log\left(P(\mathbf{X},\,\mathbf{Y},\,\mathbf{Z})\right)$ under the distribution of $P(\mathbf{Y},\,\mathbf{Z}\,|\,\mathbf{X})$ is calculated and minimised w.r.t. the parameters $(s,\,t)$, θ, $\{\mathbf{u}_k\}_{k=0}^{1}$ and $\{\boldsymbol{\pi}_j\}_{j=1}^{M}$:

$$\begin{aligned} &\mathbb{E}_{P(\mathbf{Y},\,\mathbf{Z}|\mathbf{X})}\left(\log\left(P(\mathbf{X},\,\mathbf{Y},\,\mathbf{Z})\right)\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M}\mathbb{E}\left(\mathbf{Z}_{ij}\right)\log(\boldsymbol{\pi}_j) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=0}^{1}\mathbb{E}\left(\mathbf{Z}_{ij}\mathbf{Y}_{ik}\right)\log(\mathbf{u}_k) \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=0}^{1}\mathbb{E}\left(\mathbf{Z}_{ij}\mathbf{Y}_{ik}\right)\log\left(\mathcal{N}\left(\mathbf{X}_i\,|\,f_j(s,\,t,\,\theta),\,\delta_k^2\right)\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M}\boldsymbol{\alpha}_{ij}\log(\boldsymbol{\pi}_j) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=0}^{1}\boldsymbol{\beta}_{ij}^k\log(\mathbf{u}_k) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=0}^{1}\frac{\boldsymbol{\beta}_{ij}^k\left(\mathbf{X}_i - f_j(s,\,t,\,\theta)\right)^2}{\delta_k^2} + const \end{aligned} \qquad (2.13)$$

where Equation (2.11) and (2.12) have been used. This target function is divided into three components each of which contains a group of the parameters. Note that the components of this

function that do not include any parameters were grouped into a constant term as they are irrelevant to the maximisation of the objective function.

In order to maximise Equation (2.13) w.r.t. parameters $\left\{\boldsymbol{\pi}_j\right\}_{j=1}^M$ under the constraint of $\sum_{j=1}^M \boldsymbol{\pi}_j = 1$, a *Lagrange* multiplier $\lambda_{\boldsymbol{\pi}}\left(\sum_{j=1}^M \boldsymbol{\pi}_j - 1\right)$ was added into Equation (2.13). The derivative of the objective function w.r.t. $\boldsymbol{\pi}_j$ is set to 0:

$$\frac{\partial \mathrm{E}\left(\log\left(P\left(\mathbf{X}, \mathbf{Y}, \mathbf{Z}\right)\right)\right)}{\partial \boldsymbol{\pi}_j} = \sum_{i=1}^N \frac{\boldsymbol{\alpha}_{ij}}{\boldsymbol{\pi}_j} + \lambda_{\boldsymbol{\pi}} = 0 \qquad (2.14)$$

After multiplying $\boldsymbol{\pi}_j$ on both sides of Equation (2.14) and summing over j from 1 to M then:

$$\lambda_{\boldsymbol{\pi}} = -\sum_{i=1}^N \sum_{j=1}^M \boldsymbol{\alpha}_{ij} = -N \qquad (2.15)$$

Substituting Equation (2.15) into Equation (2.14), then $\boldsymbol{\pi}_j$ is given as:

$$\boldsymbol{\pi}_j = \frac{\sum_{i=1}^N \boldsymbol{\alpha}_{ij}}{N} \qquad (2.16)$$

The $\left\{\mathbf{u}_k\right\}_{k=0}^1$ can be solved in a similar way as $\left\{\boldsymbol{\pi}_j\right\}_{j=1}^M$ by using the *Lagrange* multiplier $\lambda_{\mathbf{u}}\left(\sum_{k=0}^1 \mathbf{u}_j - 1\right)$:

$$\mathbf{u}_k = \frac{\sum_{i=1}^N \sum_{j=1}^M \boldsymbol{\beta}_{ij}^k}{N} \qquad (2.17)$$

The optimization of (2.13) w.r.t. parameters (s, t), and θ is more complex because setting the derivative of Equation (2.13) w.r.t. these parameters to zero does not give a closed solution for these parameters. Therefore, the iterative Quasi-Newton optimization algorithm (Polak, 1971) is used to find the maxima of the third component in Equation (2.13)

$$\mathrm{E} = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sum_{k=0}^1 \frac{\boldsymbol{\beta}_{ij}^k \left(\mathbf{X}_i - f_j\left(s, t, \theta\right)\right)^2}{\delta_k^2} \qquad (2.18)$$

w.r.t. (s, t), and θ in every M-step. The Quasi-Newton algorithm makes use of the gradient of the Equation (2.18) w.r.t. (s, t), and θ, which can be simply derived from Equation (2.1), (2.2) and (2.3) by using the chain rule of the derivative. Quasi-Newton algorithm stops when the following parameter convergence criteria are met: 1) the absolute difference between the values of (s, t) between two successive iterations is less than $10^{-5} \mu\mathrm{m}$; and 2) the absolute difference between the values of θ between two successive iterations is less than 10^{-5} degrees.

In all, the EM algorithm starts with initial values of the parameters (s, t), θ, $\left\{\mathbf{u}_k\right\}_{k=0}^1$ and $\left\{\boldsymbol{\pi}_j\right\}_{j=1}^M$ and iterates between the E-step and M-step. The whole process of the algorithm is summarized in Figure 2-2.
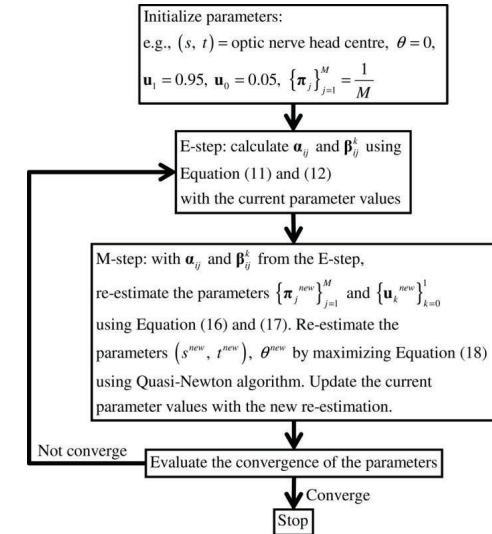
Figure 2-2. A summary of the EM algorithm.

In the E-step, $\boldsymbol{\alpha}_{ij}$ and $\boldsymbol{\beta}_{ij}^k$ in Equation (2.11) and (2.12) are calculated using initialized parameters in the first iteration or the parameters formed by the M-step in the previous iteration afterwards. In the M-step, $\boldsymbol{\alpha}_{ij}$ and $\boldsymbol{\beta}_{ij}^k$ calculated in the E-step are used and the expected complete log likelihood in (2.13) is maximized using Equation (2.16), (2.17) and Quasi-Newton optimization to form the new parameters that will be used in the next E-step. The iteration terminates if the scan circle falls outside of the borders of the fundus image or when the

following parameter convergence criteria are met: 1) the absolute difference between the values of (s, t) between two successive EM iterations is less than $10^{-5}\mu$m; 2) the absolute difference between the values of θ between two successive EM iterations is less than 10^{-5} degrees; 3) the absolute difference between the values of $\{\mathbf{u}_k\}_{k=0}^{1}$ and $\{\boldsymbol{\pi}_j\}_{j=1}^{M}$ between two successive EM iterations is less than 10^{-6}.

2.4.4. Evaluation and parameter initialisation

The algorithm was evaluated using the mean absolute angular difference (MAAD) between the matched vessels in the OCT and fundus images. The vessel matching is determined by the probability $P(\mathbf{Y}_i \mid \mathbf{X}_i)$ and $P(\mathbf{Z}_i \mid \mathbf{X}_i)$, which can be calculated from $P(\mathbf{Y}_i, \mathbf{Z}_i \mid \mathbf{X}_i)$ in Equation (2.10) by marginalization after the termination of the algorithm. The ith OCT vessel is aligned to the jth fundus image vessel if $P(\mathbf{Y}_{i1}=1 \mid \mathbf{X}_i) > P(\mathbf{Y}_{i0}=1 \mid \mathbf{X}_i)$ and $P(\mathbf{Z}_{ij}=1 \mid \mathbf{X}_i)$ is the largest among $P(\mathbf{Z}_{ij'} \mid \mathbf{X}_i)$ for $j'=1$ to M. In the unlikely case that multiple OCT vessels are so close that they are matched to the same fundus image vessel, only the OCT vessel closest to the fundus image vessel is matched. The vessel matching guarantees that the same number of vessels is matched in both OCT and fundus images so those matched vessels in two types of images form vessel pairs. The MAAD is then calculated as the mean of the absolute angular difference between these vessel pairs.

The criteria for a successful inference include: 1) $\mathbf{u}_1 \geq 0.7$; 2) the number of matched vessels is larger than 5; 3) the MAAD between the matched vessels is smaller than δ_1 which is set at $2°$ in the implementation with δ_0 set at $45°$. The choice of δ_1 and δ_0 will be explained in next section; 4) the inferred scan circle is within the border of fundus image. The first two criteria ensure that the inferred scan circle displacement is 'agreed' by adequate number of vessels. The third criterion guarantees that the distance between the matched vessels is sufficiently small.

Without losing generalisation, the parameters θ, $\{\mathbf{u}_k\}_{k=0}^{1}$ and $\{\boldsymbol{\pi}_j\}_{j=1}^{M}$ are initialised to be $\theta = 0°$, $\mathbf{u}_1 = 0.95$, $\mathbf{u}_0 = 0.05$ and $\boldsymbol{\pi}_j = \dfrac{1}{M}$. The choice of $\mathbf{u}_1 = 0.95$ results from the expectation that most OCT vessels can be aligned to the fundus image vessels. There are multiple initialisation options for the parameters (s, t) in order to cope with the potential large displacement of scan circles: they are initialised to be at nine locations shifted by 0 and $\pm 200\mu$m from the centre of the ONH on both the x- and y-axis. The algorithm starts with different parameter initialisation and if the inferred values of parameters are different with

different initialization, the displacement is chosen as the one with the lowest MAAD from a successful inference.

2.4.5. Choice of δ_1 and δ_0

The parameters δ_1 and δ_0 were set such that $\delta_1 << \delta_0$. Quantitatively, it was defined that the two intersections between the two Gaussian distributions defined by δ_1 and δ_0 in Equation (2.6) are $2.5\delta_1$ from the mean. Therefore, given the value of δ_1, δ_0 can be calculated thereafter. To find the optimal δ_1, various values of δ_1 are used and the mean MAAD and mean $\mathbf{u}_1$ are examined (Figure 2-3). To illustrate the effect of δ_1, the first two criteria of successful inference are not used because, as it will be shown below, fewer ($\mathbf{u}_1 < 0.7$) OCT vessels can be matched to the fundus image vessels with small δ_1 values.

δ_1 defines the necessary 'closeness' of the OCT vessel to the fundus image vessel in order to match the two vessels. As shown in Figure 2-3a, small δ_1 allows for a smaller difference between the OCT and fundus image vessels and thus gives better (lower) MAAD for the matched vessels. However, smaller δ_1 also excludes more OCT vessels so fewer OCT vessels can be matched to the fundus image vessels, which is quantified by lower value of $\mathbf{u}_1$ (Figure 2-3b). For instance, although $\delta_1 = 0.5°$ gives a low mean MAAD of 0.29°, but less than 70% ($\mathbf{u}_1 < 0.59$) of the OCT vessels can be matched to the fundus image vessels. On the other hand, large δ_1 allows for more OCT vessels to be matched to the fundus image vessels but the MAAD also increases at the same time.
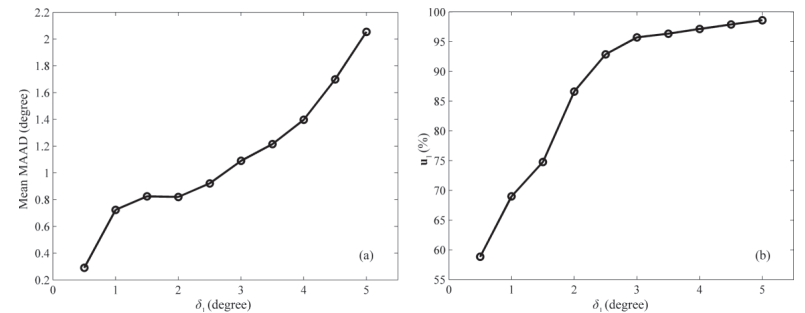
Figure 2-3. The mean MAAD and $\mathbf{u}_1$ under different values of δ_1.

Therefore, the choice of δ_1 reflects the trade-off between lower MAAD and having adequate number of matched vessels. In this implementation, δ_1 is chosen as the value giving the lowest mean MAAD with mean $\mathbf{u}_1 \geq 0.7$ as required by the criteria of successful inference. This value of δ_1 was found to be 2°, and δ_0 is calculated to be 45°.

2.4.6. Validation experiments

The OCT scan circle alignment algorithm was developed and implemented in Matlab (version 7.9.0 R2009b, The MathWorks, Inc., Natick, MA). The algorithm was then validated by investigating the impact of scan circle displacement on the RNFLT measurement and its repeatability.

2.5. Results

The algorithm was used to align all 69 OCT circular scans onto the corresponding fundus images for each eye. On average, the EM algorithm took 10.3 iterations before convergence. In each M-step, the average number of Quasi-Newton optimisation was 23.2. Computational time for aligning each OCT scan circle to the fundus image was 2.3s (SD 0.6s) on a typical desktop PC with 2.53GHz CPU and 2GB RAM.

2.5.1. Algorithm performance

An example of the results from the alignment algorithm for one of the eyes is shown in Figure 2-4. The initial location of the OCT circular scan and its vessels are described in Figure 2-1I. The EM algorithm took 11 iterations to estimate the location of the scan circle in this example. Relative rotation of this scan circle w.r.t. the fundus image was $3.5°$ and was corrected when plotting the OCT vessels in Figure 2-4I. The algorithm was successful in aligning the vessels in the images and this is quantified by the MAAD between the matched vessels having a small value of $0.4°$. In this example, most OCT vessels (from a to i) could be aligned to the fundus image vessels, as indicated by the large posterior probabilities $P(\mathbf{Y}_{i1}=1|\mathbf{X}_i)$ which are also given in Figure 2-4I. On the other hand, one OCT vessel (x) could not be aligned to any fundus image vessel because in this case $P(\mathbf{Y}_{i1}=1|\mathbf{X}_i)$=0. The OCT image in Figure 2-1II shows that the detection of this 'vessel' may be a false positive result by the OCT vessel detection algorithm because it isn't clear from looking at the image alone that there should be a vessel at that location. The vessel matching was decided by the largest posterior probabilities among $P(\mathbf{Z}_{ij'}|\mathbf{X}_i)$ for j'=1 to M which were denoted in Figure 2-4I.
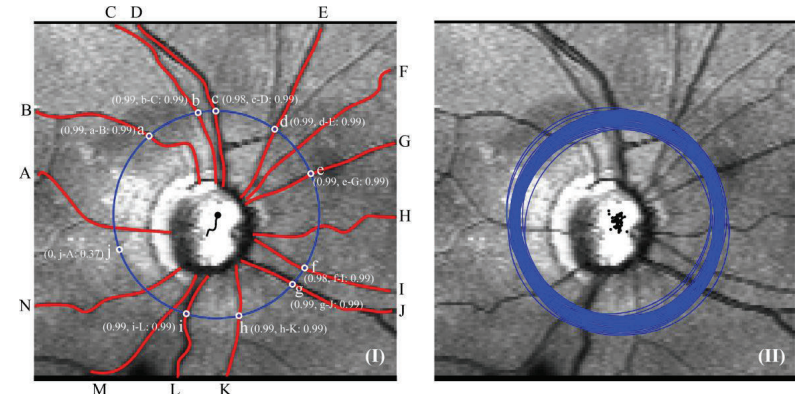
Figure 2-4. An example of OCT scan circle alignment algorithm. The initial location of the OCT scan circle and its vessels were described in Figure 2-1I and the initial rotation was 0°. The OCT scan circle and its vessels were superimposed on the fundus image (I) at the inferred location (black dot) and rotation (3.5°). The path of the scan circle centre at each step of the EM algorithm was plotted as a black curve on the fundus image. The posterior probability $P(\mathbf{Y}_{i1}=1|\mathbf{X}_i)$ and $P(\mathbf{Z}_{ij}|\mathbf{X}_i)$ for the matched fundus image vessel is denoted in the bracket beside each OCT vessel in the format of $(P(\mathbf{Y}_{i1}=1|\mathbf{X}_i)$, matched vessel pair (i-j): $P(\mathbf{Z}_{ij}|\mathbf{X}_i))$.

The algorithm produced successful inference of scan circle displacement for all OCT images. On average, the mean and SD of MAAD for all OCT circular scans (n=1242) in this sample of eyes were $0.82° \pm 0.34°$. The average number of detected vessels in these OCT images was 11.6. On average, 86% of the OCT vessels could be aligned to fundus images with posterior probabilities $P(\mathbf{Y}_{i1}=1|\mathbf{X}_i) > P(\mathbf{Y}_{i0}=1|\mathbf{X}_i)$. Therefore, the average number of matched vessels is 10 (11.6×86%), suggesting that, although 14% of the vessels were not matched, the algorithm could terminate with successful inference (criteria described in Section 2.4.4) 'agreed' by the majority of the vessels and was able to align the OCT vessels to the fundus image vessels with minimal angular difference.

Locations of the 69 repeated circular scans from the same example eye are shown in Figure 2-4(II). Although the operator aimed to scan the same circular area on the retina, the scan circles, as inferred by the algorithm, are displaced from each other and covered a wide annulus area around the ONH. The distance (relative shift in microns and as degrees of relative rotation) between the centre of each scan circle, as determined by the algorithm, was calculated for all possible pairs of scans (n=2346). In the example shown in Figure 2-4, the mean distance between scan circles was 143μm (SD of 130μm) and relative rotation was 1.9º (SD of 1.8º).

During the OCT image acquisition, 3 scans are taken consecutively within 1.92 second after the manual placement of scan circle so the locations of the scan 'triplet' can be expected to be affected less by the circle placement. To examine the assumption, the distances on the x-axis, y-axis, the overall distances in microns and the rotational distances in degree between all pairs of scan circles were calculated for each eye. These distances were compared with those calculated with scan circle pairs from three consecutive scans during the same image acquisition. The mean and SD of these distances were summarized in Table 2-1, showing that, on average, the distance between two circular scans tends to be smaller if they belong to the scan triplet from the same image acquisition.

	All scans	Three consecutive scans
Distance on x-axis	88±91μm	29±36μm
Distance on y-axis	117±103μm	34±43μm
Overall distance	153±122μm	49±77μm
Rotation difference	1.7°±1.8°	0.8°±1.6°

Table 2-1. Distances and rotation difference (mean±SD) between two circular scans.

2.5.2. Effect of scan circle displacement on RNFLT

The impact of scan circle displacement on RNFLT measurement was examined. Quadrant RNFLT measurements (temporal, superior, nasal and inferior) were plotted against x-axis and y-axis displacements of the centre of each scan circle. Linear regression then gave estimates of the average change of quadrant RNFLT caused by the displacement of scan circle. An example from one eye is shown in Figure 2-5.

The superior and nasal RNFLT are negatively correlated with the scan circle location on y- and x-axis respectively. Similarly the inferior and temporal RNFLT are positively correlated with the scan circle location on y- and x-axis respectively. On average, the inferior RNFLT increases by 3.9±0.4μm and the superior RNFLT decreases by 4.2±0.4μm when the y-coordinate of scan circle centre increases by 100μm; the temporal RNFLT increases by 3.5±0.4μm and the nasal RNFLT decreases by 4.2±0.5μm when the x-coordinate of scan circle centre increases by 100μm.
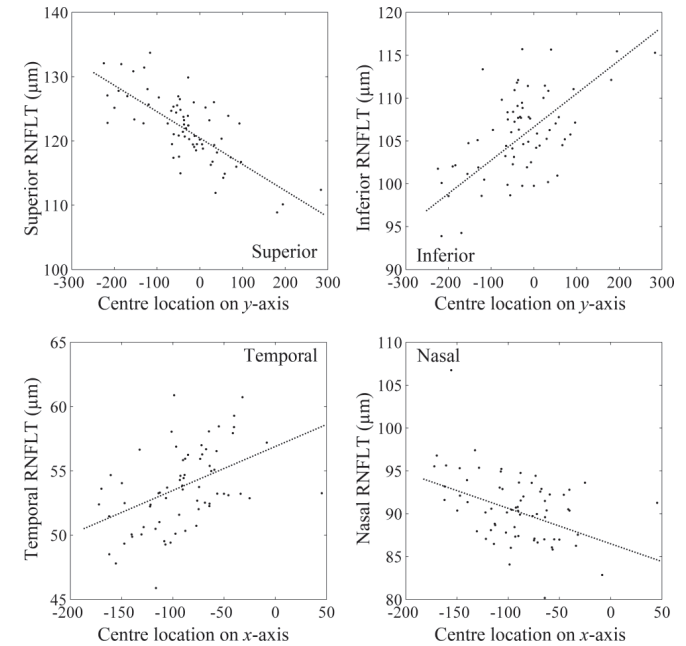
Figure 2-5. Plot of quadrant RNFLT measurements against the scan circle centre locations on x- and y-axis from one eye with fitted linear regression lines. The origin point (0μm) of the scan circle centre was arbitrarily chosen within the ONH. The slopes of the lines all differ from 0 (p<0.01).

The impact of scan circle displacement on RNFLT can also be observed by the change of RNFLT profile under scan circles at different locations. Figure 2-6 shows four scan circles relatively displaced towards superior, inferior, temporal and nasal directions. Note that the RNFLT profiles feature a 'double hump' shape where the RNFLT in superior and inferior areas is thicker than that of temporal and nasal regions. It is clear that moving the scan circle inferiorly increases the superior RNFLT and decreases the inferior RNFLT, and vice versa. Similarly, the scan circle on the nasal side has relatively thicker temporal RNFLT and thinner nasal RNFLT compared with the scan circle on the temporal side.
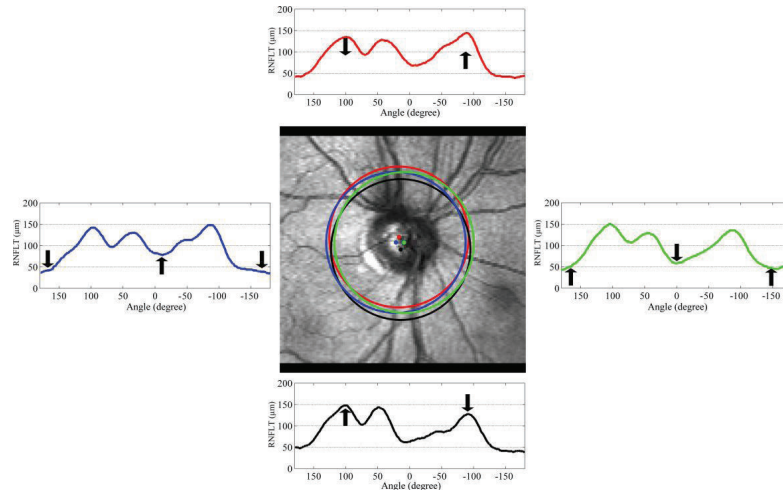
Figure 2-6: The RNFLT profile under scan circles at different locations on the retina. The scan circles at different locations were superimposed on the fundus image. These four scan circles were relatively superior (red), inferior (black), temporal (blue) and nasal (green) to each other. The tendency of change on RNFLT profile caused by the displacement of the scan circle location is indicated by the arrows in RNFLT profile plots.

2.5.3. RNFLT measurement variability

The RNFLT change caused by the scan circle displacement contributes to the variability of RNFLT measurement. The effect of scan circle displacement on mean and quadrant RNFLT measurement variability was investigated in this study. Variability was scored as two times the standard deviation of three repeated scans (Budenz et al., 2005), which were drawn from the exhaustive combination of all repeated scans of each eye. The average variability was calculated with all repeated scans and those scans with average distance among scan circle centres smaller than 50μm and larger than 300μm (Table 2-2). In short, the former represent a group of circular scans that the technique revealed to be closely matched, while the latter are scans that are more disparate.

RNFLT measurements (both mean and quadrant RNFLT) under the scan circles that are close to each other (average distance <50μm) demonstrates significantly lower (paired t-test; p<0.001) variability compared with those measured under scan circles that are far away from each other (average distance >300μm). This shows that the variability of RNFLT measurement is affected by the scan circle displacement and the scan circles that are close to each other provide RNFLT measurements with significantly better reproducibility.

	<50μm	>300μm	All scans
Mean RNFLT	4.3	7.4	6.2
Temporal RNFLT (μm)	6.8	10.7	8.9
Superior RNFLT (μm)	8.5	14.9	12.7
Nasal RNFLT (μm)	9.7	15.1	13.9
Inferior RNFLT (μm)	8.3	14.7	11.2

Table 2-2. Mean and quadrant RNFLT variability with all repeated scans and scans with different average distances among scan circle centres.

2.6. Discussion

Retinal vessels, compared with other RNFL structures, are relatively stable features for tracking a patient with glaucoma over time. This makes it possible to align multiple OCT circular scans, acquired in time, to a uniform coordinate formed by the vessel structures in the retinal fundus image. The two tasks in scan circle alignment, vessel matching and scan circle displacement inference, however, interact in a complicated way and have not been studied previously. The scan circle alignment algorithm proposed in this study integrated these two interactive steps into an EM framework: the iterative E- and M-steps in the algorithm incorporate the vessel matching, parameter inference and their interaction in a natural way. The algorithm guarantees to find a local maximum that gives an optimal alignment between two types of images.

Despite the superior specifications of the new SD-OCT, recent studies found that the diagnostic capability of TD-OCT is no worse than that of SD-OCT in clinical management of glaucoma (Leung et al., 2009, Cho et al., 2010, Sehi et al., 2009) and other retinal diseases (Forooghian et al., 2008). Particularly, the reproducibility for TD-OCT, for 'closely-matched' scans (average distance among scan circles <50μm in Table 2-2) identified by the algorithm, is close to reported reproducibility for SD-OCT (Kim et al., 2009, Mwanza et al., 2010). Therefore, many glaucoma services 'inheriting' TD-OCT from their retina specialist colleagues as they migrate to SD-OCT, may be confident that the TD-OCT provides similar monitoring capabilities for glaucoma as current SD-OCT devices.

The rate of RNFLT change caused by scan circle displacement demonstrated in Section 2.5.2 (3.5μm in temporal, 4.2μm in superior, 4.2μm in nasal and 3.9μm in inferior when scan circle displaces by 100μm on x- and y-axis) is significant when compared with the variability of RNFLT measurement in Table 2-2. Dividing the rate of RNFLT change by the RNFLT variability with all scan circles in Table 2-2 shows that, on average, 39%, 33%, 30%, 35% of the variability in the temporal, superior, nasal and inferior quadrants can be explained by the scan circle displacement of 100μm, which is a significant amount of displacement compared with the average displacement in Table 2-1 (88μm on x-axis, 117μm on y-axis and overall 153μm). The algorithm proposed in this study also helps to bridge TD-OCT to the other imaging techniques

such as SD-OCT and SLP. It has been successfully used in a recent study that assessed the axonal birefringence of RNFL by aligning the TD-OCT scan circle on the SLP image (Sehi et al., 2010).

The results from the validation experiment indicate that the variable location of the scan circle adversely affects the reproducibility of RNFLT measurements. RNFLT measurements and estimates of the corresponding displacement of the scan circles used together would be clinically useful when following a patient over time. The alignment algorithm, when applied retrospectively to data, will identify measurements that might be expected to have high variability. The technique, therefore, could provide a quality measure of scan acquisition, and this could even be achieved at the point of image acquisition. Moreover, the scan circle alignment algorithm can be used to identify those RNFLT measurements from areas that are close to each other. For multiple scans acquired at different periods of time, the RNFLT measurements that are identified as more reproducible may better reveal the real physiological change of the RNFLT in the longitudinal assessment of glaucoma. This will, therefore, have clinical impact on monitoring the progression of glaucoma over time. A recent study about the rate of RNFLT change caused by glaucoma reported decrease rates between $-1.2\mu m/year$ and $-15.4\mu m/year$ of mean RNFLT over a 5-year period (Leung et al., 2010). We have shown that the scan circle alignment yields about a 30% reduction in the variability associated with average RNFLT measurements (Table 2-2: 6.2μm to 4.3μm). This better reproducibility (lower variability) of the RNFLT measurement means that fewer scans would be needed to detect the 'progression signal', potentially saving patient visits and resources. Moreover, the better reproducibility allows for earlier identification of progression and more accurate determination of a progression rate in a shorter period of time, which leads to more appropriate treatment, targeting more aggressive treatment of fast progressors and not over-treating those falsely believed to progressing rapidly because of noisy data.

The algorithm proposed in this study also helps to bridge OCT to the other imaging techniques such as SLP so the power of these techniques can be improved by their combination. It was shown, in a recent study (Sehi et al., 2010), that the reproducibility of the calculated RNFL birefringence was improved when the OCT scan circle is aligned to the SLP image using the alignment algorithm.

The vessel matching in the scan circle alignment algorithm is 'encoded' by two latent variables $\mathbf{Y}_i$ and $\mathbf{Z}_i$. The mixture of Gaussian distributions conditioned on $\mathbf{Y}_i$ in Equation (2.6) indicates whether the OCT vessel $\mathbf{X}_i$ can be aligned to a fundus image vessel and plays a key role in the algorithm. Figure 2-7 illustrates the mixture of Gaussian distributions with the same

mean centred on a fundus image vessel but with different standard deviations $\delta_1 << \delta_0$. An OCT vessel follows a peaked distribution around the fundus image vessel and scores a high probability if these two vessels are close enough. The 'closeness' is defined by two interceptions of two Gaussian distributions that are $k\delta_1$ ($k>1$) away from the mean. On the other hand, if the distance between the two vessels are not close enough (beyond the two interceptions), the probability of the Gaussian defined by δ_1 drops under the Gaussian defined by δ_0. In this case, the OCT vessel is 'forced' to follow the more 'uniform' distribution in order to score a relatively higher probability.
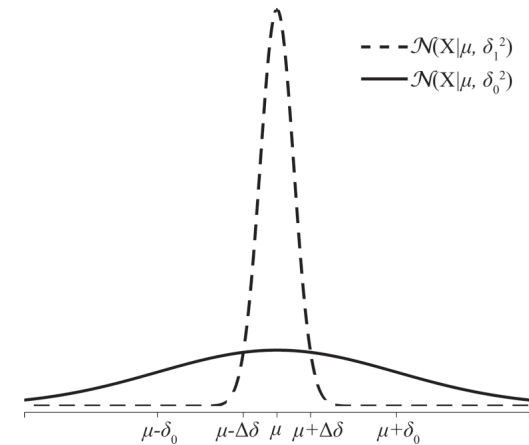
Figure 2-7. Illustration of the mixture of Gaussian distributions in Equation (2.6) for vessel matching. The two Gaussian distributions centre on the same mean $\mu = f_j(s, t, \theta)$ but have different standard deviations satisfying $\delta_1 << \delta_0$. The two Gaussian distributions intercept at two points that are $k\delta_1$ ($k>1$) from μ.

OCT vessels that cannot be aligned to any fundus image vessel (named here as 'noisy' vessels such as OCT vessel 'j' in Figure 2-4(I)) considerably mislead the parameter inference because their large distances from fundus image vessels dramatically affect the maximisation of the third term in Equation (2.13). The usage of the Gaussian mixture model helps to isolate the effect of these 'noisy' vessels. As an illustration, these 'noisy' vessels are all forced to follow a more uniform distribution (Figure 2-7) and thus have low probability values (defined in Equation (2.6)) near to zero. This, in turn, results in small $\boldsymbol{\alpha}_{ij}$ and $\boldsymbol{\beta}_{ij}^k$ for the corresponding 'noisy' vessels in the E-step. These near-to-zero values, once substituted into the objective function (2.13) in the M-step, have minimal effect on the objective function as well as its derivatives

with respect to the scan circle displacement parameters. This process ensures that the 'noisy' vessels do not interfere with the parameter inference.

Instead of using fixed values for the standard deviation δ_1 and δ_0, the model was adjusted to infer (data not shown) δ_1 and δ_0 from the data in the EM algorithm. However, the inference algorithm tended to increase δ_1 and decrease δ_0 so that more OCT vessels are matched to the fundus image vessel even with a large angular difference. This approach increased the likelihood in Equation (2.9) because more OCT vessels follow the 'peaked' Gaussian distribution in the Gaussian mixture even if they are not well aligned, but the accuracy of the alignment is worse (larger MAAD) at the same time due to the larger δ_1 Figure 2-3. Therefore, δ_1 and δ_0 are fixed as described in Section 2.4.5. The chosen standard deviation of $\delta_1 = 2°$ is small enough to meet the requirement of alignment accuracy and can incorporate the possible variance of vessel locations caused by factors such as potential eye movement during the image acquisition and possible physiological vessel shift over a long period of time. If the interceptions between the two Gaussian distributions are defined to be at $\Delta\delta = 2.5\delta_1$ away from the mean, the δ_0 is calculated to be $45°$.

Last but not least, as illustrated in Figure 2-4, if a large sample of repeated circular scans were acquired, then they might cover an annulus area around the ONH potentially allowing for a three-dimensional RNFLT profiles to be reconstructed. We have previously shown that this might be a way of bridging measurements acquired with StratusOCT and those volume measurements from more recently established SD-OCT systems (Zhu et al., 2007).

Chapter 3 : *FloatingCanvas*: quantification of 3D retinal structures from spectral-domain optical coherence tomography

3.1 Summary

Optical Coherence Tomography (OCT) has been widely used as a tool for evaluating the structure of the retina in cross-section (Huang et al., 1991, Velthoven et al., 2007). Time-domain OCT has been used in glaucoma diagnosis and follow-up by determining retinal nerve fibre layer thickness (RNFLT) since it was adopted in the clinical setting (Schuman et al., 2005, Hee et al., 1995, Carpineto et al., 2003, Bourne et al., 2005, Guedes et al., 2003). Because of its limited speed, time-domain OCT only provides RNFLT measurements in a line scan, generally a peripapillary circle but does not provide a three-dimensional (3D) RNFLT map.

The newly developed and commercialized spectral-domain OCT (SD-OCT) (Fercher et al., 1995, Wojtkowski et al., 2002) provides much faster scans (Nassif et al., 2004) with improved signal-to-noise ratio (Leitgeb et al., 2003, Boer et al., 2003) compared with time-domain OCT, for example, StratusOCT (Carl Zeiss Meditec, CA, USA). With this benefit, this technique represents a powerful and 'real-time' tool that potentially can be used in the clinic to assist the diagnosis and management of glaucoma. The extremely high image-acquisition speed allows a 3D image to be yielded. Each image is obtained by an in-depth axial scan (A-scan), a cross-sectional 2D scan (B-scan) consisting of a series of consecutive A-scans and the 3D image volume formed by consecutive B-scans. However, despite this high performance, such an imaging technique can only be useful clinically if there is a quantitative method to provide numerical information to the clinician. Moreover, the greater acquisition speed of SD-OCT means that a much greater amount of data is generated, which undoubtedly poses a technical challenge to the computer-assisted analysis.

This chapter presents a fully automated segmentation algorithm, *FloatingCanvas*, which performs a volumetric segmentation of retinal tissue layers in three-dimensional image volume acquired around the optic nerve head without requiring any pre-processing. The reconstructed layers are analysed to extract features such as blood vessels and retinal nerve fibre layer thickness. Findings from images obtained with the RTVue-100 SD-OCT (Optovue, Fremont, CA, USA) indicate that FloatingCanvas is computationally efficient and is robust to the noise and low contrast in the images. The FloatingCanvas segmentation demonstrated good agreement with the human manual grading. The retinal nerve fibre layer thickness maps obtained with this method are clinically realistic and highly reproducible compared with time-domain StratusOCT.

The work reported in this chapter has formed a paper published by *Optics Express* ((Zhu et al., 2010a); See List of supporting publications). The joint authors of this work are Haogang Zhu [HZ], David Crabb [DC], Patricio Schlottmann [PS], Tuan Ho [TH] and David Garway-Heath [DGH]. The data was collected by PS. Some of the manuscript was edited by TH, and the work was supervised by DC and DGH. Everything else described in this chapter was completed by HZ. The results in this chapter have also been presented in part as read paper at the Association for Research in Vision and Ophthalmology Meeting, Fort Lauderdale, Florida, USA in May 2008 (See List of supporting publications).

3.2. Purpose and previous studies

Algorithms for the segmentation of images have been extensively studied since the early days of computer vision and image processing. In the studies of OCT images, some segmentation algorithms were implemented by thresholding the OCT A-scan profile, which consists of a series of 'peaks' and 'valleys' that represent various high and low tissue reflectivity (Ishikawa et al., 2005, Fernández et al., 2005, Ahlers et al., 2008). A tissue layer boundary was found when the pixel intensity reached a target threshold that was sometimes adaptively chosen. These methods are computationally efficient, but are susceptible to intensity inconstancy within the individual layers. This problem was partly resolved by an efficient maximum intensity search based approach proposed by Fabritius et al (Fabritius et al., 2009). In other studies, a gradient was used to form an automatic active contour to minimize the overall energy, which was defined by gradient, boundary smoothness and edge density (Mujat et al., 2005). This method is less affected by the intensity variations, but is sensitive to morphological features such as blood vessels. Koozekanani et al (Koozekanani et al., 2001) utilized a Markov model to select and organize the edges to form a coherent boundary structure. A minimum-cost closed set approach was developed by Haeker et al (Haeker et al., 2007) and Niemeijer et al (Niemeijer et al., 2008) to identify retinal layers based on a linear combination of domain-specific cost functions. Mishra et al (Mishra et al., 2009) used the image gradient to derive an external force through an adaptive kernel function and used dynamic programming to identify the continuous retinal layers within the OCT images. Garvin et al (Garvin et al., 2008) attempted to find a closed set in a geometric 3-D graph that minimizes the associated costs and constraints by using an optimal graph search method (graph-cut). This method was then extended (Garvin et al., 2009) such that the constraints and cost functions were learned from a training set, and by using a multiscale 3-D graph search (Quellec et al., 2010, Lee et al., 2010): the retinal surfaces were detected in a subvolume constrained by the retinal surface segmented in a down-sampled image volume. Chiu et al (Chiu et al., 2010) also extended the graph-cut segmentation using dynamic programming and this helps to improve the computational efficiency when it was applied on individual B-scans. Moreover, segmentation using intrinsic tissue property such as

depolarization, or polarization scrambling, on backscattered light has also been investigated using other types of OCT instruments (Götzinger et al., 2008).

The majority of these studies (Fernández et al., 2005, Ishikawa et al., 2005, Ahlers et al., 2008, Mujat et al., 2005, Fabritius et al., 2009, Garvin et al., 2009, Quellec et al., 2010, Chiu et al., 2010) filtered the images to remove distractive features such as the speckle noise. These filtering methods were controlled by subjectively selected parameters and had difficulties in 'balancing' the deduction of high speckle noise and preservation of structural edges, especially in images with low contrast. Haeker et al (Haeker et al., 2007) and Garvin et al (Garvin et al., 2008), on the other hand, proposed image averaging to create composite images from repeat scans. The composite images had higher signal-to-noise ratio, but multiple scans (six repeat scans in this case) were needed, which may exacerbate the detrimental effects of eye movement between the scans.

A few studies performed real volumetric segmentation in 3D space (Haeker et al., 2007, Niemeijer et al., 2008, Garvin et al., 2008, Garvin et al., 2009, Quellec et al., 2010, Lee et al., 2010), but most other segmentation was done on individual B-scans and with the 3D layer topography created by filtering across the B-scans. The segmentation of individual B-scan is computationally efficient, but the filtering may reduce sensitivity to detect a structural abnormality and the change of an abnormality over time. A key benefit of combining the information from neighbouring B-scans is that it can reduce measurement variability, especially in volumetric scans with noise or low contrast in some B-scans. Computationally speaking, the core technique behind these real 3D segmentation methods is the 3D graph search algorithm (Li et al., 2006) that has an efficient polynomial time complexity.

The aim of this study was to develop a new segmentation algorithm, *FloatingCanvas*, that has a balance between the robustness and efficiency. *FloatingCanvas* was implemented to quantify the retinal structures in 3D image volumes around the optic nerve head (ONH) obtained with SD-OCT. It was used to process the whole image volume simultaneously and to reconstruct analytical surfaces for tissue layers or their boundaries. This method was designed to be robust to noise and artefact in image volumes and thus required no pre-processing such as filtering or image averaging. It made use of the first and higher order gradient as the natural boundary between tissue layers. In this case, the algorithm searches for the retinal pigment epithelium (RPE) and retinal nerve fibre layer (RNFL) boundaries which consequently form the RNFLT measurement. Although RNFL and RPE are in theory the two tissue layers with the strongest reflectivity in these OCT images, they and their boundaries become less identifiable in images with overall or local artefacts. *FloatingCanvas* was tested on images taken from both healthy and glaucomatous subjects, and was compared with manual segmentation by the human expert.

It was demonstrated that the algorithm was robust enough to detect the tissue layer boundaries in images with low contrast. The RNFLT maps obtained with this method were also compared with those derived from time-domain StratusOCT in healthy and glaucomatous subjects.

3.3. Methods

In this study, the SD-OCT images were acquired with the RTVue-100 using the 4mm × 4mm 3D volume scan protocol around the ONH with a depth of 2mm. This provides volumetric images with 101 B-scans comprised of 513 A-scans, each of 768 pixels in depth. Therefore, the distance between two B-scans is about 5 times that of two neighbouring A-scans.

The axis in the image is subsequently denoted as x for the direction of the B-scan, y in the direction across all B-scans and z for the direction of A-scan, and the location of a pixel in the image Im is described by a vector $\langle x, y, z \rangle^{\mathrm{T}}$ or a two-tuples $(\mathbf{x}, z)$, where $\mathbf{x}$ is a column vector $\langle x, y \rangle^{\mathrm{T}}$. The positive direction in an A-scan is defined to be from the top to the bottom of the image and will be described as 'downward' subsequently. Therefore, the value of the pixel in image Im is represented as $Im(x, y, z)$ or $Im(\mathbf{x}, z)$. The pixel coordinates were all converted to a scale in microns.

3.3.1. Analytical surface modelled by Gaussian Process

FloatingCanvas searches for a tissue layer, or its boundaries, in the image by deforming a 3D analytical surface that is efficiently modelled by a Gaussian Process (GP) (Williams, 1995). The analytical surface is spanned by a sample of 'skeleton' points $\{(\mathbf{x}_i, w_i)\}_{i=1}^{N}$, where $\mathbf{x}_i = \langle x_i, y_i \rangle^{\mathrm{T}}$ is a column vector containing coordinates on the x- and y-axis for the ith 'skeleton' point, and w_i is the coordinate on the z-axis. The skeleton points were evenly placed along the x- and y-axis to form a regular grid in the x-y space. The interval of the 'skeleton' point grid to model the anterior and posterior RNFL boundaries was chosen to be 100μm on both x- and y-axis, and the interval for RPE was 300μm given that RPE is expected to be smoother than RNFL boundaries around ONH. The GP model acts on these 'skeleton' points to determine a function $f(\mathbf{x})$ that provides a calculation of the surface coordinate z on z-axis for any vector coordinate $\mathbf{x}$ on x- and y-axis.

Similar to a Gaussian distribution, the GP is defined by a mean function and a covariance function: $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}_*))$, where the mean function $m(\mathbf{x}) = E(f(\mathbf{x}))$ is the expectation of $f(\mathbf{x})$, and the covariance function is defined as the expectation: $k(\mathbf{x}, \mathbf{x}_*) =$

$E((f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}_*) - m(\mathbf{x}_*)))$. In this case, the GP defines the joint probability between the skeleton points and the values $f(\mathbf{X}_*)$ at arbitrary locations $\mathbf{X}_*$ to be a Gaussian distribution:

$$\begin{bmatrix} \mathbf{w} \\ f(\mathbf{X}_*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \delta_n^2 \mathbf{I} & \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \\ \mathbf{K}(\mathbf{X}_*, \mathbf{X}) & \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix} \right) \qquad (3.1)$$

In Equation (3.1), $\mathbf{w}$ is a column vector containing all $\{w_i\}_{i=1}^{N}$ and $\mathbf{X}$ is a matrix with $\{\mathbf{x}_i\}_{i=1}^{N}$ in the columns, and a similar notation is used for $\mathbf{X}_*$. $\mathbf{K}(\mathbf{X}, \mathbf{X}_*)$ is a matrix with element $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_{*j})$, where $\mathbf{x}_i$ and $\mathbf{x}_{*j}$ are the ith and jth columns in $\mathbf{X}$ and $\mathbf{X}_*$ respectively; $k(\mathbf{x}_i, \mathbf{x}_{*j})$ is defined to be an un-normalized Gaussian kernel function between data points $\mathbf{x}_i$ and $\mathbf{x}_{*j}$: $k(\mathbf{x}_i, \mathbf{x}_{*j}) = e^{-\frac{1}{2}\left(\frac{((\mathbf{x}_i^1 - \mathbf{x}_{*j}^1)^2 + (\mathbf{x}_i^2 - \mathbf{x}_{*j}^2)^2)}{l^2} \right)}$, where l is the length-scale of the Gaussian and is the main parameter to control the smoothness of the analytical surface, $\mathbf{x}^1$ and $\mathbf{x}^2$ represent the coordinates of point $\mathbf{x}$ on the x- and y-axis, respectively. A similar notation applies to $\mathbf{K}(\mathbf{X}, \mathbf{X})$, $\mathbf{K}(\mathbf{X}_*, \mathbf{X})$ and $\mathbf{K}(\mathbf{X}_*, \mathbf{X}_*)$. δ_n^2 is the prior Gaussian noise variance of $\mathbf{w}$ and is fixed at 100μm in the algorithm. The parameter l was set to be 150μm for the surfaces modelling the anterior and posterior RNFL boundaries, and was 450μm for the surface to detect RPE.

Using the joint probability in Equation (3.1), the conditional distribution $p(f(\mathbf{X}_*) | \mathbf{X}, \mathbf{X}_*, \mathbf{w}) \sim \mathcal{N}(\bar{f}(\mathbf{X}_*), \mathrm{cov}(\bar{f}(\mathbf{X}_*))$ can be derived as (Williams, 1995):

$$\bar{f}(\mathbf{X}_*) = \mathbf{K}(\mathbf{X}_*, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \delta_n^2 \mathbf{I})^{-1} \mathbf{w} \qquad (3.2)$$

$$\mathrm{cov}(\bar{f}(\mathbf{X}_*)) = \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \delta_n^2 \mathbf{I})^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \qquad (3.3)$$

where $\mathbf{I}$ in Equation (3.2) and (3.3) is an identity matrix with 1 on the diagonal and 0 off the diagonal.

If $\mathbf{X}_*$ contains the coordinates of all points on the x- and y-axis, $\bar{f}(\mathbf{X}_*)$ returns the corresponding coordinates of the analytical surface on the z-axis.

In contrast to the conventional usage of GP where $\mathbf{w}$ are the fixed inputs for a regression problem, $\mathbf{w}$ are treated as parameters of the model in this algorithm. Therefore, the task of the algorithm is to search for the parameters $\mathbf{w}$ to form an analytical surface that is sufficiently close to a tissue layer or its boundaries (target surface).

3.3.2. Analytical surface deformation

In *FloatingCanvas*, an analytical surface is initialized to be at a regular location such as the top or bottom of the A-scan depending on the target surface. The analytical surface is deformed by updating the parameter $\mathbf{w}$ according to the forces applied on the points in the surface. Without specifying a certain target surface, the deformation process is described as a differential equation with regard to an artificial time t defined on the parameters $\mathbf{w}$:

$$\frac{d\bar{f}(\mathbf{X}_*)}{dt} = \bar{\Theta}\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right)F_{\mathbf{g}}\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right) + \Theta\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right)F_{Img}\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right) \tag{3.4}$$

where $F_{\mathbf{g}}$ is a function related to a constant 'gravity' force $\mathbf{g} = < 0,\ 0,\ g >^{\mathrm{T}}$ that drives the points on the surface towards one direction along the z-axis; F_{Img} is a force driving the surface to be attached on the target surface; Θ is a binary function consisting of one or several necessary conditions of $\mathbf{x}_{*_i}$ in $\mathbf{X}_*$ being a point on the target surface, and $\bar{\Theta}$ is the negation of Θ. The notation $F_{\mathbf{g}}\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right)$ assembles the values calculated from all $\mathbf{x}_{*_i}$ in $\mathbf{X}_*$ into a column vector, and the same notation also applies to F_{Img} and Θ. Moreover, for notational simplicity, the resulting column vectors calculated by these three functions are simplified by removing the common variables '$\left(\mathbf{X}_*, \bar{f}(\mathbf{X}_*)\right)$'.

Equation (3.4) describes a process, as indicated by the name of the algorithm, where the analytical surface acts as a 'canvas' floating in the 3D image and is driven by different forces: the analytical surface is initially 'moved' towards the target surface by a gravity force $F_{\mathbf{g}}$; when the analytical surface is close to the target surface (i.e. the corresponding Θ function outputs 1), the force F_{Img} takes over $F_{\mathbf{g}}$ and attaches the analytical surface on the target surface. The switch between $F_{\mathbf{g}}$ and F_{Img} is controlled by the function Θ.

To form an equation about the parameter $\mathbf{w}$, the left part of Equation (3.4) is expanded by inserting Equation (3.2) into Equation (3.4) and applying the chain rule of derivative:

$$\mathbf{K}(\mathbf{X}_*, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \delta_n^2\mathbf{I})^{-1}\frac{d\mathbf{w}}{dt} = \bar{\Theta}F_{\mathbf{g}} + \Theta F_{Img} \tag{3.5}$$

To form a concise notation, the matrix $\mathbf{K}(\mathbf{X}_*, \mathbf{X})(\mathbf{K}(\mathbf{X}, \mathbf{X}) + \delta_n^2\mathbf{I})^{-1}$ is substituted by Λ. Multiplying $\left(\Lambda^{\mathrm{T}}\Lambda\right)^{-1}\Lambda^{\mathrm{T}}$ on both side of Equation (3.5) gives:

$$\frac{d\mathbf{w}}{dt} = \Lambda^{\dagger}\left(\bar{\Theta}F_{\mathbf{g}} + \Theta F_{Img}\right) \tag{3.6}$$

where $\Lambda^{\dagger}$ is the *pseudo-inverse* (Golub and Kahan, 1965) of Λ and is given by $\left(\Lambda^{\mathrm{T}}\Lambda\right)^{-1}\Lambda^{\mathrm{T}}$. $\Lambda^{\dagger}$ acts as a projection matrix, which propagates the information from the pixels on the analytical surfaces to the skeleton points.

Equation (3.6) establishes the analytical surface deformation in *FloatingCanvas*: given the old parameter $\mathbf{w}^{old}$, the new $\mathbf{w}^{new}$ can be updated by:

$$\mathbf{w}^{new} = \mathbf{w}^{old} + \Lambda^{\dagger}\left(\bar{\Theta}^{old}F_{\mathbf{g}}^{old} + \Theta^{old}F_{Img}^{old}\right)\Delta t \tag{3.7}$$

according to the definition of the derivative in Equation (3.6). The functions $F_{\mathbf{g}}$, F_{Img} and Θ are labelled with 'old' because the input $\bar{f}(\mathbf{X}_*)$ in these three functions is calculated using $\mathbf{w}^{old}$, and Δt is a sufficiently small time increment which is set to 0.1 in the algorithm.

The deformation in Equation (3.7) is repeated, and in each iteration, the $\mathbf{w}^{old}$ is substituted with $\mathbf{w}^{new}$ in last iteration. The algorithm stops when the value of $\mathbf{w}$ converges, or pragmatically when the change of $\mathbf{w}$ becomes sufficiently small (<0.26μm (≈0.1 pixel on z-axis) in the implementation).

3.3.3. Searching for tissue layers or their boundaries

In *FloatingCanvas*, different tissue layers or their boundaries can be found by configuring functions $F_{\mathbf{g}}$, F_{Img} and Θ and the 'gravity' force $\mathbf{g}$. To search for the anterior boundary of RNFL, the parameter $\mathbf{w}$ is initialised to be 0 such that the analytical surface is at the top of the A-scan. The $\mathbf{g}$ is set to be $\mathbf{g} = < 0,\ 0,\ 30 >^{\mathrm{T}}$ and the functions $F_{\mathbf{g}}$, F_{Img} and Θ are:

$$F_{\mathbf{g}}\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right) = \left(\nabla Im\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right) + \mathbf{g}\right)\bar{f}(\mathbf{x}_{*_i}) \tag{3.8}$$

$$F_{img}\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right) = -\nabla\left|\nabla Im\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right)\right|\bar{f}(\mathbf{x}_{*_i}) \tag{3.9}$$

$$\Theta\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right) = \begin{cases} 1 & \left|\nabla Im\left(\mathbf{x}_{*_i}, \bar{f}(\mathbf{x}_{*_i})\right)\right| \geq |\mathbf{g}| \\ 0 & otherwise \end{cases} \tag{3.10}$$

where '•' is the dot product; '∇' is the gradient operator when it is applied to an image. For instance, $\nabla Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)$ is the gradient of the image Im at the location $\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)$. Consequently, $\nabla\left|\nabla Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)\right|$ in Equation (3.9) is the gradient of the magnitude of the image gradient $\nabla Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)$. This value can easily be calculated using the second derivative of the image. ∇ is the surface normal operator when it is applied to a function. For example, $\nabla\overline{f}(\mathbf{x}_{*_i})$ is the normal of the surface $\overline{f}(\mathbf{x}_{*_i})$ given by $\left\langle -\frac{\partial\overline{f}(\mathbf{x}_{*_i})}{\partial\mathbf{x}_{*_i}^1}, \ -\frac{\partial\overline{f}(\mathbf{x}_{*_i})}{\partial\mathbf{x}_{*_i}^2}, \ 1 \right\rangle$, which can be analytically calculated from Equation (3.2). The hat '~' above the gradient operator in $\nabla\overline{f}(\mathbf{x}_{*_i})$ in Equation (3.8) and (3.9) means that the normal has been normalised to have a length of 1.

Equation (3.8) acts as a 'gravity' pulling the analytical surface downwards from the top of the A-scan. The Θ function in Equation (3.10) guaranties that this gravity function stops working when the analytical surface comes across a significant gradient that is larger than $|\mathbf{g}|$. Equation (3.9) 'attaches' the surface to the local maximum of the image gradient magnitude when the function Θ outputs 1.

Similarly, the RPE layer is searched for by initialising the surface to be at the bottom of the A-scan, $\mathbf{g} = <0, \ 0, \ -10>^{\mathrm{T}}$ and the functions F_g, F_{Img} and Θ are:

$$F_g\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right) = -\left|\mathbf{g}\right| \tag{3.11}$$

$$F_{img}\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right) = -\nabla Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)\blacksquare\overline{f}(\mathbf{x}_{*_i}) \tag{3.12}$$

$$\Theta\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right) = \begin{cases} 1 & Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right) \le Im_{\max}\left(\mathbf{x}_{*_i}\right) \\ 0 & otherwise \end{cases} \tag{3.13}$$

where $Im_{\max}\left(\mathbf{x}_{*_i}\right)$ is the maximum intensity of the pixels within a 150μm window below the surface point $\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)$.

The force in Equation (3.11) pulls the surface upwards from the bottom of the A-scan when Θ in Equation (3.13) shows that the intensity on the surface is larger than the current local maximum intensity $Im_{\max}\left(\mathbf{x}_{*_i}\right)$. Equation (3.12) 'attaches' the surface to the locations with the maximum local intensity when the function Θ outputs 1.

To search for the posterior RNFL, the surface is initialized to be the detected RPE. Two necessary condition constraints are applied during the search of the posterior RNFL: the intensity on the posterior RNFL should be no less than 70% of the intensity of the RNFL anterior boundary, and the gradient of the RNFL posterior boundary should be 'downward' ($\nabla_z Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)>0$). These two constraints are integrated into the function Θ:

$$\Theta\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right) = \begin{cases} 1 & \left|\nabla Im\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)\right| \ge |\mathbf{g}| \text{ and both constraints are met} \\ 0 & otherwise \end{cases} \tag{3.14}$$

The $\mathbf{g}$ is set to be $\mathbf{g} = <0, \ 0, \ -15>^{\mathrm{T}}$, and the functions F_g and F_{Img} are the same with Equation (3.8) and (3.9) if $\mathbf{x}_{*_i}$ is not in vessel region. Otherwise, F_g and F_{Img} are set to 0. The vessel detection will be described in the subsequent section. The function Θ in Equation (3.14) including the two constrains only brings the analytical surface near to the RNFL posterior boundary, which is eventually decided by the gradients in function F_{Img}.

Benefiting from the segmentation in 3D space, the 4mm × 4mm RNFLT map can simply be calculated as the difference between the segmented anterior and posterior RNFL, without smoothing out or interpolating the individually segmented B-scans (Mujat et al., 2005, Fernández et al., 2005, Ishikawa et al., 2005).

The deformation procedure illustrated in Figure 3-1 demonstrates the intermediate and the final forms of the analytical surface searching for the anterior RNFL boundary. The corresponding curve in the surface was superimposed on one of the B-scans. The B-scan has been cropped so that only the region containing the tissue layers was displayed. The analytical surface is initialized as a flat panel at the top of the image volume. Before the analytical surface contacts the target boundary, it is mostly driven by the gravity $\mathbf{g}$ and the speckle noise in the image. Therefore, the surface becomes noisy and irregular (Figure 3-1(a, d)). It is also possible that the speckle noise and some tissues above the anterior RNFL form strong gradients that set the function Θ to be 1 before the analytical surface is close to the target boundary. An example of such an area is indicated by an arrow in Figure 3-1a. However, the deformation of the analytical surface is driven not only by the force at one point but also by the forces in the neighbouring regions. Therefore, these local exceptions that are not 'supported' by the neighbouring regions affect the deformation less and are overcome in the subsequent iterations (Figure 3-1b). In Figure 3-1(b, e), the force $F_{img}\left(\mathbf{x}_{*_i}, \overline{f}(\mathbf{x}_{*_i})\right)$ from Equation (3.9) 'attaches' the surface to the target boundary in the areas where the surface is sufficiently close to the target boundary. This

procedure is repeated until all parts of the target boundary are found and the surface deformation stops (Figure 3-1(c, f)).
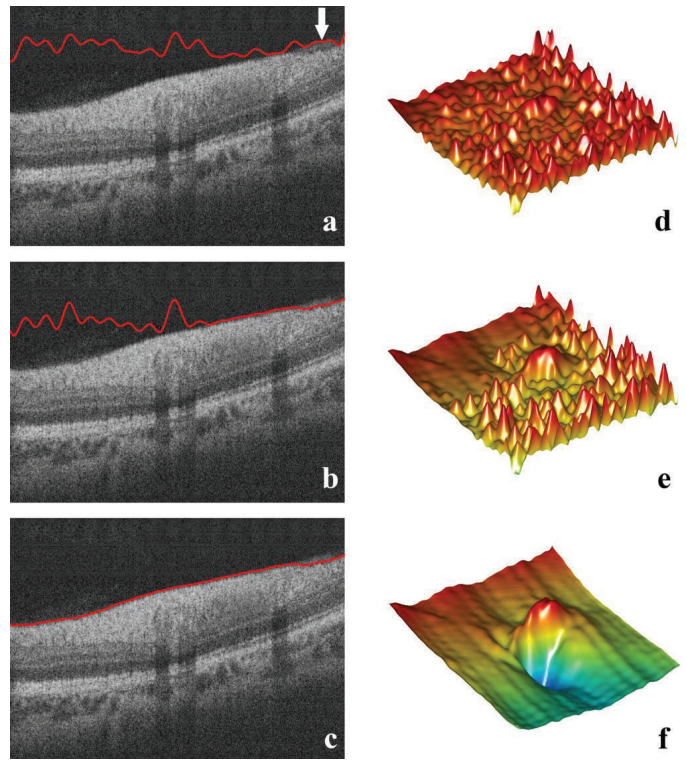
Figure 3-1. Steps in the deformation procedure to search for the anterior RNFL boundary. (a-c) a B-scan in an image volume superimposed with the corresponding curve in the analytical surface at the 30th, 45th and the last (96th) iterations of the deformation procedure. Plates (d-f) show an overview of the analytical surface at these 3 iterations. The arrow in (a) shows an area with strong gradient that sets the Θ function to be 1 before the analytical surface is close to the target boundary.

There are four main free-parameters in *FloatingCanvas*: the skeleton point interval, length-scale l of the Gaussian kernel in Equation (3.1), the time increment Δt in Equation (3.7) and the gravity **g**. The choice of these parameters represents the trade-off between the computational efficiency and the robustness of the algorithm. The skeleton point grid interval and length-scale l are closely related and they quantify how much of the 'neighbourhood' information is taken into account for the segmentation of a particular point. The algorithm considers no information

from the neighbouring area and becomes a segmentation of individual B-scans if these two parameters are very small. On the other hand, large values of these two parameters make the analytical surface too rigid to model the necessary morphology variety of the target surface. Moreover, small skeleton point interval also results in larger $\mathbf{K}(\mathbf{X}, \mathbf{X})$ and $\Lambda^{\dagger}$ matrices in Equation (3.2) and (3.6) and the algorithm is no longer affordable by normal computing platforms. The gravity **g** and time increment Δt do not affect the flexibility of the analytical surface, and they mainly control the convergence performance of the algorithm. Large values of these two parameters causes quicker convergence of the analytical surface deformation, but this risks the deformation procedure 'over-stepping' the target surface and causing convergence on the wrong surface.

3.3.4. Vessel detection

There has been much recent discussion about how blood vessels influence current OCT segmentation algorithms causing bias in estimates of RNFLT (Hood et al., 2008, Hood et al., 2009). In fact, RNFLT tends to be significantly overestimated or underestimated within the area of blood vessels. It is therefore necessary to mark and delineate as far as possible the blood vessels before detecting the RNFL posterior boundary.
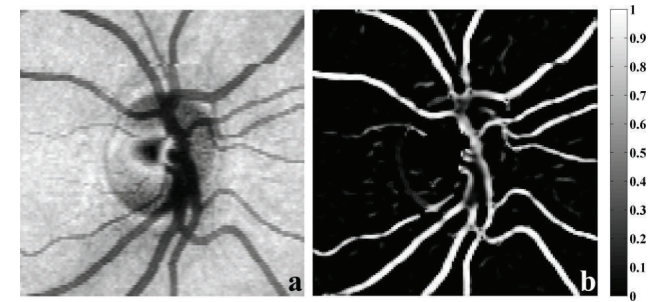
Figure 3-2. En-face image and pixel vesselness. (a) the en-face image calculated by averaging the 50 pixels below and above the detected RPE. (b) the pixel vesselness in grayscale.

FloatingCanvas identifies blood vessels by using the en-face image $\mathbf{EF}(\mathbf{x})$ obtained by averaging the 50 pixels below and above the analytical RPE (Jiao et al., 2005, Fabritius et al., 2009) (Figure 3-2a). This detection scheme computes the 'vesselness' for each pixel in the en-face image. 'Vesselness' is a definition based on the analysis of eigenvalues of the Hessian matrix of image intensity (Frangi et al., 1998):

$$S_{vn} = \begin{cases} 0 & v_{max} \leq 0 \\[2ex] \dfrac{e^{-\frac{(v_{min}/v_{max})^2}{0.5^2}}}{e^{\frac{1}{2}} - e^{-\frac{v_{min}^2 + v_{max}^2}{15^2}}} & v_{max} > 0 \end{cases}$$

where v_{min} and v_{max} are the smaller and larger eigenvalues of a Hessian matrix which consists of the second-order derivatives of the en-face image $\mathbf{EF}(\mathbf{x})$:

$$\begin{bmatrix} \nabla_{x_i^1 x_i^1} \mathbf{EF}(\mathbf{x}_i) & \nabla_{x_i^1 x_i^2} \mathbf{EF}(\mathbf{x}_i) \\[1ex] \nabla_{x_i^2 x_i^1} \mathbf{EF}(\mathbf{x}_i) & \nabla_{x_i^2 x_i^2} \mathbf{EF}(\mathbf{x}_i) \end{bmatrix}$$

where, for example, $\nabla_{x_i^1 x_i^2} \mathbf{EF}(\mathbf{x}_i)$ is the second-order derivative of $\mathbf{EF}(\mathbf{x}_i)$ with respect to $\mathbf{x}_i^1$ and then $\mathbf{x}_i^2$. As it is shown in Figure 3-2b, the 'vesselness' score S_{vn} provides a clear distinction between the vessel and non-vessel pixels, and even a simple thresholding ($S_{vn} > 0.2$) yields a satisfactory result. Eventually, the vessels are detected in the binary image as the area with more than 100 connected pixels.

3.3.5. Optic nerve head approximation

Physiologically, the RNFL near and within the ONH area changes direction and becomes the neural rim of the ONH. The idea of detecting the ONH area in this study is to exclude the ONH when displaying the RNFLT map. Therefore, the method described below was not designed to find an accurate contour line of the ONH, but to derive an approximated area of the ONH. The ONH is detected as the area bounded by the end-tips of the RPE.

The RPE tip is detected using the anterior RNFL and en-face image used in vessel detection. The detected anterior RNFL surface by *FloatingCanvas* is approximated by the combination of a quadratic and a Gaussian surface, which is a similar to a method proposed by Swindale et al when modelling scanning laser ophthalmoscope topography (Swindale et al., 2000). The initial estimate of the ONH centre is set to be the centre of the fitted Gaussian component. The local intensity gradients at every pixel in the en-face image are then calculated in the *x-y* plane. The candidate pixels of RPE tips are required to meet three criteria. First, RPE tips should have a sufficiently large gradient (e.g. above the 75th percentile of the distribution of gradients). Second, considering the gradients of the intensity gradients at each pixel in the en-face image, the gradient of the intensity gradient of qualifying pixels is required to be near to 0. The gradient of the intensity gradient is near to 0 when the intensity gradient is at a local maximum. Third, there are two vectors of interest for each pixel in the en-face image: the local intensity gradient

at the pixel (which is at right-angles to edges) and the vector connecting this pixel with the initial estimate of the ONH centre. It is required that the angle formed by these two vectors is smaller than 45° for the candidate RPE tip pixels. This criterion removes most pixels on vessel edges which also have strong gradient, because the angles between vectors for pixels on vessel edges are generally large (e.g. close to 90°). An ellipse is fit to the candidate RPE tip pixels using a Random Sample Consensus (RANSAC) parameter estimation (used by Li et al (Li et al., 2005) for ellipse detection in noisy data). A cubic spine is then fitted to the pixels that are close to the fitted ellipse to form the approximated contour of the RPE tips. The ONH centre is finally calculated as the geometric mean of the contour. The ONH area is removed when the RNFLT map is displayed.

The algorithms in *FloatingCanvas* described above were implemented under Matlab (version 7.4.0 R2007a, The MathWorks, Inc., Natick, MA).

3.4. Validation

To validate the segmentation algorithm, 26 glaucomatous subjects (mean age of 52 (range 22 to 91) years) and 14 healthy subjects (mean age of 50 (range 16 to 67) years) were recruited. The study was approved by an ethics committee and informed consent, according to the tenets of the Declaration of Helsinki, was obtained prior to examination from each subject. Each subject was imaged 3 times with the RTVue-100 system 4mm×4mm 3D volume scan protocol. Images were acquired in both eyes of each subject during the same testing session by the same observer following the manufacturer instructions. Patient identifiers were removed from the data and the 3D volumes were transferred to a secure computer. *FloatingCanvas* was then applied to these 240 image volumes to extract the RNFLT measurement.

The first validation compared the automated segmentation by *FloatingCanvas* with segmentation by human manual grading. One of the three repeat image volumes of a randomly chosen eye from each subject was randomly selected for manual segmentation. For each selected image volume, the 101 B-scans were evenly divided into 10 sections, each of which contains about 10 B-scans. One B-scan was then randomly chosen from each section for manual segmentation. With 40 subjects, 400 B-scans were manually delineated and the segmented surface positions were compared with those produced by *FloatingCanvas*. The mean and standard deviation (SD) of signed and absolute difference between the manual and *FloatingCanvas* segmentation were then evaluated for healthy and glaucomatous subjects respectively.

The second validation hypothesis was that, if the method is reliable, the estimated RNFLT from the *FloatingCanvas* segmentation would be equivalent across different width annuli around the

ONH. Therefore, overall mean and quadrant mean RNFLT were estimated using two different calculation annuli: wide (0.58mm wide from the inner margin radius of 1.170mm) and narrow (0.29mm wide from the inner margin radius of 1.315mm) annuli (Figure 3-3). The two annuli were centred on the same circle with a radius of 1.460mm, with one annulus twice as the width of the other.
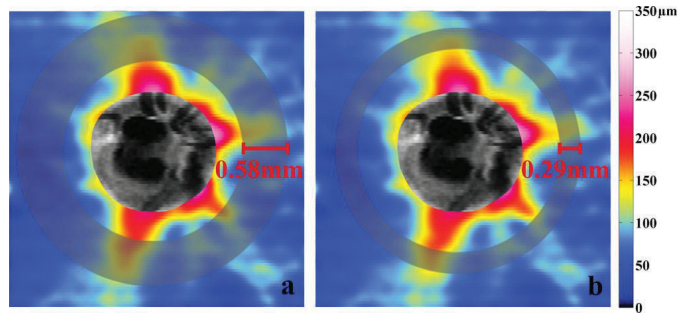
Figure 3-3. An illustration of the location and width of two different annuli used to calculate the mean and quadrant RNFLT. The wide (0.580mm wide from the inner margin radius of 1.170mm) annulus was twice as wide as the narrow one (0.290mm wide from the inner margin radius of 1.315mm). Both annuli were centred on the same circle with a radius of 1.460mm.

Moreover, one of the most important parameters for the quantitative analysis of imaging techniques is the reproducibility, which directly relates to the reliability of the techniques and their ability to separate physiological changes from measurement variability and also to detect progressive RNFL loss over time. The reproducibility of RNFLT measurements was evaluated by estimating test-retest variability based on the three repeated measurements and the coefficient of variation (CV) for mean and quadrant RNFLT. We defined test-retest variability of RNFLT, expressed in micrometers, as twice the SD of the three repeated measurements. The coefficient of variation was calculated as the SD of the three measurements divided by the mean.

3.5. Results

FloatingCanvas segmented the retinal structures in all 240 SD-OCT volume scans without clinically spurious results. On average, it took 5.6±1.2 minutes to process a large image volume (513×101 A-scans) on one core of a Intel Core 2 Duo 2.4GHz CPU and 8GB RAM with single thread.

Figure 3-4 shows RNFLT maps from a healthy subject and a glaucomatous patient. The RNFLT map was colour-coded in micrometers with the 'hotter' colour denoting thicker RNFL. It should

be noted that the healthy retina (Figure 3-4a) has a much thicker RNFL in the superior and inferior quadrants compared with the nasal and temporal quadrants. This is consistent with the known normal retinal anatomy. The reduced RNFLT, especially in the superior and inferior quadrants, in the glaucomatous eye (Figure 3-4b) can be observed and is consistent with clinical knowledge.
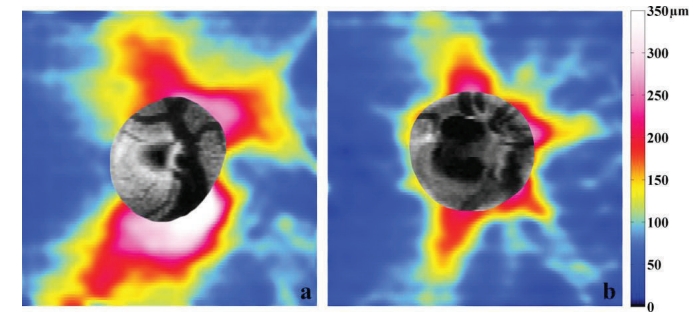
Figure 3-4. The RNFLT map calculated from the segmented retinal structures of a healthy subject (a) and a glaucomatous patient (b). The RNFLT map was colour-coded in micrometers. Note the significantly thicker RNFLT, especially in the superior and inferior areas, in the healthy example (a) compared with that of the glaucomatous example (b).

The mean and SD of the signed and absolute difference between the manual and *FloatingCanvas* segmentation for glaucomatous and healthy subjects are given in Table 3-1.

Surface	Glaucoma (MEAN±SD, μm)		Healthy (MEAN±SD, μm)	
	Signed	Absolute	Signed	Absolute
All surfaces	-0.8±2.1	4.8±2.7	-0.7±2.3	3.6±2.1
Anterior RNFL	1.2±0.7	3.3±0.9	1.5±0.9	2.2±0.7
Posterior RNFL	-1.9±1.8	6.8±1.8	-1.7±2.5	4.7±1.4
RPE	−1.6±2.0	4.3±1.8	−1.8±1.6	3.5±1.9

Table 3-1. Mean and SD of signed and absolute difference between the manual and *FloatingCanvas* segmentation for glaucomatous and healthy subjects. The difference values were summarized with all 3 surfaces and anterior, posterior RNFL and RPE respectively.

For all manually segmented B-scans, the mean and SD of the absolute difference for all three boundaries are 4.3±2.0μm. Therefore, on average, the *FloatingCanvas* segmentation differs from that of the human expert by 4.3μm, which is equivalent to 1.7 pixels on *z*-axis. The mean absolute difference between the manual and algorithm segmentation is relatively higher for glaucomatous retina compared with that of healthy retina but the difference is not statistically significant.

RNFLT profile in the healthy and glaucomatous eyes is summarized in Table 3-2. There were no statistically significant differences between RNFLT measurements using the calculation annuli with different widths, which suggests that *FloatingCanvas* is robust and stable across the 3D volume. The quadrant RNFLT shows a difference between healthy and glaucomatous eyes. In general and on average, the healthy eyes, as expected, have a thicker RNFL, especially in the superior and inferior quadrants.

	Glaucoma (MEAN±SD, µm)		Healthy (MEAN±SD, µm)	
Calculation circle	**0.58mm**	**0.29mm**	**0.58mm**	**0.29mm**
Total RNFLT	86±16	85±18	110±6	111±7
Temporal RNFLT	71±16	70±13	77±16	77±14
Superior RNFLT	98±21	97±22	133±15	134±12
Nasal RNFLT	76±16	75±16	96±14	97±13
Inferior RNFLT	99±30	98±32	134±16	135±15

Table 3-2. Mean and SD of total and quadrant retinal nerve fibre layer thickness (RNFLT) of healthy and glaucomatous eyes. RNFLT was determined with two types of calculation annuli (0.58mm and 0.29mm wide, respectively).

The reproducibility of the segmented RNFLT using SD-OCT was compared with the typical reproducibility of StratusOCT as reported in the literature using the standard scan protocol (Budenz et al., 2005) in Table 3-3 (healthy subjects) and Table 3-4 (glaucomatous subjects). From Table 3-3 and Table 3-4, it can be seen that test-retest reproducibility in RNFLT measurements is better for both healthy and glaucomatous eyes with SD-OCT. RNFLT measurements were least reproducible in the nasal quadrant, with both SD-OCT and StratusOCT, while the segmented nasal measurement with SD-OCT showed markedly better reproducibility (~7µm vs 10.2µm in both normal and glaucomatous eyes). Moreover, RNFLT measurements in glaucomatous eyes were more variable than those of healthy eyes with both SD-OCT and StratusOCT, but SD-OCT showed much less variability and better reproducibility compared with StratusOCT, especially in the superior and inferior quadrants, which are the most important areas for glaucoma diagnosis. These results are consistent with the literatures about the reproducibility on another SD-OCT platform (Cirrus, Carl Zeiss Meditec, CA, USA) (Kim et al., 2009, Mwanza et al., 2010).

	0.58mm annulus		0.29mm annulus		Stratus OCT (Budenz et al., 2005)	
	CV(%)	Test-retest variability (µm)	CV(%)	Test-retest variability (µm)	CV(%)	Test-retest variability (µm)
Total RNFLT	1.5	3.3	1.5	3.3	1.7	3.5
Temporal RNFLT	3.8	5.8	3.9	6.0	5.1	7.5
Superior RNFLT	2.6	7.5	3.0	8.2	3.8	9.6
Nasal RNFLT	3.5	6.6	4.3	8.0	6.7	10.2
Inferior RNFLT	1.5	4.4	1.6	4.6	3.7	9.7

Table 3-3. Coefficient of variation and test-retest variability of total and quadrant retinal nerve fiber layer thickness (RNFLT) of healthy retina measured by *FloatingCanvas*. RNFLT was calculated at two widths of calculation annuli (0.58mm and 0.29mm). A typical reproducibility of StratusOCT is given for comparison.

	0.58mm annulus		0.29mm annulus		Stratus OCT (Budenz et al., 2005)	
	CV(%)	Test-retest variability (µm)	CV(%)	Test-retest variability (µm)	CV(%)	Test-retest variability (µm)
Total RNFLT	2.2	4.0	2.2	4.1	3.7	5.2
Temporal RNFLT	4.5	5.2	5.0	5.6	5.3	5.6
Superior RNFLT	3.8	7.8	4.2	8.6	6.4	10.7
Nasal RNFLT	4.9	6.7	6.1	7.4	9.0	10.2
Inferior RNFLT	2.1	4.7	1.9	4.3	6.6	10.6

Table 3-4. Coefficient of variation and test-retest variability of total and quadrant retinal nerve fiber layer thickness (RNFLT) of glaucomatous subjects measured by *FloatingCanvas*. RNFLT was calculated at two widths of calculation annuli (0.58mm and 0.29mm). A typical reproducibility of StratusOCT is given for comparison.

3.6. Discussion

FloatingCanvas has been developed as an effective 3D segmentation method for SD-OCT volume scans centred on the ONH. It is important that automatic segmentation should be compared with the manual segmentation as the gold standard. *FloatingCanvas* segmentation demonstrated good agreement with the human manual grading. It also provides a repeatable estimation of the RNFLT in the image volume. As opposed to the sparse area covered by the circular scans used in StratusOCT, the RNFLT maps cover a larger and clinically more useful area allowing for a more reliable measure of the RNFLT. The method has been tested on 240 3D volume scans acquired from both healthy and glaucomatous eyes of 40 subjects without spurious results under visual inspection. The results indicate that the RNFLT map gives a highly reproducible evaluation of a larger retina area compared with the last-generation time domain StratusOCT.
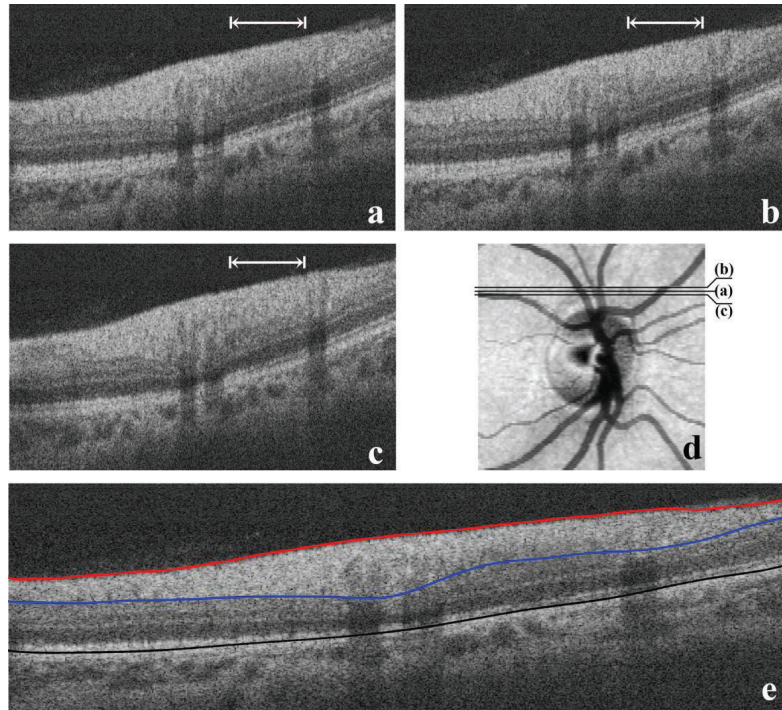
Figure 3-5. The benefit of segmentation in 3D space. (a) a B-scan with a segment of RNFL with an indistinct posterior boundary. The region with an indistinct posterior RNFL boundary is denoted by a white arrow in the image. (b, c) the RNFL boundary in the B-scans adjacent to the one in (a) is more distinct; the location of indistinct RNFL boundary in (a) is also marked. (d) the locations of B-scans (a, b, c) in the image volume. (e) the segmented RPE and anterior and posterior RNFL boundaries in image (a) by *FloatingCanvas*.

The main novelty of *FloatingCanvas*, compared with previous algorithms, is that it processes the whole volume of data in its 'raw' format without pre-processing, such as filtering or image averaging. The algorithm needs no segmentation of individual B-scans so that it benefits from the fact that the covariance among neighbouring B-scans helps to make the analytical surface deformation robust to local noise or errors in individual B-scans. To illustrate the benefit of the volume segmentation, an individual B-scan with local low contrast is shown in Figure 3-5a. Although this image volume has good overall image quality (image quality score >60 in RTVue system), a part of the posterior RNFL boundary in this B-scan is not well defined due to the low contrast in the region marked by the white arrow in Figure 3-5a. Segmentation of the indistinct

posterior RNFL boundary in this region would pose problems even for expert clinicians and B-scan-based segmentation algorithms would give spurious results in this case. However, this potential source of segmentation error can be resolved by taking into account of the neighbouring B-scans. Figure 3-5(b, c) show the B-scans adjacent to the one in Figure 3-5a. The locations of these three B-scans in the image volume are shown in Figure 3-5d. The adjacent B-scans have a more distinct posterior RNFL boundary in the region where it is indistinct in Figure 3-5a. The covariance among B-scans modelled by the GP model in *FloatingCanvas* can 'borrow' the information from the neighbouring B-scans to aid the segmentation in a local area. Therefore, the posterior RNFL boundary in Figure 3-5a can be correctly identified by the algorithm (Figure 3-5e) even if the information is incomplete in this individual B-scan.

Furthermore, *FloatingCanvas* is computationally efficient as a segmentation algorithm in 3D space. There are two computationally intensive components in the algorithm: 1) the matrix inversion involved in the calculation of the projection matrix $\Lambda^\dagger$ in Equation (3.7). The time complexity of matrix inversion scales as $O(n^3)$ given the number of skeleton points n. However, the projection matrix $\Lambda^\dagger$ is not changed during the deformation of an analytical surface and thus only need to be computed once before the loop of the deformation. Practically, $\Lambda^\dagger$ can be pre-computed before the algorithm and loaded into memory when it is needed because the $\Lambda^\dagger$ matrix is decided only by the skeleton point interval and length-scale l which are all fixed parameters in the algorithm. This costs nearly no time on a modern computing platform; 2) the matrix multiplication and addition in Equation (3.7). The right part of Equation (3.7) has to be evaluated at every iteration of the deformation. However, matrix multiplication and addition are all low cost computational operations and can thus be implemented efficiently. Therefore, *FloatingCanvas* has a lower order polynomial computational complexity that is as efficient as the 3D graph search approach (Li et al., 2006). Garvin et al (Garvin et al., 2009) and Quellec et al (Quellec et al., 2010) have demonstrated that by using improved implementation, the execution time of the 3D graph search algorithm was significantly reduced compared with their original implementation (Garvin et al., 2008). Similarly, the current Matlab implementation of *FloatingCanvas* is relatively slow due to the nature of this interpreted programming language and the lack of the implementation optimisations, but the computational efficiency would be significantly improved if it is implemented using C/C++ programming language, multi-threading techniques together with the optimisation using higher order gradient information during the search.

FloatingCanvas was applied on the SD-OCT scan centred on the ONH which is designed for the assessment of the RNFLT and the optic disc; these are the standard and important

examinations for the management of glaucoma. Therefore, the algorithm removed other intra-retinal boundaries by applying the constraints in Equation (3.14) to allow for the direct search of the posterior RNFL. The study of *FloatingCanvas* for segmentation of macula scans with more intra-retinal boundaries is part of the ongoing work. To search the intra-retina layers, the parameters in *FloatingCanvas* were altered such that the gravity force became dynamic and was decided by the gradient profile that analytical surfaces pass through during the search.

Overall, *FloatingCanvas* provides a robust and efficient delineation and evaluation of RNFL and RPE structures around the ONH. It can be a useful tool for clinically interpreting SD-OCT volumes for glaucoma diagnosis. The reproducible results can potentially be used for monitoring RNFLT changes in longitudinal studies. The larger scan area also improves the chance of achieving a stronger relationship of RNFLT measurements with visual function. This structure-function relationship will be extensively studied in chapter 4.

Chapter 4 : Predicting visual function from the measurements of retinal nerve fibre layer structure

4.1 Summary

Damage from glaucoma is clinically assessed by both retinal structure and visual function. It was shown in last chapter that the optical coherence tomography (OCT) technique, once equipped with quantification tools like the segmentation algorithm, can be used to extract important clinical *knowledge* about retinal structures such as nerve fibre layer (RNFL) thickness. On the other hand, visual function loss caused by glaucoma is clinically evaluated with standard automated perimetry (SAP) reviewed in chapter 1.

The structural and functional measurements yield complex data that are often difficult to interpret. In this chapter, the extracted information about RNFL structure from imaging technique is taken into the next iteration of *knowledge mining*. A mathematical framework is developed and evaluated to correlate the retinal structure and visual function. RNFL thickness (RNFLT) measurements from GDxVCC scanning laser polarimetry (SLP) and visual field (VF) sensitivity from standard automated perimetry were made available from 535 eyes from three centres. In a training dataset, structure-function relationships were characterized using linear regression and a type of neural network: Radial Basis Function (RBF) customised under a Bayesian framework. These two models were used in a test dataset to 1) predict sensitivity values at individual VF locations from RNFLT measurements and 2) predict the spatial relationship between VF locations and positions at a peripapillary RNFLT measurement annulus. Predicted spatial relationships were compared with a published anatomical structure-function map (Garway-Heath et al., 2000). The Bayesian Radial Basis Function (BRBF) generates clinically useful relationships that relate topographical maps of RNFL measurement to VF locations and allows the VF sensitivity to be predicted from structural measurements. This method may allow clinicians to evaluate structural and functional measures in the same domain. Because the spectral domain OCT (SD-OCT) is only recently established as a clinical tool for the assessment of RNFL, a large-scale dataset with both SD-OCT images and the VFs is still not available. Therefore, the method is demonstrated on the RNFLT acquired by SLP technique. However, the developed mathematical framework does not make specific assumption about the imaging and SAP techniques and can thus be generalized to other types of measures.

[HL], Nicolaas Reus [NR], Paul Healey [PH], Paul Mitchell [PM], Tuan Ho [TH] and David Garway-Heath [DGH]. Some of the data was collected by PS and other data was made available retrospectively from studies conducted by HL, NR, PH and PM. Some of the manuscript was edited by TH and the work was supervised by DC and DGH. Everything else described in this chapter was completed by HZ. Some of the technical detail of the method was also published in proceedings from *The 3rd International Conference of Bioinformatics and Biomedical Engineering published by IEEE* ((Zhu et al., 2009); See List of supporting publications). Some of the initial work was also presented as a prize-winning read paper at the 17th International Symposium of the Imaging and Perimetry Society Meeting in Portland, Oregon USA in July 2006 (See List of supporting publications).

4.2. Purpose and previous studies

Assessing the way in which structural and functional measures in glaucoma interact is clinically important. Visual loss is assumed to follow from, and correlate to, structural loss caused by the disease process. It would be clinically useful to know the magnitude and location of structural loss that will result in visually important functional loss. However, current clinical devices for measuring structural and functional deficits are far from accurate and have imperfect precision. Structural measures of optic nerve head (ONH) topography or RNFLT can be made using optical imaging techniques, but these only provide surrogate measures of the biological variable of real interest, namely retinal ganglion cell (RGC) count and function. At the same time, SAP, the clinical cornerstone of functional testing in glaucoma, is subject to considerable measurement variability and is also a poor surrogate for RGC count and function (Anderson, 2006). Despite their limitations, these techniques are currently central to the diagnosis and management of glaucoma. It would, therefore, be beneficial if structure and function measurements were directly linked in some way, allowing clinicians to corroborate decisions by considering the measurements in tandem.

There have been several studies attempting to quantify the structure-function relationship using clinical measurements. These typically proceed by taking one summary value to represent function (for example Mean Deviation, MD, of the visual field from SAP) and one number to represent the structural data (for example, average neuroretinal rim area or mean RNFLT), then assessing the curve-linear (e.g. log-linear) or monotonic association between the two variables via R^2, Pearson or Spearman coefficients (Brigatti and Caprioli, 1995, Weinreb et al., 1995, Iester et al., 1997, Teesalu et al., 1997). This approach suffers from two major flaws: the use of summary data loses spatial information and may reduce power, and these association measures and regression models assume a particular shape of the relationship. Furthermore these analyses fail to take account of spatial associations in the data, an integral attribute of glaucomatous loss.

While some studies have attempted to overcome these limitations by analysing smaller regions from the structural measures (for example, sectoral values of rim area) and groups or individual points in the visual field (VF) (Garway-Heath et al., 2002, Garway-Heath et al., 1999, Anton et al., 1998, Gardiner et al., 2005, Mai et al., 2007, Bowd et al., 2006, Schlottmann et al., 2004), it is anticipated that much more could be gained by analysing the data in its 'high resolution form'. For example, in scanning laser polarimetry (SLP), RNFLT estimates are yielded over an image space of several thousand pixels; this is high-dimensional data and any method linking structural measures to the 50 or so individual points in the VF ideally should take this into account. Moreover, individual points from both structure (pixel or sector values) and function (areas of VF or individual locations) are more likely to interact as 'groups' rather than single independent measurements.

In recent years, a special class of artificial neural networks using RBF have received considerable attention (Powell, 1987, Orr, 1996, Poggio and Girosi, 1990, Bishop, 1996). This class of networks differs from the artificial neural networks (ANNs) that use a multilayer perceptron (MLP) approach, in that the nonlinearity of the model is embedded in a hidden layer of the network: this hidden layer consists of basis functions. The key element of the RBF is that predictions can be made in a multi-dimensional space that consists of 'mini-distributions' of possible predictions, analogous to a kernel or window when interpolating data. RBF networks are particularly useful when mapping two- or three-dimensional images where interpolation is a prerequisite, with some image features being preserved while others not necessarily being mapped exactly. In this way they seem most suitable to the problem of mapping the individual points from the structural and functional measures. Moreover, the RBF method developed for this study is formulated under a Bayesian probabilistic framework in order to tackle the considerable variability in measurements as well as to form an automatic learning process: this customisation leads to what is denoted as a BRBF network.

The purpose of this study is to develop a general statistical method for linking clinically used measures of structure and function. The use of the method is demonstrated on measurements of the RNFLT derived from SLP with variable corneal compensation (GDxVCC) and individual VF locations from SAP with the aim of providing predictions of one measure from another and to generate a topographical map of the spatial relationship between the imaged structure data and individual points in the visual field. It is imperative that any new method should be developed and tested on more than one data set: for this work, three large data sets collected at three independent clinical centres are used. The method, which aims to make predictions, is compared to the correlation approach for mapping structural and functional measures typified

by Gardiner and colleagues (Gardiner et al., 2005). It is also demonstrated that the method can be extended to high resolution measurements.

4.3. Subjects and datasets

The study sample was derived from three independently acquired populations from Moorfields Eye Hospital, London (MEH), Rotterdam Eye Hospital, the Netherlands (REH) and the Blue Mountains Eye Study, Australia (BMES)

4.3.1. Moorfields Eye Hospital (MEH) data

Thirty-four healthy subjects (34 eyes), 43 glaucomatous patients (43 eyes) and 30 ocular hypertension subjects (30 eyes) were enrolled. Inclusion criteria for the healthy subjects were a normal VF, intraocular pressure (IOP) <21 mm Hg, no previous history of ocular disease, and no family history of glaucoma in a first degree relative. For the glaucomatous patients, inclusion criteria were previously raised IOP (>21mmHg), reproducible VF defects and absence of other disorders that might cause VF loss. The VF defects were defined as a defect of two or more contiguous points in the Humphrey pattern deviation probability maps with P <1% loss or greater, or three or more contiguous points with P <5% loss or greater, or a 10 dB difference across the nasal horizontal midline at two or more adjacent points in the total deviation plot. Ocular Hypertension subjects were recruited from clinic on having a raised IOP >21 mmHg in two consecutive visits and having at least two normal VFs. Optic disc appearance was not used to categorize subjects. However, the optic disc was evaluated to exclude anatomical abnormalities such as coloboma or drusen. For all participants, one eye was randomly selected for study if both were eligible. All subjects had visual acuity of 20/40 or better, with ametropia <7 diopters, and had no other significant ocular abnormality or concomitant ophthalmic condition.

4.3.2. Rotterdam Eye Hospital (REH) data

Forty six healthy subjects (46 eyes) and 76 glaucomatous patients (76 eyes) took part in a study that has been described in detail in Reus and Lemij (2005) (Reus and Lemij, 2005). In short, healthy subjects had normal VF, optic discs without structural abnormalities, IOP <21 mmHg, no previous history of ocular disease, and no family history of glaucoma. The patients had a glaucomatous appearance of the optic disc and a corresponding nerve fibre bundle VF defect, as described by Keltner et al (2003) (Keltner et al., 2003) with SAP. All subjects had a visual acuity of 20/40 or better, and had no other significant ocular abnormality. For all participants, one eye was randomly selected for the study.

4.3.3. Blue Mountains Eye Study (BMES) data

VFs and images were available for 1540 subjects from a large population-based study of visual impairment, common eye diseases, and other health conditions from an elderly community in Australia. A description of the Blue Mountain Eye Study (BMES) is given elsewhere (Mitchell et al., 1996). Two hundred and thirty healthy subjects (230 eyes) and 76 patients (76 eyes) diagnosed with glaucomatous optic neuropathy were selected from this population under strict measurement quality criteria (described below). Only one randomly selected eye per subject was used throughout. The criteria used for defining glaucomatous VF loss was an abnormal Humphrey Glaucoma Hemifield Test (GHT) plus one or more of the following VF defect classifications (top row of test points in 24-2 pattern were excluded to reduce the effect of lid artefact), which could not be explained on the basis of non-glaucomatous ocular, or neurological, causes: (1) at least 4 contiguous points on pattern deviation plot depressed at p<0.5% level; (2) at least two horizontal points in nasal step locations with pattern deviation plot depressed at p<0.5% level; (3) advanced glaucomatous field loss (hemispheric or severe generalised field loss, with residual temporal or central islands). Glaucoma was diagnosed if glaucomatous VF loss was present, combined with matching neural rim loss at the optic disc, and gonioscopy showed no evidence of angle-closure, rubeosis or secondary glaucoma, other than pseudoexfoliation syndrome.

All three datasets were collected in accordance with the tenets of the Declaration of Helsinki from studies that had local research ethics committee approval with all participants giving informed consent. Data were anonymized and transferred to a secure database held at City University London.

4.3.4. Measurements and data

This study made use of measurements of the RNFL yielded from GDxVCC SLP (Carl Zeiss Meditec AG, Jena, Germany) and individual VF locations from SAP. The principles of SLP have been described by Greenfield et al (Greenfield, 2002). This instrumentation estimates the thickness of the peripapillary RNFLT by measuring the summed retardation in a polarized scanning laser beam reflected from the fundus. Retardation measurements at various points around the optic disc are used to construct a thickness map of the RNFL in micrometers. SLP with variable corneal compensation (VCC) has been shown to improve the estimates of RNFL thickness compared to earlier versions of the technology (Reus et al., 2003). All GDxVCC images, from all three centres, had quality scores greater than or equal to 8 and typical scan scores greater than or equal to 80. Single GDxVCC images were available from all the subjects from all the centres. The 64-sector thickness profile on the peripapillary annulus (provided by

the GDx software) for each subject and the raw images (.mif files) were transferred to a secure database.

The VF measurements were acquired in all cases using a Humphrey Field Analyzer (HFA) II (Carl Zeiss Meditec, CA, USA) using either full-threshold testing or the Swedish Interactive Threshold Algorithm (SITA) Standard test program in the standard 24-2 test pattern. For the MEH and REH data, all VFs were considered reproducible as well as reliable. The VFs from MEH data were all tested with the SITA standard program. VF reliability indices applied were: fixation losses ≤15%; and false-positive and false-negative response rates ≤25%. In the REH data, 29 healthy subjects and 73 glaucomatous subjects were tested with the full-threshold program and the others with the SITA standard program. Reliability indices applied were: fixation losses ≤25%; and false-positive and false-negative response rates ≤20% for full-threshold test and ≤7% for SITA standard test. Higher false-negative response rates were accepted in eyes with advanced field loss (up to 33%). All VFs in BMES were tested with the full-threshold program. Stricter reliability criteria were applied: fixation loss, false positive and false negative all ≤15%. In all eyes, the two VF test points above and below the blind spot were excluded and the remaining 52 raw sensitivity values (dB) were transferred to a secure database for analysis.

In this study, the MEH and REH datasets (229 eyes in total) were used for development of the models, and the BMES dataset (306 eyes in total) was used for independent testing in order to demonstrate the generalization of the method. The image and VF quality criteria were used only for the purpose of selecting reliable measurements and were not used in the modelling. Moreover, the proposed model does not need the initial labelling of the subject data as being from glaucomatous or healthy eyes, and under the Bayesian framework, no manual input parameters are required.

4.4. Statistical Models

With the task of correlating two types of measurements, the study aimed to generate a model that will 'map' an individual structural measurement (RNFLT values from a software generated peripapillary profile or an individual RNFLT value at a pixel in the image) to a sensitivity (dB) value at an individual VF location. Such a model will allow the prediction of a VF sensitivity value from RNFLT values and vice versa. The remaining part of this section compares the technical aspects of the BRBF model to the linear model that has been used widely to assess the association between structural and functional measurements.

4.4.1. Linear regression model

Starting with a linear model, consider the case where an individual VF sensitivity value denoted $\hat{y}_d$ (where d is one of the 52 locations in a 24-2 HFA VF) is predicted from a series of RNFLT values denoted $\mathbf{x}_i$ for i from 1 to m. This can be expressed by the following over-simplified but illustrative equation:

$$\hat{\mathbf{y}}_d = \mathbf{w}_{d1}\mathbf{x}_1 + \mathbf{w}_{d2}\mathbf{x}_2 + ... + \mathbf{w}_{dm}\mathbf{x}_m + c_d \qquad (4.1)$$

where c_d is a constant offset. The symbol on $\hat{\mathbf{y}}_d$ indicates that it is a prediction rather than the real measured value denoted by $\mathbf{y}_d$. In this example, the equation has 64 peripapillary thickness profile values (m=64) each with its own coefficient which quantifies the contribution of each $\mathbf{x}$ value to the prediction. So each $\mathbf{y}$ value can be predicted by a 'linear combination' of $\mathbf{x}$ values. With some actual data some real numbers for the $\mathbf{w}$ terms can be found by least squares regression; this attempts to 'fit' an equation that minimises the difference between the predicted and measured values. This yields an individual equation for each $\mathbf{y}$ that can be enumerated across all the points to predict a complete VF from a given vector of $\mathbf{x}$ values. This classic multiple linear regression can be adapted to select only those $\mathbf{x}$ values that are statistically significant for the prediction of $\mathbf{y}$ values using techniques such as stepwise multiple regression (Hocking, 1976) with, for instance, the forward selection scheme. In a linear model described by Equation (4.1), the $\mathbf{w}$ values (divided by their standard errors) quantify the amount of meaningful contribution made by $\mathbf{x}$ values to predict the $\mathbf{y}$ values. The largest absolute $\mathbf{w}$ term (with respect to the variability in estimating the term) would indicate the $\mathbf{x}$ value that affects the $\mathbf{y}_d$ value the most, in the sense that change in this $\mathbf{x}$ value results in the largest change in $\mathbf{y}_d$. Similarly, the next largest absolute $\mathbf{w}$ term would indicate the second most important term and so on. Equivalently, one could look at the relationship between the $\mathbf{y}_d$ value and each $\mathbf{x}$ value separately and simply calculate the correlation coefficient based on the raw data or on the ranks of the data (Spearman's rho) and end up with a similar result. Loosely speaking this is the approach of Gardiner and colleagues (Gardiner et al., 2005), who used neuroretinal rim area estimated from scanning laser ophthalmoscope measurements (Heidelberg Retina Tomograph, Heidelberg Engineering, Heidelberg, Germany) as the surrogate structural measurement for glaucomatous damage. Particularly, in the linear model implemented for this study, the visual field sensitivity was 'unlogged' from the decibel value and the prediction from this model was converted back to the decibel scale when compared with the measured sensitivity. We will refer to this method as a 'classic linear model'.

This classic linear model approach makes several restrictive and incorrect assumptions about the data. First it assumes that each $\mathbf{x}$ value is independent of all the other $\mathbf{x}$ values (Cohen, 1968), whereas in reality the $\mathbf{x}$ and $\mathbf{y}$ values are topographically and physiologically related and may interact as groups. Although one could try to demarcate these groups based on a physiological relationship between the $\mathbf{x}$ values, or an anatomical map, it would be preferable for the groups to be 'learnt' from the data rather than imposing any relationship from incomplete prior knowledge. Second, this approach assumes that the relationship between $\mathbf{y}$ and $\mathbf{x}$ is either linear or linear after some transform (typically logarithmic). In reality, the relationship between $\mathbf{y}$ and $\mathbf{x}$ may be more complex with the nature of the association probably changing across the measurement range of $\mathbf{y}$. Put simply, at different stages of disease, the apparent relationship between $\mathbf{y}$ and $\mathbf{x}$ could switch from being linear, to noisy, to curvilinear and occasionally being censored due to the imprecision and range of the measurement. This notion that the association may change at different levels of functional loss was also asserted by published studies (Harwerth et al., 2005, Gonzalez-Hernandez et al., 2009). The third difficulty with the classic linear model is that outlier points exert an overly strong influence, and can yield a false association.

4.4.2. Bayesian radial basis function

The proposed method to link the structural and functional measurements is formalized by a function $\mathcal{G}$, which would predict functional measurement $\mathbf{y}$ from a structural one $\mathbf{x}$: $\mathbf{y} = \mathcal{G}(\mathbf{x})$. This sets a mathematical framework for the question: what would be the corresponding functional change ($\Delta\mathbf{y}$) for the structural change $\Delta\mathbf{x}$: $\Delta\mathbf{y} = \mathcal{G}(\mathbf{x} + \Delta\mathbf{x}) - \mathcal{G}(\mathbf{x})$? Because we are interested in subtle structural change to model the slow progression of RNFL damage, we assume that $\Delta\mathbf{x}$ becomes very small and tends to 0. The structure-function relationship is then defined with the general equation:

$$\underset{\Delta\mathbf{x}\to 0}{Lim}\frac{\Delta\mathbf{y}}{\Delta\mathbf{x}} = \underset{\Delta\mathbf{x}\to 0}{Lim}\frac{\mathcal{G}(\mathbf{x} + \Delta\mathbf{x}) - \mathcal{G}(\mathbf{x})}{\Delta\mathbf{x}} = \nabla_{\mathbf{x}}\mathcal{G} \qquad (4.2)$$

where $\mathcal{G}$ is differentiable w.r.t. $\mathbf{x}$, and $\nabla_{\mathbf{x}}\mathcal{G}$ is the gradient of $\mathcal{G}$ at $\mathbf{x}$. Since we are sampling subjects rather than considering the whole population, the final term in this equation must be expressed as a statistical expectation of $\nabla_{\mathbf{x}}\mathcal{G}$ or, for simplicity, the mean of $\nabla_{\mathbf{x}}\mathcal{G}$.

The BRBF was extended from the Relevance Vector Machine (RVM) (Tipping, 2000), where the output was originally only one-dimensional. The extension was similar to the model derived by Thayananthan et al (Thayananthan et al., 2006). In particular, VF sensitivity is assumed to be:

$$\mathbf{y}_d^n = \mathbf{w}_d^{\mathrm{T}}\phi(\mathbf{x}^n) + \varepsilon_d \qquad (4.3)$$

where $\mathbf{y}_d^n$ is the dth element in the measurement of the nth subject, $\mathbf{x}^n$ is the RNFLT measurement (64-point profile or SLP image) vector, $\mathbf{w}_d$ is a weight vector, ε_d is an additive zero-mean Gaussian noise $\mathcal{N}(0, \delta_d^2)$ with variance δ_d^2, and the radial basis function vector $\phi(\mathbf{x}^n)$ is defined to be M+1 dimensional for M bases: $\phi(\mathbf{x}^n) = \left(1, \phi_1(\mathbf{x}^n), \phi_2(\mathbf{x}^n),..., \phi_M(\mathbf{x}^n)\right)^{\mathrm{T}}$, where each element is a radial basis function with centre θ_m and an isotropic covariance:

$$\phi_m(\mathbf{x}^n) = \exp\left(-\eta_m\left(\mathbf{x}^n - \theta_m\right)^{\mathrm{T}}\left(\mathbf{x}^n - \theta_m\right)\right) \qquad (4.4)$$

If all weight vectors $\mathbf{w}_d$ are organized into a matrix $\mathbf{W}$ columnwise, then using Bayesian methodology a prior over $\mathbf{W}$ is assigned:

$$\mathbf{w}_d \sim \mathcal{N}(0, \Psi) \quad \text{and} \quad p(\mathbf{W}\,|\,\Psi) = \prod_d \mathcal{N}(0, \Psi) \qquad (4.5)$$

where Ψ is a diagonal matrix whose elements are $\alpha_1^{-1}, \alpha_2^{-1},..., \alpha_M^{-1}$ on diagonal and 0 otherwise. Each α_m^{-1} represents the average variance of the weights for the mth basis.

According to Bayesian methodology, priors of hyper-parameters are defined over α_m and $\beta_d = \delta_d^{-2}$:

$$p(\mathbf{\alpha}) = \prod_m \mathrm{Gamma}(\alpha_m\,|\,a,\,b), \quad p(\mathbf{\beta}) = \prod_d \mathrm{Gamma}(\beta_d\,|\,c,\,d) \qquad (4.6)$$

where $\mathbf{\alpha}$ and $\mathbf{\beta}$ are vectors of α_m and β_d respectively, and $\mathrm{Gamma}(\alpha_m\,|\,a,\,b)$ is a Gamma distribution with parameters a and b.

The framework above forms the objective function, which is the probability of all parameters, $\mathbf{W}$, $\mathbf{\alpha}$, $\mathbf{\beta}$, given the observations $\mathbf{Y}$:

$$p(\mathbf{W},\,\mathbf{\alpha},\,\mathbf{\beta}\,|\,\mathbf{Y}) = p(\mathbf{W}\,|\,\mathbf{Y},\,\mathbf{\alpha},\,\mathbf{\beta})p(\mathbf{\alpha},\,\mathbf{\beta}\,|\,\mathbf{Y}) \qquad (4.7)$$

where $\mathbf{Y}$ is a matrix, the columns of which are $\mathbf{y}^n$ for n from 1 to N.

In the first item in Equation (4.7), it is straightforward to infer that $\mathbf{w}_d$ is independent from any other weight vectors given $\mathbf{Y}$, $\mathbf{\alpha}$ and $\mathbf{\beta}$. Consequently:

$$p(\mathbf{W} \,|\, \mathbf{Y},\ \boldsymbol{\alpha},\ \boldsymbol{\beta}) = \prod_d \frac{p(\mathbf{y}_d \,|\, \mathbf{w}_d,\ \beta_d)\, p(\mathbf{w}_d \,|\, \boldsymbol{\alpha})}{p(\mathbf{y}_d \,|\, \boldsymbol{\alpha},\ \beta_d)} = \prod_d \mathcal{N}(\mu_{\mathbf{w}_d},\ \Sigma_{\mathbf{w}_d}) \qquad (4.8)$$

where the covariance matrix and mean for the posterior distribution of $\mathbf{w}_d$ are:

$$\Sigma_{\mathbf{w}_d} = \left(\beta_d \Phi(\mathbf{X})^{\mathrm{T}} \Phi(\mathbf{X}) + \Psi^{-1}\right)^{-1} \text{ and } \mu_{\mathbf{w}_d} = \beta_d \Sigma_{\mathbf{w}_d} \Phi(\mathbf{X})^{\mathrm{T}} \mathbf{y}_d \qquad (4.9)$$

where $\Phi(\mathbf{X})$ is a matrix: $\Phi(\mathbf{X}) = \left(\phi(\mathbf{x}^1),\ ...,\ \phi(\mathbf{x}^N)\right)^{\mathrm{T}}$.

The latter item in equation (7) can be calculated as: $p(\boldsymbol{\alpha},\ \boldsymbol{\beta}\,|\,\mathbf{Y}) \propto p(\mathbf{Y}\,|\,\boldsymbol{\alpha},\ \boldsymbol{\beta})\,p(\boldsymbol{\alpha})\,p(\boldsymbol{\beta})$, where

$$\begin{aligned} p(\mathbf{Y}\,|\,\boldsymbol{\alpha},\ \boldsymbol{\beta}) &= \int p(\mathbf{Y}\,|\,\mathbf{W},\ \boldsymbol{\beta})\,p(\mathbf{W}\,|\,\boldsymbol{\alpha})\,d\mathbf{W} \\ &= \int_{\mathbf{w}_1} \cdots \int_{\mathbf{w}_D} \prod_d p(\mathbf{y}_d\,|\,\mathbf{w}_d,\ \beta_d)\,p(\mathbf{w}_d\,|\,\boldsymbol{\alpha})\,d\mathbf{w}_1 ... d\mathbf{w}_D \\ &= \prod_d \mathcal{N}(0,\ \beta_d^{-1} I + \Phi(\mathbf{X})\Psi\Phi(\mathbf{X})^{\mathrm{T}}) \end{aligned} \qquad (4.10)$$

With no prior knowledge on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, these two parameters are assumed to be "uniformly" distributed so $p(\boldsymbol{\alpha})$ and $p(\boldsymbol{\beta})$ have little impact on $p(\boldsymbol{\alpha},\ \boldsymbol{\beta}\,|\,\mathbf{Y})$. Using the similar approximation as that of RVM (Tipping, 2000), the hyper-parameters are optimized by setting the derivative of Equation (4.10) to be zero:

$$\alpha_m = \frac{D}{\sum_d \left((\mu_{\mathbf{w}_d}^m)^2 + \Sigma_{\mathbf{w}_d}^{mm}\right)} \text{ and } \beta_d = \frac{N - \sum_m (1 - \alpha_m \Sigma_{\mathbf{w}_d}^{mm})}{\left(\mathbf{y}_d - \Phi(\mathbf{X})\mu_{\mathbf{w}_d}\right)^{\mathrm{T}}\left(\mathbf{y}_d - \Phi(\mathbf{X})\mu_{\mathbf{w}_d}\right)} \qquad (4.11)$$

where $\mu_{\mathbf{w}_d}^m$ is the mth element of the vector $\mu_{\mathbf{w}_d}$, and $\Sigma_{\mathbf{w}_d}^{mm}$ is the diagonal element at mth row and column in $\Sigma_{\mathbf{w}_d}$.

The parameters are inferred by iterating between Equation (4.9) and (4.11) until convergence. The radial basis centres θ_m are initialized to contain all $\mathbf{x}^n$. As with an RVM, many of α_m become infinite during the training process, so the corresponding radial bases are removed accordingly. The radial basis parameter η_m is optimized by gradient descent.

Given the inferred parameters, the distribution of predicted values $\mathbf{y}_d'$ given a test example $\mathbf{x}'$ is computed as:

$$\begin{aligned} p(\mathbf{y}_d' \,|\, \mathbf{y}_d,\ \boldsymbol{\alpha},\ \boldsymbol{\beta}) &\\ = \int p(\mathbf{y}_d'\,|\,\mathbf{w}_d,\ \boldsymbol{\beta})\,p(\mathbf{w}_d\,|\,\mathbf{y}_d,\ \boldsymbol{\alpha},\ \boldsymbol{\beta})\,d\mathbf{w}_d &\\ = \mathcal{N}(\mu_{\mathbf{w}_d}^{\mathrm{T}}\phi(\mathbf{x}'),\ \beta_d^{-1} + \phi(\mathbf{x}')^{\mathrm{T}}\Sigma_{\mathbf{w}_d}\phi(\mathbf{x}')) & \end{aligned} \qquad (4.12)$$

where the prediction is made by the mean of the distribution: $\mu_{\mathbf{w}_d}^{\mathrm{T}}\phi(\mathbf{x}')$.

Therefore, the structure-function relationship in Equation (4.2) is implemented by

$$\mu_{\mathbf{w}_d}^{\mathrm{T}} \frac{\partial \phi(\mathbf{x}^n)}{\partial \mathbf{x}^n} \qquad (4.13)$$

where $\dfrac{\partial \phi(\mathbf{x}^n)}{\partial \mathbf{x}^n} = \left(\dfrac{\partial \phi_1(\mathbf{x}^n)}{\partial \mathbf{x}^n}, ..., \dfrac{\partial \phi_M(\mathbf{x}^n)}{\partial \mathbf{x}^n}\right)^{\mathrm{T}}$, and $\dfrac{\partial \phi_m(\mathbf{x}^n)}{\partial \mathbf{x}^n}$ is a vector: $-2\eta_m \phi_m(\mathbf{x}^n)(\mathbf{x}^n - \theta_m)$ according to Equation (4.4).

The RBF attempts to model the relationship between $\mathbf{y}$ and $\mathbf{x}$ without the limiting assumptions associated with the classic linear model described above. As an illustrative example, consider one $\mathbf{x}$ value, say $\mathbf{x}_1$, appearing to be co-related to a $\mathbf{y}$ value, say $\mathbf{y}_1$. This apparent relationship might be explained by $\mathbf{x}_1$ being very strongly related to $\mathbf{x}_2$ which in turn is very strongly associated with $\mathbf{y}_1$. Thus the apparent $\mathbf{x}_1$ and $\mathbf{y}_1$ relationship may well be much weaker or not even be significant. As it will be shown in the derived structure-function relationship (Figure 4-3 to Figure 4-5), this covariance in the relationship between the $\mathbf{y}$ and $\mathbf{x}$ values is modelled better with the RBF approach. Furthermore, the central idea of the RBF are the basis functions, each of which performs very much like a dynamic window or kernel that moves across the data, both spatially and at various stages in disease severity, identifying groups of measurements that appear to behave in a similar pattern. The non-normalized Gaussian basis function used in this study has an activation field which has a centre: that is, a particular input value at which they have a maximal output. The output tails off as the input moves away from this point. In this way, those hidden basis functions that have centres similar to the input $\mathbf{x}$ patterns will have stronger activation and will thus contribute more to the prediction of $\mathbf{y}$. On the other hand, those basis functions with weak activation will be 'isolated' and will not affect the prediction. Moreover, RBF 'learns' the parameters from the data and this is customised for our purpose by using a Bayesian framework (BRBF). The method makes predictions in multiple dimensions by extending the standard relevance vector machine (Tipping, 2000). The VF sensitivities at different locations are implicitly correlated by sharing the same basis functions and parameters of weights $\boldsymbol{\alpha}$. In the BRBF model, the original visual field sensitivity value in dB was used.

In general, BRBF is not restricted by the number of inputs in the model, or more precisely by the dimensionality of the data, so that it can be adapted to use the RNFLT values at each pixel, rather than only the 64-sector peripapillary thickness profile. In this study, examples of the SLP images used for predicting VF are shown in Figure 4-2. We used an annular region, centred on the ONH. The inner and outer diameters were 2.3mm and 4.9mm, respectively, so that the annulus was 26 pixels wide (compared with the 8-pixel calculation ring in the GDx software from which the 64-sector RNFLT peripapillary profile is computed). In this annulus, there are 16512 pixels each with a retardation value. It was hypothesised that the predicted spatial structure-function relationships would be strengthened by avoiding the data reduction to 64 sector RNFLT values. This hypothesis will be validated by the improved structure-function relationship derived from the BRBF by using the measurement in its higher dimensional form.

Another challenge in the modelling process involves handling the large dimensionality of the SLP data. If the dimensionality M and the number of data points N of a dataset satisfy $M \geq N$, the dimensionality of the dataset can be reduced from M to $N-1$ with minimal loss of information using Principal Component Analysis (PCA) (Jolliffe, 1986), because the N data points span at least one linear *hyperplane* in this M-dimensional space. Using this technique the 16512 dimensional SLP image vectors were reduced to 228 dimensions for analysis and transformed back to the original SLP vectors for the purpose of visualization and evaluation of results. Using the classic linear model on a reduced SLP image vector with 228 elements and a 24-2 VF with 52 sensitivity values will still result in a prohibitive number of 'weights' (11856) to be fitted, which will cause significant overfitting to the 'noise' in the data.

4.5. Testing the model

It is well established that if validation of a modelled relationship between dependent and independent variables is performed on the same dataset as the selection of input variables and weights, then the model estimates will be overly optimistic (Copas, 1983, Altman and Royston, 2000). It is also well known that in developing a model, the input variables and weights selected may vary across different samples (Miller, 1984). Therefore, the models were developed on the MEH and REH data alone, leaving the BMES data as a test dataset.

The predictive performance of the classic linear model and the BRBF was evaluated by point-by-point analysis of the predictions of the VF sensitivity in the 306 VFs from the BMES dataset. The predictive performance was summarised by the mean of the absolute prediction errors in the 52 points of the VF.

Matlab code was written to display the results and present the output of the linear and BRBF models via a graphical user interface. One graphical output was a structure-function map in a

similar format to that described by Gardiner et al (Gardiner et al., 2005). The structure-function relationship was defined by the corresponding functional change given a subtle structural change, which is mathematically modelled by the derivative of the BRBF in Equation (4.13). This describes the relationship between **y** (VF sensitivity) and **x** (RNFLT) at each individual VF location. The other output from this analysis consisted of point-by-point predictions of each subject's VF as represented by the HFA greyscale (which was replicated for this purpose). These outputs were considered for 1) the classic linear model, 2) the BRBF I model (based on the 64 summary RNFLT values output from the GDx software) and 3) for the BRBF II model (based on the reduced 228 dimensional data derived from the PCA on the 16512 individual pixel retardation values in the broad annulus centred on the ONH).

4.6. Results

A summary of the measurements: HFA mean deviation (MD), HFA pattern standard deviation (PSD); GDx Nerve Fibre Indicator (NFI)) for each of the datasets is given in Table 4-1.

The mean absolute prediction error of VF sensitivities in the 306 eyes from the BMES data was 4.9dB (standard deviation (SD) 4.0dB) for the classic linear model. In comparison, both BRBF models yielded a nearly two-fold improvement (P<0.001; paired t-test) in performance: BRBF I 2.9dB (SD 3.7dB), BRBF II 2.8dB (SD 3.8dB). In BRBF I and BRBF II models, the training process selected 49 and 73 basis functions respectively in the hidden layer.

	Subjects	Gender	Age mean ± SD (min, max)	MD (dB) mean ± SD (min, max)	PSD (dB) mean ± SD (min, max)	NFI mean ± SD (min, max)
MEH	Healthy	19M, 11F	40.6 ± 15.6 (21, 75)	1.26 ± 0.76 (0.09, 3.30)	1.47 ± 0.28 (1.12, 2.27)	16 ± 8 (2, 30)
	OHT	13M, 21F	60.7 ± 11.1 (21, 75)	1.07 ± 0.82 (0.10, 3.06)	1.35 ± 0.27 (1, 2.1)	20 ± 10 (5, 43)
	Glaucoma	28M, 15F	60.8 ± 13.1 (31, 84)	-4.02 ± 2.55 (-12.00, 1.15)	4.93 ± 2.93 (1.49, 12.50)	42 ± 18 (11, 80)
REH	Healthy	23M, 23F	60.4 ± 12.1 (23, 77)	0.38 ± 0.99 (-1.55, 2.73)	1.63 ± 0.26 (1.13, 2.30)	21 ± 9 (2, 43)
	Glaucoma	47M, 29F	62.2 ± 10.1 (30, 82)	-9.52 ± 8.43 (-30.39, 1.25)	8.35 ± 4.32 (1.99, 15.92)	63 ± 21 (21, 98)
BMES	Healthy	99M, 131F	69.3 ± 6.5 (60, 87)	0.21 ± 1.07 (-1.26, 3.03)	1.53 ± 0.29 (1.12, 2.55)	19 ± 10 (2, 49)
	Glaucoma	25M, 51F	72.0 ± 6.3 (61, 78)	-7.94 ± 6.55 (-29.67, 1.56)	6.97 ± 3.76 (1.67, 15.56)	65 ± 24 (27, 98)

Table 4-1: A summary of the measurements (HFA mean deviation (MD) and pattern standard deviation (PSD); GDx Nerve Fibre Indicator (NFI)) for the three datasets (MEH, REH, BMES) used in this study.

Figure 4-1 summarises the predictive performance of the classic linear model and the BRBF I model across the range of VF sensitivity measurements. The BRBF II demonstrated similar

prediction performance as BRBF I. Published test-retest differences (5th and 95th percentile limits) (Artes et al., 2002) across all locations from two VF tests from each of 49 individuals have been superimposed on Figure 4-1b. On inspection, these limits are similar to the 90% prediction limits when using the BRBF to predict the VF from a GDx RNFLT measurement of the same individual. Note that predictions at higher sensitivities (>30dB) tend to be slightly lower than the actual values whilst at lower sensitivities (<20dB) the predictions tend to be higher.
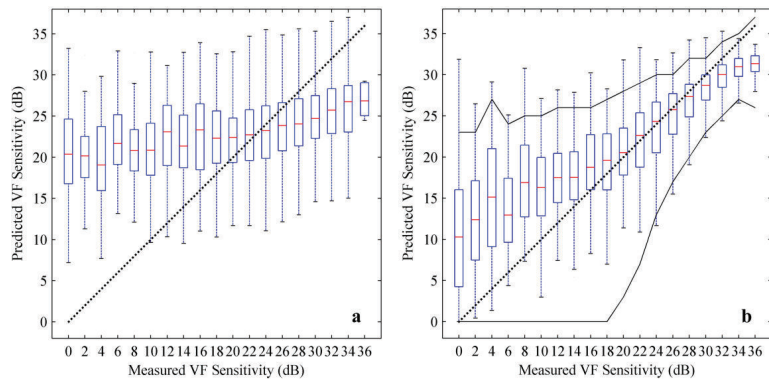
Figure 4-1. Distributions of the error between the predicted and the measured sensitivity for each visual field (VF) location in 306 eyes from the Blue Mountains Eye Study data, stratified by VF sensitivity. Each error bar summarises the predictive performance over a 2dB range from 0 to >36dB. Thin vertical lines indicate 90% prediction limits (5th and 95th percentile of error), the 'boxes' indicate the interquartile range of the prediction error (25th and 75th percentile error) with the line in the box indicates the median error. The dotted line of unity indicates perfect prediction (no error). (a) illustrates the predictive performance of the classic linear model. (b) illustrates the predictive performance of the Bayesian Radial Basis Function I model. Previously published (5th and 95th percentiles) test-retest limits for VF data derived from the pointwise differences between two VFs tested over a short period of time have been superimposed as black lines on (b).

Figure 4-2 gives some case examples of the predictions. In some cases, the classic linear model overestimates the defect severity of the VF (I and III) and in other cases predicts a less damaged VF (IV), when compared to the true measured VF. In Figure 4-2II the classic linear model matches the overall average sensitivity of the VF but fails to capture the spatial location of this loss. In each case the BRBF better estimates the true VF, with spatial features of the measured defects generally retained. In Figure 4-2IV the BRBF model manages to predict the advanced defect severity.
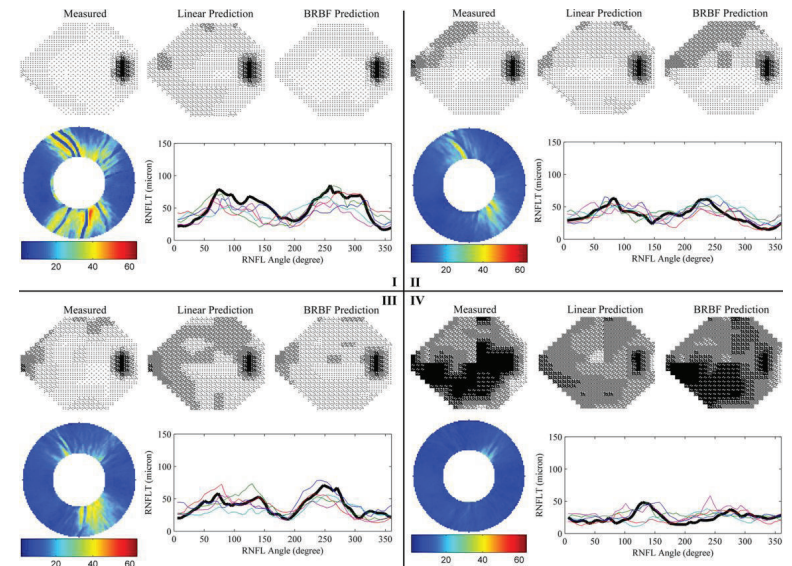
Figure 4-2. Model predictions for four cases from the Blue Mountains Eye Study dataset. For each case (I to IV), the top row shows, from left to right, Humphrey Field Analyzer visual field (VF) grayscales for the measured VF, the VFs predicted from the classic linear regression and the Bayesian Radial Basis Function (BRBF) with 64-sector retinal nerve fibre layer thickness (BRBF I), respectively. The predicted VF from BRBF II with 16512-pixel annulus from the GDxVCC scanning laser polarimetry (SLP) images was in general similar to the one from BRBF I. The row of graphics below show the GDxVCC SLP image annulus and the corresponding 64-sector retinal nerve fibre layer thickness (RNFLT; thick black line) used to predict the VFs. The coloured lines in the 64-sector RNFLT plot indicate the top five radial basis functions with the strongest activation (those contributing the most to the prediction).

Figure 4-3, Figure 4-4 and Figure 4-5 provide topographical maps of the spatial relationship between RNFLT positions on the peripapillary annulus and VF sensitivity at all the points in the VF, by means of classic linear regression, BRBF I and BRBF II, respectively.

The agreement between the derived structure-function relationship and the anatomical benchmark was also summarised by calculating the absolute angular difference between the direction with the strongest derived relationship and the anatomical benchmark. This was calculated as the median of the absolute angular difference across all locations in a spatial relationship map such as Figure 4-3, Figure 4-4 and Figure 4-5.

Both visual inspection on the spatial relationship and the quantified agreement with anatomical prior knowledge showed that the structure-function relationship derived by the classic linear regression model (Figure 4-3) has little concordance (median absolute angular difference of 62°) with the anatomical benchmark derived in Garway-Heath et al (2000) (Garway-Heath et al., 2000). This correspondence improves (P<0.001; paired t-test) with the BRBF I model (median absolute angular difference of 15°) in general, and improves further (P<0.001; paired t-test) still with the BRBF II model (median absolute angular difference of 12°), especially at the points around central vision and blind spot. This is consistent with the BRBF models learning to 'encode' the structure-function relationship during the training.

Figure 4-3. A topographical map describing the relationship between the 64-sector retinal nerve fibre layer thickness (RNFLT) profile and individual visual field (VF) locations as described by classic linear regression. The whole figure forms a Humphrey Field Analyzer 24-2 VF for a right eye. Each location is represented by a circular graph made up of 64 bars representing the correlation value between the VF location and the RNFLT at the corresponding angle. In this instance, the 'correlation values' are derived from scaled values of the weights given in Equation (4.1). The length of the bar indicates the magnitude of the correlation, with red bars showing a positive correlation and blue bars indicating a negative correlation. The green bar with an asterisk indicates the location of expected strongest correlation on the basis of the anatomically derived map from Garway-Heath et al (2000).

Figure 4-4. A topographical map describing the relationship between the retinal nerve fibre layer thickness (RNFLT) profile and individual visual field (VF) points as described by Bayesian Radial Basis Function model I using the 64-sector RNFLT profile. The composition of the graph is the same as Figure 4-3 with the green bar with an asterisk indicates the location of expected strongest correlation on the basis of the anatomically derived map from Garway-Heath et al (2000). In this instance the 'correlation values' are derived from scaled values in Equation (4.13).

Figure 4-5. A topographical map describing the relationship between the retinal nerve fibre layer thickness (RNFLT) profile and individual visual field (VF) points as described by Bayesian Radial Basis Function model II using the 16512 pixel retardation values from the GDxVCC scanning laser polarimetry image. The composition of the graph is the same as Figure 4-3 and Figure 4-4 with the green bar with an asterisk indicates the location of expected strongest correlation on the basis of the anatomically derived map from Garway-Heath et al (2000). In this instance the 'correlation values' are derived from scaled values in Equation (4.13) and are summarized into 64 sectors for the purpose of comparison.

4.7. Discussion

The application of ANNs to both functional and structural measurements in glaucoma is not a new idea (Bowd et al., 2002, Goldbaum et al., 1994, Bengtsson et al., 2005, Brigatti et al., 1996, Uchida et al., 1996, Brigatti et al., 1997, Spenceley et al., 2007). Most of these applications have tended to use a 'conventional' MLP ANN. Technically there are several advantages of the RBF over MLP ANNs. For example, with the latter there is an input and the distributed pattern 'lights up' all hidden units to contribute to the prediction of the output, which makes them combine and interfere with each other. This typically yields a highly non-linear training process with mathematical difficulties that result in a slow convergence of a training procedure (Bishop, 1996). Moreover, the complexity of hidden unit patterns causes difficulty when interpreting the

result because the connection among units and the hidden unit output do not have any physical or realistic meaning: they are simply some numbers that can produce a correct output. Hence, MLP ANNs are less suitable for the mapping of points in different measurement spaces, which requires a detailed understanding of the hidden layer output and other manipulation (e.g. derivatives) within the network. In contrast, RBF substitutes hidden layers in MLP with a set of basis functions, which leaves the solution of weights in a linear space. The Gaussian basis function forms a local representation in the hidden unit, each of which can be understood as a representative of similar input patterns (Figure 4-2). With a given RNFLT input, only a few 'representatives' will be activated and contribute to the VF prediction (Figure 4-2). Furthermore, the RBF handled within a Bayesian probabilistic framework (Tipping, 2000), as developed for our purposes, is unlike most other ANN approaches because it is independent of any subjective input parameter, and thus requires no model validation on a test dataset provided that the training dataset is representative and sufficiently large to enable modelling of the many different states of the VF. However, for this study the 'trained' model is still tested on a separate, independent dataset to illustrate the generalisation of the model performance. One 'technical' limitation of the current BRBF model, despite its good performance, is that it assumes that the variability in the VF measurements is largely Gaussian, which is not optimal given that it is often skewed and heavily-tailed (Artes et al., 2002).

VF locations have previously been related to sectors of the ONH using an anatomical map (Garway-Heath et al., 2000) derived by overlaying an appropriately scaled VF grid on RNFL photographs and tracing nerve fibre bundles or defects from various VF locations to the ONH margin. Gardiner and colleagues (Gardiner et al., 2005) produced a topographical map of the relationship between sectors of the ONH and locations in the VF by considering the linear correlation between Heidelberg retina tomography neuroretinal rim area and VF sensitivity at each point in the VF. The classic linear model is akin to that derived by Gardiner et al except that the SLP measurements from GDx imaging are used. The maps derived from the BRBF (Figure 4-4 and Figure 4-5) indicate a closer concordance with the anatomically derived landmark than the map derived from the classic linear model (Figure 4-3). In addition, the predictions on pointwise VF sensitivities in the validation dataset were, on average, better for the BRBF method than the classic linear model. A plausible explanation for the improvements afforded by the BRBF technique is that it models the spatial and quantitative structure-function relationship more precisely. The technical and statistical advantages of the BRBF model over the classic linear model would support this notion. For example, the BRBF considers that the VF points, and indeed RNFLT values acquired from different discreet areas, interact as groups rather than as independent measurements. The BRBF also accommodates the covariance and non-independence of the measurements; it is less affected by outlying observations and makes

no assumption about the linearity of relationships. Of course, one difficulty in generating any type of map, driven by data or anatomical observation, is the restriction in the sampling of VF points: these are, for example, probably not optimally placed for estimating RGC density nearer the fovea.

The range and distribution of differences between the measured VF sensitivity values and those predicted from the RNFLT by the BRBF, at different levels of sensitivity, is shown in Figure 4-1b. This profile is remarkably similar to published limits for test retest variability when two VFs are measured within a short space of time (Artes et al., 2002). This suggests that, on average, a VF predicted by the BRBF from RNFLT values has measurement noise equivalent to that found in a newly measured field. This isn't as exciting as it may first appear because it is well established that the measurement noise in VFs is already very high, prohibiting straightforward clinical diagnosis of glaucomatous defects and monitoring progression. Nevertheless, this finding illustrates that the range and scale of the average predictive performance of the BRBF model is much better than the classic linear model approach, which completely fails to predict the full range of VF values (Figure 4-1a).

Although the mean absolute prediction error is reduced with BRBF I and BRBF II, the standard deviation of the absolute prediction error of both BRBF models are still large (3.7dB and 3.8dB). This can also be observed in more detail in the prediction limits shown in Figure 4-1b. Despite of the general similarity between the prediction limits and VF test-retest limits, predictions at the normal end of the range still tend to be lower than the actual VF measurements, and at the damaged end, predictions tend to be higher than actual VF measurements. The reported retest differences (Artes et al., 2002), on average, are very small at the normal or 'healthy' end. At the damaged end, the retest values tend to be higher, but the median BRBF predictions are a little higher than the retest values. This likely reflects the difficulty the prediction from RNFLT images have in identifying small, focal defects. Moreover, the 'floor effect' in the VFs and GDxVCC SLP images (Wyatt et al., 2007, Blumenthal et al., 2006) and the atypical scan pattern in GDxVCC SLP images, which may be associated with glaucoma severity (Yanagisawa et al., 2009), may be additional causes of the overestimation at the lower end of the VF sensitivity. Furthermore, because the diagnosis of subjects in REH datasets includes structural criteria, the normal subjects in the training dataset may have 'super-normal' structure and the glaucomatous subjects have greater than average structural damage. This potential bias on subject selection might distort the structure-function relationship. Therefore, the training process may be improved by including a range of glaucoma severity in subjects defined only by VF loss. Another factor that may confound the reported prediction accuracy is that the models were trained with VFs tested with both SITA and full-threshold programs, while they were evaluated

on VFs tested only with full-threshold program. This may, in part, account for the observed tendency for the prediction to overestimate the sensitivities in the test dataset due to the higher sensitivities obtained with SITA compared with the full-threshold program (Sharma et al., 2000). However, this effect would be small because the sensitivity difference with two programs is just 1.3dB on average (Sharma et al., 2000).

Recent investigations attempting to uncover the structure-function relationship in glaucoma generally have the aim of assessing the relative accuracy of structural and functional tests throughout the course of the disease (Hood et al., 2007, Hood and Kardon, 2007). For example, loss of function without loss in structure that does not adhere to a particular structure-function model's prediction might be an indicator of a non-glaucomatous process, measurement imprecision, or some artefact in the image of the structure or function test. The long term aim is to use the BRBF technique to provide a relevant clinical tool that indicates concordance between the VF and the chosen surrogate measure for structural loss. For example, when a VF and structural measure derived from one of GDx, HRT or OCT are available, a chart mapped in VF space will be provided indicating areas where the measurements are in concordance (within a certain range of accuracy and precision) and where they are not: this could provide clinically useful information about the reliability of the individual measurements or diagnostically useful information.

It is an imperative that any new statistical method should be developed and tested on more than one dataset (Altman and Royston, 2000). The study had access to three large independent datasets, each collected at one of three clinical centres. The inclusion criteria for glaucomatous and healthy subjects were generally consistent across the three samples. However, as the aim of this study was not to determine diagnostic performance, the precise definitions for glaucoma were less important. In fact, the mixture of data can be viewed as an advantage in the study design. However, further testing on different datasets, especially where realistic estimates of measurement precision have been performed (from test-retest measurement), is still required and this is being undertaken in our laboratory and elsewhere.

The BRBF method is not limited to one type of input of structural measurement or imaging device. The BRBF was shown to handle input of the GDx RNFLT peripapillary profile (64 values) as well as the PCA-reduced RNFLT values at 16512 pixels in a wide peripapillary annulus. The BRBF method could be used on neuroretinal rim area values from scanning laser opthalmoscope technology or RNFLT values derived from OCT technology, or any other surrogate measure of glaucomatous structural loss. Moreover, the study demonstrated, albeit qualitatively with the maps in Figure 4-5, that using the surrogate measures of RNFLT in their high dimensional form provides a closer mapping to the expected structure-function relationship.

It is speculated that the next generation of SD-OCT instrumentation, now finding its way into clinics and providing volumes of data for RNFLT, will be particularly amenable to this method.

In conclusion, this study introduced a new statistical method for describing the relationship between functional and structural measurements used in the clinical evaluation of glaucoma. Evidence from a dataset independent of those used to derive the model indicates that the BRBF method has advantages over standard statistical approaches for modelling these relationships, and estimates of functional deficits from structural measures yielded from this method are better than those derived from a classic linear regression approach. This method can provide a platform from which clinically useful tools can be derived for mapping and charting discordance between VF and RNFLT measurements in glaucoma: this is the topic of the next chapter.

Chapter 5 : Quantifying discordance between structure and function measurements in the clinical assessment of glaucoma

5.1 Summary

The Bayesian radial basis function (BRBF), reported in last chapter, was shown to be capable of learning the *knowledge* about the structure-function relationship from the measurements of visual field (VF) and retinal nerve fibre layer thickness (RNFLT). In this chapter, the BRBF approach is used to develop a clinically useful methodology to quantify and visualize structure-function discordance between the retinal structure and visual function.

Five GDxVCC RNFLT scans and 5 Humphrey SITA VF tests were obtained from 50 glaucomatous eyes from 50 patients. A best available estimate of the 'true' VF was calculated as the point-wise median of these 5 replications. This 'true' VF was compared with every single RNFLT-predicted VF from BRBF and every single measured VF. Predictability of VFs from RNFLT was established from previous data. A structure-function pattern discordance map (PDM) and structure-function discordance index (SFDI; values 0 to 1) were established from the predictability limits for each structure-function measurement pair to quantify and visualise the discordance between the structure-predicted and measured VFs. The structure-function PDM highlights locations and the extent of structure-function discordance and the SFDI provides a summary index quantifying overall discordance. These tools may help clinicians trust the mutually confirmatory structure-function measurements with good concordance, or identify unreliable ones with poor concordance.

The work reported in this chapter has formed a paper published by *Archives of Ophthalmology* ((Zhu et al., 2011a); See List of supporting publications). The joint authors of this work are Haogang Zhu [HZ], David Crabb [DC], Marie-Josée Fredette (MJF), Douglas Anderson [DA] and David Garway-Heath [DGH]. The data was made available from previous studies conducted by MJF and DA. The work was supervised and edited by DC and DGH. Everything else described in this chapter was completed by HZ. The results in this chapter have also in part been presented as a read paper at the Association for Research in Vision and Ophthalmology Meeting, Fort Lauderdale, Florida, USA in May 2009, and at the 18th International Symposium of the Imaging and Perimetry Society Meeting in Nara, Japan in May 2008 (See List of supporting publications).

5.2. Introduction and previous studies

As previously discussed, glaucoma is a progressive optic neuropathy in which structural damage is evident at the optic nerve head (ONH) and the retinal nerve fibre layer (RNFL). This structural damage results in the loss of visual function during the progression of the disease. Glaucoma is clinically evaluated by various tests, which, in general, examine structural or functional properties of the optic disc or retina. Ideally, the functional loss would be precisely predictable by the magnitude and configuration of structural loss. However, the relationship between the structural and functional measurements is still not well understood, mainly due to the poor precision and accuracy of the current clinical devices measuring the structural and functional deficits. Particularly, various structural measures use optical imaging techniques to examine ONH parameters or RNFL thickness (RNFLT) as surrogates for the biological variable of interest, namely the number of (functioning) retinal ganglion cells. Visual function can be examined with visual field (VF) tests of various kinds, some of which are similar enough to be grouped under the term 'standard automated perimetry' (SAP). SAP is the clinical cornerstone in the assessment of glaucoma but is also subject to considerable measurement variability and inaccuracy (Anderson, 2006). Despite their limitations, these techniques are presently the state of art tools for the diagnosis and management of glaucoma.

When multiple types of measurement assessing glaucoma damage are available for the same patient, an important question is: "Are they consistent with each other, and if not, to what degree and why are they discordant?" Thinking about this question would help clinicians to avoid being misled by inaccurate or imprecise measurements. Recently, clinical software has been developed that presents structure and function classification analysis (probability levels) for corresponding regions of the VF and ONH (Heidelberg Eye Explorer (HEYEX) version 1.7.0, Heidelberg Engineering, Heidelberg, Germany). This enables visualization of structure-function agreement, or spatial concordance, at an ONH sector level in both the structural and functional measurements, but the degree of concordance is not quantified.

Attempts have been made to correlate regions from the structural measures (e.g., mean RNFLT in 6 sectors) and groups of, or individual points in the VF, and assess the curve linear (e.g., log-linear) or monotonic association between the two variables via R2, Pearson or Spearman coefficients (Brigatti and Caprioli, 1995, Weinreb et al., 1995, Iester et al., 1997, Teesalu et al., 1997, Garway-Heath et al., 2002, Garway-Heath et al., 1999, Anton et al., 1998, Gardiner et al., 2005, Mai et al., 2007, Bowd et al., 2006, Schlottmann et al., 2004). As methods of examining and demonstrating regional structure-function relationships, these approaches make restricted assumptions and, by their nature, do not quantify the frequency or magnitude of structure-function agreement (concordance) or disagreement (discordance) at a high spatial resolution (for

example, at each VF test location). Firstly, the analyses suffer from considerable data reduction by summarizing the structural image into a few discrete measurements (for instance, the sectoral RNFLT derived from the GDxVCC [Carl Zeiss Meditec, Dublin, CA, USA]), or a grouping of visual field locations, so the spatial resolution of the measurements is lost. Secondly, these analyses assume, and are limited to, a particular 'shape' of the structure-function relationship, while in reality this relationship may be more complex with the nature of the association possibly changing across the stages of the disease. Thirdly, these studies treat individual points from structure and function measurements as independent and fail to address the inherent spatial relationships within the individual measurements. Last but not least, owing to measurement variability and anatomic variation, the structure-function relationship in these studies is reported as a statistical association based on measurements from a population as average typical relationship, which only approximates the relationship of any particular individual in the population. However, the structure-function discordance should be quantified for the measurements of individual patients and the method should be generalized to the individuals not included in the dataset used to establish the structure-function relationship.

In contrast to association-based methods, a more flexible non-linear prediction model, the Bayesian Radial Basis Function (BRBF), has been developed and validated to predict the VF from the RNFLT (Zhu et al., 2010b, Zhu et al., 2009). The model was built and tested on the measurements of 535 eyes from 535 subjects from three independent centres. It was shown that the BRBF model, on average, can predict a patient's visual field from the RNFLT almost as well as another visual field from the patient. Moreover, the BRBF derived a high resolution structure-function relationship that is consistent with known typical location of retinal ganglion cells that project their axons toward particular meridians of the optic disc (Garway-Heath et al., 2000). The VF predicted from the RNFLT by the BRBF can be understood as the expectation or representation of retinal structure in the domain of the VF. Therefore, structure-function discordance is converted to discordance between the structure-expected VF and the measured VF.

5.3. Subjects and datasets

A test-retest dataset was used in this study. Fifty patients were recruited from a hospital-based glaucoma ophthalmology practice (Bascom Palmer Eye Institute, Anne Bates Leach Eye Hospital, University of Miami Miller School of Medicine, Miami, Florida). Glaucoma was diagnosed by detailed medical and ocular histories and by standard ophthalmic examination. The ophthalmic examination included best-corrected visual acuity, anterior segment examination, Goldmann applanation tonometry, dilated fundus examination and SAP with the 24-2 Swedish Interactive Thresholding Algorithm (SITA)-Standard program of the Humphrey

Field Analyzer (HFA, Carl Zeiss Meditec Inc, Dublin, CA, USA). The dataset covers a wide spectrum of glaucoma severity from early to advanced stage, judged by their treating physicians to be under control with stability of nerve damage and VF loss. Inclusion criteria for these patients included: (1) 18 years old or more, (2) controlled intraocular pressures, (3) best corrected visual acuity of 20/40 or better, (4) less than 5 diopters of spherical, and 3 diopters of cylindrical, refractive error, (5) pupil diameter of 2 mm or more, (6) no history of ocular or neurological disease or surgery that might produce non-glaucomatous structural or functional abnormality, (7) no history of amblyopia, (8) mental and physical incapacity to perform the tests, and (9) willingness to participate in the study. Only one eye had to fit the eligibility criteria for the patient to enter the study. If both eyes met the eligibility criteria, the study eye was selected to ensure representation of a wide spectrum of the disease.

Each subject attended five sessions within an eight-week period. During each of the five sessions, the subjects had SAP, scanning laser polarimetry (SLP) and optical coherence tomography (OCT) measurements done. SAP was performed with the 24-2 SITA-Standard program of the HFA with the Goldmann size III stimulus. The SAP test used to establish the diagnosis and eligibility was not included in the 5 repeat tests obtained for the study experiments. The SLP measurements were taken with the GDxVCC. All scans were acquired through undilated pupils. OCT images were taken with the Stratus OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA) using the 3.4mm fast scan protocol.

The study followed the tenets of the Declaration of Helsinki, and the protocol was approved by the Human Subjects Institutional Review Board of the University of Miami. Patients agreed to participate as subjects in the study and signed an informed consent form after explanation of the nature and possible consequences of the study. The data were collected in the last half of 2004. They were later anonymised and transferred to a secure database held at City University London for the conduct of the reported analysis.

5.4. Methods

5.4.1. Structure-predicted VF

The structure-function relationship was previously modelled with a BRBF which was used to predict the VF from the RNFLT profile from the GDxVCC. The model had been trained on RNFLT and VF measurements of 229 eyes, and derived a structure-function relationship that is consistent with the retinal anatomy such that the structure-predicted VF can be understood as the RNFLT representation in the VF domain (Zhu et al., 2010b, Zhu et al., 2009). The structure-

function discordance, as described below, is defined as the 'difference' between this structure-predicted and the measured VF.

5.4.2. Best available estimate of 'true' VF

In order to compare the predicted and measured VFs, the true VF should be known, but high variability of testing means it cannot be known from any particular VF test. Therefore, a best available estimate of the true VF was calculated as the point-wise median of the 5 repeated VFs. It is then assumed that the best available estimate VF provides closer estimate of the true VF than any single measurement. The best available estimate VF will be referred as the 'true' VF in the subsequent description. The 'true' VF was used as the reference to identify the location, extent and possible cause of any discordance when the structure-predicted VF and the measured VFs were discordant. However, the methodology proposed in this study is not constrained by the availability of a 'true' VF.

The difference between each structure-predicted VF and the 'true' VF was quantified. This difference was compared with the difference between each single VF from the 5 test-retest set and the 'true' VF.

5.4.3. VF predictability and quantification of structure-function discordance

The discordance between the structure-predicted VF and measured VF may be easily calculated as the point-wise difference of the two sensitivities. However, the predictability of the VF varies with the VF sensitivity itself. In Figure 5-1a, as reported previously (Zhu et al., 2010b, Zhu et al., 2009), the range of prediction errors decreases with a higher measured VF sensitivity. This means that a higher VF sensitivity is more predictable than a lower sensitivity; this needs to be taken into account in the calculation of the discordance. The predictability can be quantified as the percentile of the predicted VF sensitivity for any given measured VF sensitivity. This maps the VF sensitivity in dB to a prediction percentile ranging from 0 to 1. The percentile curves for the predicted VF sensitivity are illustrated in Figure 5-1b at the measured VF sensitivities of 5dB, 15dB and 25dB. The variable predictability at different measured sensitivities is illustrated by the different shapes of the percentile curves in Figure 5-1b: the corresponding curve shifts towards the right and has a steeper slope (increases from 0 to 1 more rapidly) at higher measured sensitivity.

Due to the greater predictability at high sensitivity, any given predicted/measured difference has greater importance if the measured sensitivity is high than if it is low. To compensate for this, the discordance between the structure-predicted and measured VF sensitivity is calculated as the percentile difference on the predictability curve (corresponding to the measured sensitivity)

between the structure-predicted and measured sensitivity values. For instance, for a 5dB predicted/measured difference at a measured sensitivity of 5dB, the discordance between a 0dB prediction and 5dB measurement is 5.1% (on the 5dB predictability curve; Figure 5-1b), but for a 5dB difference at a 15dB measured sensitivity, the discordance between 10dB prediction and 15dB measurement is significantly higher at 17.3% (on the 15dB predictability curve; Figure 5-1b). Similarly, the discordance between 20dB prediction and 25dB measurement is even higher: 30.1%. Therefore, by using the predictability percentiles, the difference between the structure-predicted and measured VF sensitivity in dB is normalised with respect to the predictability at any given level of measured sensitivity.
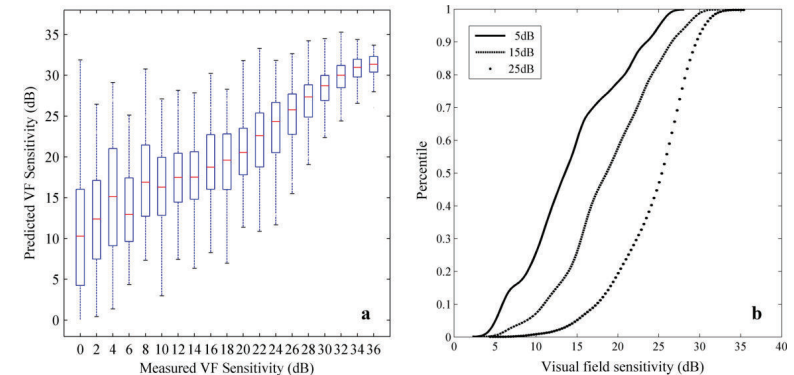
Figure 5-1. BRBF predictability and percentile curves. (a) BRBF prediction ranges (predictability) at different visual field sensitivities. Thin line tips and 'box' ends indicate 90% (5%—95%) and interquartile (25%—75%) prediction limits. (b) Percentile curves of predictability at measured sensitivities of 5dB, 15dB and 25dB.

The signed point-wise discordance between the structure-predicted and measured VF sensitivity quantifies the structure-function discordance: the sign is positive if the RNFLT overestimates the measured VF and vice versa. The value ranges from -1 to +1.

5.4.4. Pattern discordance map (PDM) and structure-function discordance index (SFDI)

The point-wise structure-function discordance can be visualised with a novel tool, named the structure-function pattern discordance map (PDM), with a technique similar to the Hinton diagram, which has been used to visualise the weight parameters in a multilayer perceptron neural network (Bishop, 1996). As shown in Figure 5-3, the structure-function PDM is a 24-2 greyscale representation of the VF with red or green squares superimposed on each location of

the measured VF. A green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result (discordance>0), while a red square predicted a greater visual defect than was found in VF testing (discordance<0). The size of the square represents the magnitude of the discordance. Like the pattern standard deviation map that is used in a HFA VF chart to describe the deviation from the normative database, the PDM provides an easily-understood clinical tool to identify the point-wise deviation of the structure-predicted VF from the measured VF.

To summarize the overall discordance, a structure-function discordance index (SFDI) was defined as the mean of absolute (unsigned) point-wise discordance across all locations in the VF. It ranges between 0 and 1, where 0 indicates no discordance and 1 means complete discordance. SFDI quantifies the average difference between the structure-predicted and measured VF, and therefore acts in a similar style as a global index from a VF chart.

The SFDI and the PDM were used on the test-retest dataset to evaluate the structure-function discordance in this glaucomatous sample. The 'true' VF was used as the 'arbiter' to judge which measurement is correct when the discordance was flagged by either of the tools. The OCT scans for the same eyes were also used to identify the possible source of discordance.

5.5. Results

The mean absolute difference (MAD) between the structure-predicted VF and the 'true' VF was 3.9dB (standard deviation (SD) 4.3dB). This compares with the MAD between a single measured VF and the 'true' VF of 2.6 (SD 3.5dB).

The SFDI was calculated for each of the 5 pairs of structural and functional measurements taken in the same session. In this glaucomatous sample, the mean SFDI was 0.34 (SD 0.11), and the distribution of SFDI is given by the histogram in Figure 5-2. Although SFDI is a continuous variable, with a suggestion of bimodal distribution, an arbitrary threshold (SFDI=0.3) was chosen to divide the measurements into two groups: a group with good overall concordance and the other with noticeable discordance. Sixty-one percent of the measurements fall in the group with good concordance. For the group of measurements with good overall structure-function concordance (SFDI≤0.3), the mean SFDI was 0.25 (SD of 0.04) compared with the SFDI (0.41±0.08) for the group with noticeable discordance (SFDI>0.3).
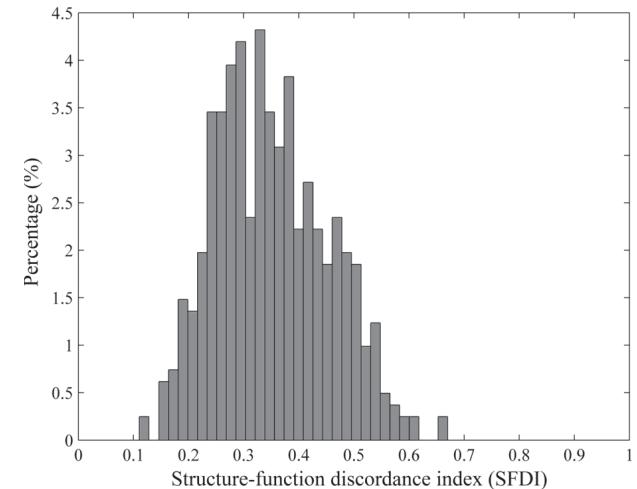
Figure 5-2. Histogram of structure-function discordance index (SFDI) in this test-retest sample of glaucomatous eyes.

Four examples with little, moderate and significant structure-function discordance are illustrated in Figure 5-3 to Figure 5-6. The SFDI and structure-function PDM are demonstrated as clinical tools to identify the disagreement between structural and functional measurements in these four examples.
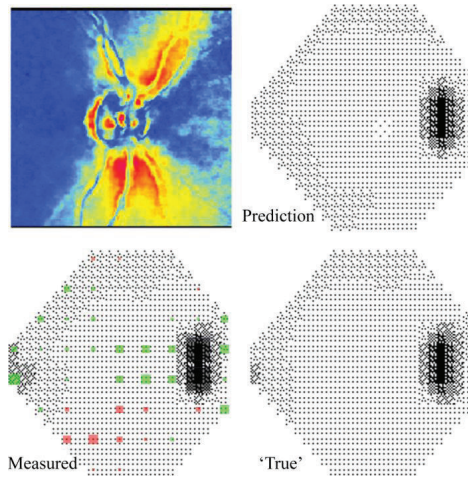
Figure 5-3. An example of good structure-function concordance (SFDI=0.15). The GDxVCC RNFL thickness predicted early nasal defect that was also indicated in measured and 'true' visual fields (VFs). Structure-function pattern discordance map (PDM) superimposed on the measured VF shows small discordance flags (green/red squares).

In Figure 5-3, the structure-predicted, measured and 'true' VFs all indicated an early defect in the nasal, nasal-superior and nasal-inferior areas. The SFDI of 0.15 demonstrates a good overall structure-function concordance and the structure-function PDM showed no significant flags with large 'discordance squares'.
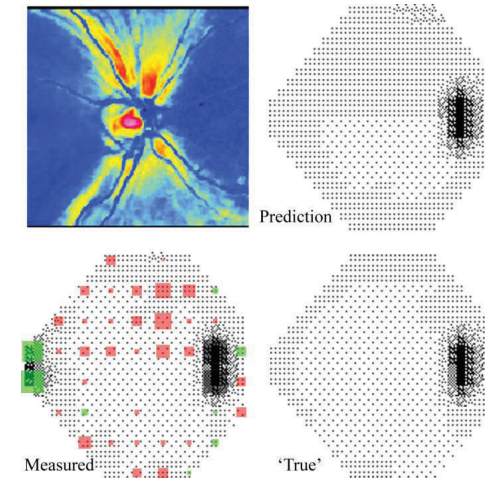
Figure 5-4. An example with satisfactory overall structure-function concordance (SFDI=0.28) but with local discordance. The RNFL thickness did not predict the nasal defect in the measured visual field, which was flagged by the large green squares in structure-function pattern discordance map.

The SFDI evaluates the overall structure-function discordance but it may fail to pick up local discordance. For example, the SFDI of 0.28 in Figure 5-4 is still within the arbitrary threshold of a satisfactory overall concordance but the summary index did not capture the discordance in the nasal area that was flagged by the PDM. In this example, RNFLT did not predict the nasal defect in the measured VF; this disagreement was flagged by two large green 'discordance squares' in the nasal area.
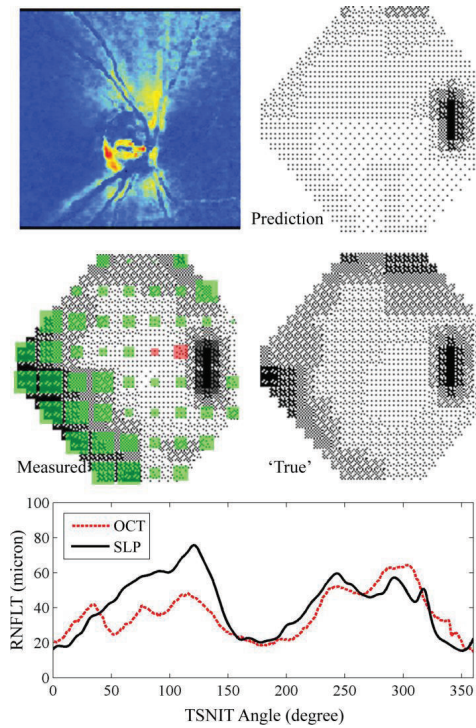
Figure 5-5. An example with substantial discordance (SFDI=0.58). The RNFL thickness prediction overestimated the measured visual field, especially in the inferior hemifield nasally, as flagged by the pattern discordance map. TSNIT RNFL thickness profile measured by OCT (appropriately rescaled (Leung et al., 2005)) and GDxVCC were compared.

Figure 5-5 is an example with moderate overall discordance (SFDI=0.58) and the locations of the discordance were further described by the PDM. The RNFLT overestimated the measured VF especially in the inferior part of the VF and the single measured VF was more consistent with the 'true' VF, showing that the structural measurement is more likely to be the cause of the discordance.
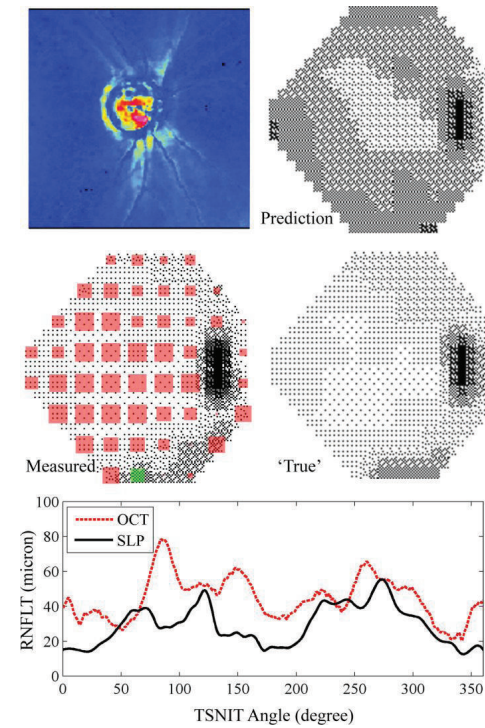
Figure 5-6. An example of significant structure-function discordance (SFDI=0.66). The RNFL thickness prediction significantly underestimated the whole measured visual field, as flagged by the pattern discordance map. TSNIT RNFL thickness profile measured by OCT (appropriately rescaled (Leung et al., 2005)) and GDxVCC were compared.

The last example, with high discordance (SFDI=0.66) is demonstrated in Figure 5-6. The overall discordance was flagged across the whole VF in the structure-function PDM. The single measured VF was closer to the 'true' VF and showed that the discordance was more likely to be caused by the GDxVCC RNFLT measurement.

5.6. Discussion

Clinicians are expected to use both structural and functional assessment of the optic nerve to evaluate and monitor glaucoma patients (European Glaucoma Society 2008). Quantitative tools for measuring visual function (for instance, SAP) and aspects of optic nerve anatomy, such as the RNFL (for instance, the GDxVCC), are in common clinical usage, but, as yet, there have been no analysis algorithms available to describe the level of agreement between these different

modes for the assessment of retinal ganglion cell integrity. The technique described in this chapter addresses this clinical need.

The BRBF technique maps the structural measurement into the same domain as the functional measurement so that the clinicians can consider both measurements in tandem. The SFDI and structure-function PDM are proposed as tools to quantify and visualise any disagreement between the two different types of measurements. Just as with a repeat visual field test, the clinician should not expect exact agreement, but serious disagreement may indicate that one or the other test was faulty, and sometimes the tests may need to be repeated.
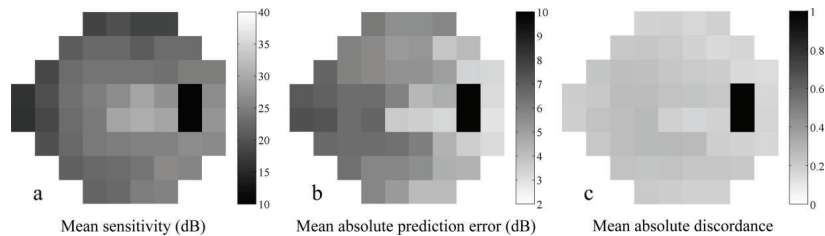
Figure 5-7. Illustration of spatial predictability. In this glaucomatous sample, the mean sensitivity (a), mean absolute prediction error (b) and mean absolute structure-function discordance (c) are shown for each visual field location with a grey scale.

The use of the predictability of the VF from the RNFLT by the BRBF for the discordance calculation played a key role in this method. Figure 5-7 demonstrates that any regional variation of discordance within the VF can be accounted for by sensitivity values at various locations. Figure 5-7a illustrates the average distribution of VF sensitivity in the glaucoma patients and Figure 5-7b illustrates the distribution of mean absolute prediction errors; there is a clear regional variation which is similar to the distribution of sensitivity. However, in Figure 5-7c, the mean unsigned structure-function discordance is calculated at each location in the VF. Unsigned structure-function discordance normalises the absolute prediction errors in dB to be the unsigned percentile difference according to the VF predictability at different levels of measured VF sensitivity (Figure 5-7). The use of predictability based on prediction percentiles attenuates the regional variation of the structure-function discordance, demonstrating that the prediction errors have been normalised with respect to the measured VF sensitivity. Therefore, the discordance value was less affected by the change in predictability caused by the level of VF sensitivity and tended to reflect the 'intrinsic' difference between the structural and functional measurements.

The most important potential use of the PDM is in alerting the clinician to the possibility of 'bad data' such as those shown in Figure 5-4, 5 and 6. Bad data may be either in the functional or structural measurement. In the functional domain, errors in quantifying neural loss may be systematic, such as the bias introduced by media opacity, or random, such as imprecision introduced by learning affects, fatigue, false positive responses, lens rim artifacts, and many other factors. In the structural domain, systematic error may be introduced by factors such as atypical scan pattern in GDxVCC images and measurement differences induced by variation in axial length; imprecision may be introduced by factors such as vitreous opacities, a poor tear film, poor fixation, and other known and unknown factors. Almost all sources of error will result in less than optimal concordance between structural and functional measurements and the SFDI and PDM provide an easily interpretable flag to suboptimal data.

The PDM in Figure 5-4 shows local discordance in the nasal area and indicates relative underestimation of the sensitivity in the measured VF. The structure-predicted and measured VFs were compared with the reference 'true' VF, which showed that this discordance was caused by test variability in the measured VF. The moderate sized red squares in the upper hemifield indicate that possible early RNFL loss has not manifested as VF loss in this patient.

Overall discordance is alerted by moderate SFDI and visual warning (large 'discordance squares') in the inferior part of PDM in Figure 5-5. The OCT measurement of RNFLT was used to validate the RNFLT measurement by GDxVCC. Due to measurement scaling differences between GDxVCC and OCT, the OCT measurements were adjusted by the scaling factors described by Leung et al (Leung et al., 2005). The overestimation of RNFLT by GDxVCC, when compared with the OCT measurement, was apparent especially in the superior region and, in this case, may be caused by the atypical scan pattern in the GDxVCC RNFLT map.

Discordance across the whole VF is indicated by SFDI and PDM in Figure 5-6. The agreement between the single and 'true' VFs shows that the inaccurate RNFLT measurement made by GDxVCC was more likely to be the cause of the discordance. This was confirmed by the OCT measurement indicating significantly thicker RNFLT than that from GDxVCC.

In the absence of a better reference standard, 'true' VF is used here as a means to compare how closely a measured VF resembles the presumed true field, and similarly how the field predicted from structural data compares to the presumed true field. Although the 'true' VF is derived from the measured VFs on these subjects, and is therefore biased to resemble any of the measured VFs, it was found that in some cases the structural data predicted a worse VF, and sometimes the measured VF test was worse than the 'true' field. Comparison of the mean difference of the structure-predicted and of the measured VFs from the 'true' VF showed that,

on average, the measured VF was closer to the 'true' VF. However, this may be, at least in part, bias from the fact that the 'true' VF was determined from measured VFs themselves. Simply put, a VF from the 5 VFs comprising the 'true' VF is likely to be close to the 'true' VF because of the method for calculating the 'true' VF. The structure-predicted VF can be compared with the measured VF to score the structure-function discordance without the 'true' VF reference used for this scientific study. So it should be emphasised that the construction of a 'true' VF test from available replicate testing was used in this study for illustrative purposes to demonstrate features of the method, but is not necessary when the method is applied in a clinical setting or in other scientific studies.

The principles reported in this study could be developed for any structural measurement. The various commercial products of other imaging techniques (such as spectral-domain OCT) need to be evaluated independently. Furthermore, there is already a new version of the SLP (GDxECC, Carl Zeiss Meditec, Dublin, CA, USA), which is likely to have a better prediction accuracy compared to the GDxVCC used in this study.

In conclusion, the work presented in this chapter presents, for the first time, a methodology that quantifies discordance between structural and functional measurements of glaucoma damage. The analysis is presented in a novel and clinically useful way such that high discordance may alert clinicians to an instance of poor quality test data, and low discordance may be reassuring that structural and functional findings match.

Chapter 6 : Conclusions and future work

Current clinical measurements of structure and function are simply surrogates of the real target biomarker for glaucoma, namely the integrity of retinal ganglion cells and their axons. These measurements are 'contaminated' by various noise and artefacts during the data acquisition. Moreover, as illustrated in the previous chapters these measurements yield high dimensional data that typically exhibits relationships that are non-linear and interact in complex ways. Extracting intrinsic *knowledge* from these complex measurements is critical for the assessment of glaucoma and the understanding of the physiological mechanism. Moreover, acquisition of clinical measurements is 'expensive' and consumes large amount of time and resources. These precious clinical measurements should be utilised in the most efficient way so more *knowledge* can be extracted from them in their singular and combined form; this theme has been central to the work presented in this thesis and is further elaborated on in this concluding chapter.

6.1. Applications of knowledge mining

Knowledge mining isn't a simply data analysis technique, but is a procedure of dealing with complex problems using an iterative divide-and-conquer solution to large-scale questions: smaller problems can be solved within manageable scales and the solutions to these individual problems are further used as the input to generate *knowledge* with more complexity and/or better quality. These individual problems and their solutions form the building blocks of *knowledge mining*, which can be divided into four categories: *feature extraction and quantification*, *information restoration*, *information integration* and *accuracy/precision improvement*. What follows is a summary of the *knowledge mining* approach illustrated by the work and conclusions presented in this thesis.

Example tasks of *feature extraction and quantification* can be image processing (separating and quantifying useful information from image space) or machine classification (knowledge extraction from a raw measurement condensed into a binary variable). The volumetric image segmentation of spectral-domain optical coherence tomography (SD-OCT) images reported in chapter 3 is also an example of feature extraction and quantification: the morphology of retinal tissue layers was extracted from image volumes and the retinal nerve fibre layer thickness (RNFLT) and vessel structures were quantified as the condensed *knowledge* from the 3D image space. Extracted features, e.g. RNFLT, can be used as the input to further *knowledge mining* process such as the structure-function relationship model reported in chapter 4.

Information restoration becomes useful when desired *knowledge* about the measurement is missing during the data acquisition. Because the target information is not complete in the original measurement, information restoration is normally assisted by other measurements,

which may be repeated measurements of the same type, or other types of measurements. Examples of *information restoration* include image 'super-resolution' where the lost information in a single low-resolution image is reconstructed into a higher resolution image by combining multiple repeated low-resolution images (Park et al., 2003). Another example, from ophthalmology, is 3D-reconstruction of the optic disc structure from stereo fundus image (Tang et al., 2010, Xu and Chutatape, 2004) in which the 3D information about the optic disc is restored from 2D stereo fundus photos taken at different angles. In this thesis, the optical coherence tomography (OCT) scan circle alignment reported in chapter 2 is another example of *information restoration*. The displacement of the OCT circular scan is critical information but it is not recorded or precisely estimated during image acquisition. Assisted by the retinal fundus image, the proposed method in chapter 2 restored information by aligning the vessel features in both OCT and fundus images. Restored information about the displacement of OCT circular scan has significant impact on the reproducibility of RNFLT measurement (see chapter 2) and is therefore important in controlling and interpreting the measurement variability.

Information integration focuses on 'bridging' multiple types of measurements. It might be that two or more types of measurement are taken independently in practice and differ, for example, by acquisition technique, target variable, unit, dimensionality and dynamic range, but these measurements may be correlated in a complicated way. In many cases, a clear understanding about the relationship among these measurements is incomplete or not available. For instance, retinal structure measurements acquired by imaging techniques and the visual function measurements acquired by perimetry are taken independently in the clinical assessment of glaucoma. In spite of the reasonable assumption that the structure and function measurements are related according to the anatomy (Garway-Heath et al., 2000), good longitudinal studies indicate the contrary, especially when it comes to disease changes. For example, Strouthidis et al (Strouthidis and Garway-Heath, 2009) examined longitudinal changes in structural and functional measures and concluded, "....*agreement between optic disc progression and visual field progression is the exception rather than the rule*..." and this observation is supported by other carefully conducted longitudinal studies (Artes and Chauhan, 2005). Such observations cause a dilemma when developing *information integration* approaches attempting to bridge the structural and functional measurements. One solution to this problem involves using the existing but incomplete prior knowledge about the structure-function relationship. In this 'direct' approach, the prior knowledge is explicitly used to build the model that links the structural and functional measurements. For example, many studies (Anton et al., 1998, Gardiner et al., 2005, Mai et al., 2007, Bowd et al., 2006, Schlottmann et al., 2004) used the estimated anatomical map (Garway-Heath et al., 2000) between the visual field and peripapillary sectors in retinal structure measurement to examine the association between the retinal structure and visual

function in glaucoma. These models built from 'direct' use of prior knowledge are relatively simple and easy to understand, but they are also limited by the incompleteness of the prior knowledge. On the other hand, an 'indirect' approach attempts to learn the relationship from the data, leaving the prior knowledge for the purpose of validation. An improperly designed model may be easily misled by the measurement noise and variability, potentially producing erroneous *knowledge* that then contradicts the prior knowledge. For example, the classic log linear model explored in chapter 4 failed to quantify a meaningful structure-function relationship in a high-resolution format. However, the novel BRBF model proposed in chapter 4 was able to quantify a structure-function relationship with high resolution by 'learning' from the data and the results are consistent with the anatomical prior knowledge. The way of using prior knowledge in this 'indirect' approach forms unbiased criteria for validating the method. Results from the *information integration* can then be used to develop more *knowledge* in the form of tools that can be applied to practical and clinical problems, such as the new visualization and quantification tools for the structure-function measurement discordance derived and illustrated in chapter 5.

The goal of the *knowledge mining* used in the context of the work presented in this thesis is *accuracy/precision improvement* of the measurements used in the clinical assessment of glaucoma. The OCT scan circle alignment method (chapter 2) and *FloatingCanvas* for quantification of retinal structures in SD-OCT images (chapter 3) serve as tools to improve the measurement variability in the assessment of RNFLT. Moreover, the quantified structure-function relationship and its discordance quantification developed in chapter 4 and 5 can also be used to improve the reproducibility of the measurements. This forms another example of the goal of using *knowledge mining*, which is an attempt to see how a combined measurement of the RNFLT and the visual field can better be used for tracking changes or progression detection in glaucoma. What follows is a very brief summary of this work. Some of the data and graphs presented in this next section have been presented as a read paper at the Association for Research in Vision and Ophthalmology Meeting, Fort Lauderdale, Florida, USA in May 2009 (See List of supporting publications).

6.2 Combining structural and functional measurements to improve reproducibility of follow up data in glaucoma

Briefly, in this work the RNFLT and the measured visual field (VF) were combined by calculating the weighted average of the RNFLT-predicted and measured VF. The weight of the predicted VF is arbitrarily chosen to be equal to the structure-function discordance index (SFDI) proposed in chapter 5 with an upper limit of 0.5. The weight of the measured VF is one minus that of the predicted VF. When the RNFLT and measured VF have a good concordance (SFDI

is near to 0), the predicted and measured VFs are similar so the combined VF is set to be equal to the measured VF. More weight is given to the predicted VF as the SFDI increases, and this proportion has an upper limit of 0.5 if SFCI≥0.5 when the combined VF is the average of structure-predicted and measured VFs.

Some of the test-retest dataset used in chapter 5 was used to illustrate the calculation of combined VF, in comparison with the predicted, measured and 'true' (calculated as the median of five repeated VFs) VFs. The combined VFs for four examples with their SFDI values in chapter 5 are shown in Figure 6-1. Note that the combined VFs appear to be closer to the 'true' VFs in these examples.

GDxVCC SLP	Predicted VF	Measured VF	Combined VF	'True' VF
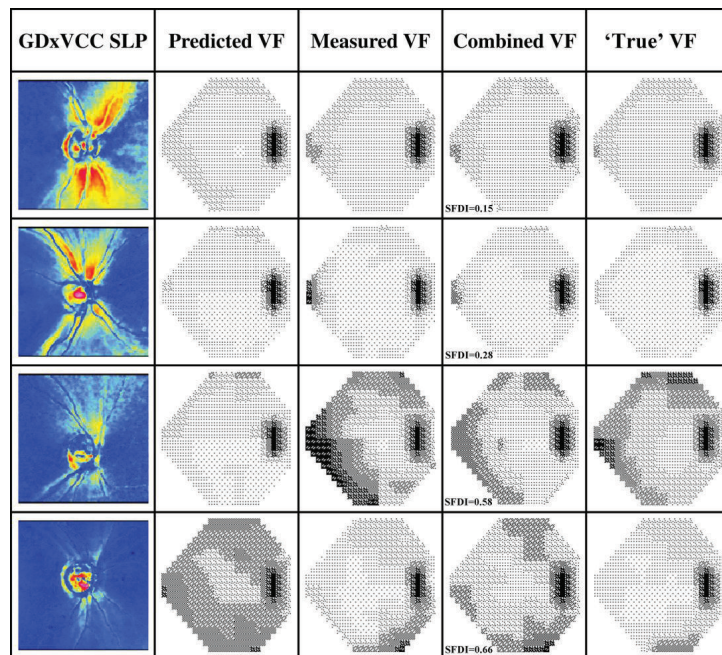
Figure 6-1: The calculated combined VF compared with predicted, measured and 'true' VFs for four patient examples. The corresponding GDxVCC SLP image and VFs are listed in a row for each eye. The SFDI values are given beside the combined VF.

The hypothesis for this study was that the reproducibility of the combined VF would be better than the reproducibility of an individual VF. This was examined by plotting point-wise sensitivities in one combined VF against those in another and was shown in a box plot similar to

that used in chapter 4. The combined VF pairs were exhaustively drawn all possible pairs from the 5 repeat measures of RNFLT and VF for each eye. The reproducibility of single VF test was analysed in a similar fashion and was compared with that of combined VF in Figure 6-2.

The 90% retest limits (5th and 95th percentile of retest sensitivity) in Figure 6-2 describe the range of the retest VF sensitivity (within the 90% confidence interval) with respect to the baseline VF sensitivity. Recollect, the retest limits of a perfectly precise and accurate measurement would overlap on the diagonal line of unity. Therefore, with highly reproducible measurement, the retest limits should be closely centred on the diagonal. In Figure 6-2a, the single VF test demonstrates poor reproducibility especially at the lower end of the VF sensitivity and this has been shown before (Artes et al., 2002). The combined VF, on the other hand, shows much less variability indicated by the narrower 90% retest limit in Figure 6-2b.

Figure 6-2: The reproducibility of (a) single VF sensitivity and (b) combined VF sensitivity. In both plots, the distribution of retest VF is stratified by the base line VF. Each box-whisker bar summarises the retest distribution over a 2dB range from 0 to >36dB. Thin vertical lines indicate 90% retest limits (5th and 95th percentile of retest sensitivity), the 'boxes' indicate the interquartile range of the retest VF (25th and 75th percentile) with the line in the box indicates the median. The 90% retest limits are connected by thick black curves.

Because the combined VF was calculated as the weighted mean of two VFs: structure-predicted and measured VFs, it would be unfair to compare the reproducibility of the combined VF against that of the single VF. In order to generate a fairer comparison, the reproducibility of the mean of two measured VFs was examined. To do this, two VFs were exhaustively drawn from the five repeat VFs and the point-wise mean of the two VFs was calculated. In Figure 6-3a, the mean VFs were plotted against another. The reproducibility of the mean VF is higher compared to the single VF (Figure 6-2a), but is still slightly poorer than that of the combined VF (Figure

6-3b). These graphs therefore provide some evidence of the benefits of taking a structural measure and a VF, in preference to two VFs, when the measurement are combined as described in this thesis
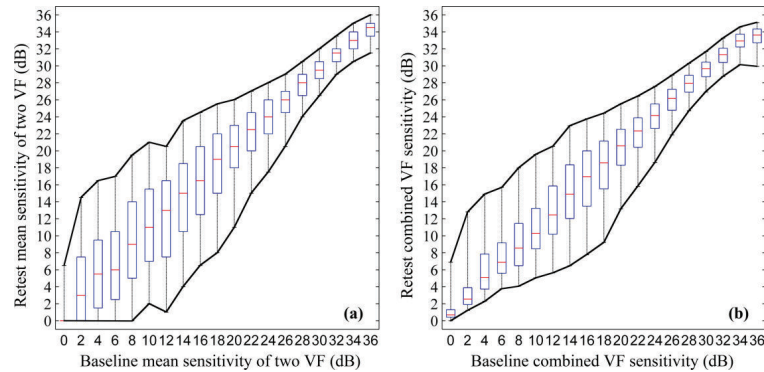
Figure 6-3: The reproducibility of (a) mean of 2 VF sensitivities compared with (b) combined VF sensitivity. In both plots, the distribution of retest VF is stratified by the base line VF. Each box-whisker bar summarises the retest distribution over a 2dB range from 0 to >36dB. Thin vertical lines indicate 90% retest limits (5th and 95th percentile of retest sensitivity), the 'boxes' indicate the interquartile range of the retest VF (25th and 75th percentile) with the line in the box indicates the median. The 90% retest limits are connected by thick black curves.
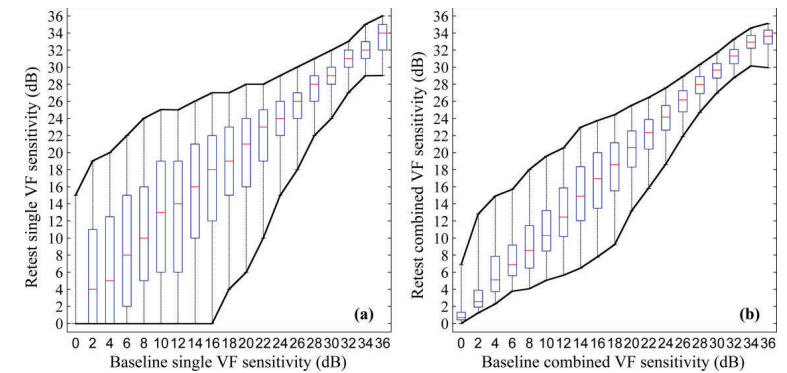
The reproducibility of the combined VF makes it potentially more robust in detecting the worsening of glaucoma over time. To show this, a *pseudo* time series of VFs with no underlying worsening was generated by considering the five repeat measurements taken over 2.5 years with a test interval of 6 month. With five repeat tests it was possible to generate 120 of these pseudo series by reshuffling the order of the five tests. Linear regression was used to detect the change $\geq$-1dB/year at each location of the single and combined VF. The progression of the whole VF (both single and combined VFs) time series is defined as $\geq$2 contiguous point being marked as progressing by >-1dB/year.

This assumption of no expected change in the pseudo series means that any flagged change is a false positive result. This false-positive rate of linear regression change detection with various levels of *p*-value for the single and combined VFs is summarized in Figure 6-4. It is evident that the single VF gives a higher false-positive rate in the linear regression change detection with all *p*-values. For example, at *p*-value of 0.05, the combined VF had lower false-positive rate (3.1%) compared with that of the single VF (5.8%). In practical terms, by using the combined measurement, instead of a single visual field, it might be possible to halve the false positive rate

associated with a standard method for progression detection in glaucoma. This result suggest that a combined measure might provide a new biomarker for progression detection but a more thorough examination, including a consideration of the sensitivity of the method is the subject of planned future work (below).
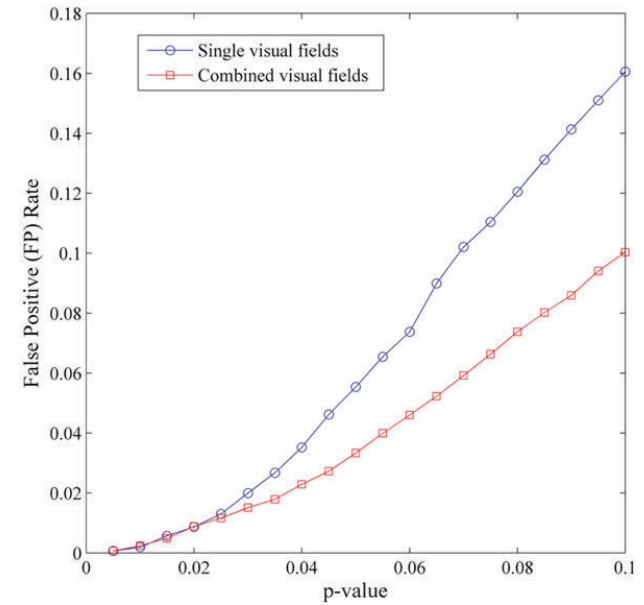
Figure 6-4: The false-positive rate of linear regression change detection using simulated time series of single VF (blue line) and combined VF (red line). The linear regression was carried out using various *p*-value (*x*-axis) criteria.

6.3. Future work

This thesis has illustrated the application of newly developed *knowledge mining* procedures to improve the acquisition and/or accuracy of the clinical measurements used in glaucoma assessment. This has led to new analysis tools to detect and manage glaucoma; translating these into clinical tools is the subject of future studies, some of which are briefly outlined here:

(1) Validating the sensitivity of detecting disease progression when using the new combined structure-function biomarker introduced in this chapter (and chapter 4 and 5) on a real longitudinal data set.

(2) Testing and extending the reported methods to see how generalizable in their application to other retinal imaging platforms. For example, the OCT scan circle alignment algorithm (chapter 2) could be realised on SD-OCT circular scan protocols.

(3) *FloatingCanavs* (chapter 3) could be incorporated into the RTVue-100 system (Optovue, Fremont, CA, USA), and extended to detect intra-retinal tissue layers in macula images taken from SD-OCT.

(4) The BRBF model (chapter 4 and 5) could be applied to RNFLT measurements produced by SD-OCT. The training dataset of BRBF could also be expanded by combining datasets from various institutions to cover a wider spectrum of the disease.

(5) Incorporating temporal information into the structure-function relationship model. Investigating the evolvement of structure-function relationship over time and its impact on the rate of progression during the worsening of glaucoma, with the aim of developing a progression biomarker that combines both structural and functional measures.

List of supporting publications

Peer reviewed papers

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., LEMIJ, H., REUS, N. J., HEALEY, P. R., MITCHELL, P., HO, T. & GARWAY-HEATH, D. F. 2010. Predicting Visual Function from the Measurements of Retinal Nerve Fibre Layer Structure. *Invest Ophthalmol Vis Sci*, 51, 5657-5666.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., HO, T. & GARWAY-HEATH, D. F. 2010. FloatingCanvas: quantification of 3D retinal structures from spectral-domain optical coherence tomography. *Opt Express*, 18, 24595-24610.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., WOLLSTEIN, G. & GARWAY-HEATH, D. F. 2011. Aligning Scan Acquisition Circles in Optical Coherence Tomography Images of The Retinal Nerve Fibre Layer. IEEE Trans Med Imaging, Accepted. In Press. DOI 10.1109/TMI.2011.2109962.

ZHU, H., CRABB, D. P., FREDETTE, M., ANDERSON, D. R. & GARWAY-HEATH, D. F. 2010. Quantifying Discordance Between Structure and Function Measurements in the Clinical Assessment of Glaucoma. *Arch Ophthalmol*, Accepted. In press.

Paper in conference proceedings

ZHU, H., CRABB, D. P. & GARWAY-HEATH, D. F. A Bayesian Radial Basis Function Model to Link Retinal Structure and Visual Function in Glaucoma. *The 3rd International Conference of Bioinformatics and Biomedical Engineering*, June 2009 Beijing, China. IEEE, 1-4.

Meeting abstracts

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., LEMIJ, H., REUS, N. J., HEALEY, P. R., MITCHELL, P., HO, T. & GARWAY-HEATH, D. F. 2006. A Bayesian Retinal Structure-Function Relationship Quantifier. *17th Imaging and Perimetric Society Symposium*. Portland, Oregon, USA.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G. & GARWAY-HEATH, D. F. 2007. Aligning Sequential Stratus OCT RNFL Scans - Solving the Problem. *Invest Ophthalmol Vis Sci*, 48, 516.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G. & GARWAY-HEATH, D. F. 2008. 3DSEGMENT: Quantitative Analysis of Retinal Structures From Fourier-Domain Optical Coherence Tomography (FD-OCT) 3D Volume Scans. *Invest Ophthalmol Vis Sci*, 49, 3764.

ZHU, H., FREDETTE, M. J., ANDERSON, D. R., GARWAY-HEATH, D. F. & CRABB, D. P. 2008. Quantifying concordance between structure and function measurements in the clinical assessment of glaucoma. *18th Imaging and Perimetric Society Symposium*. Nara, Japan.

ZHU, H., CRABB, D. P., ANDERSON, D., FREDETTE, M. J. & GARWAY-HEATH, D. F. 2009. Combining structural and functional measurements to improve reproducibility of follow up data in glaucoma. *Invest Ophthalmol Vis Sci*, 50, 2572.

References

AHLBERG, J. H., NILSON, E. N. & WALSH, J. L. 1967. The theory of splines and their applications. New York: Academic Press.

AHLERS, C., SIMADER, C., GEITZENAUER, W., STOCK, G., STETSON, P., DASTMALCHI, S. & SCHMIDT-ERFURTH, U. 2008. Automatic segmentation in three-dimensional analysis of fibrovascular pigmentepithelial detachment using high-definition optical coherence tomography. *Br J Ophthalmol,* 92**,** 197-203.

ALTMAN, D. G. & ROYSTON, P. 2000. What do we mean by validating a prognostic model. *Statistics in Medicine,* 19**,** 453-473.

ANDERSON, R. S. 2006. The psychophysics of glaucoma: Improving the structure/function relationship. *Prog Ret Eye Res,* 25**,** 79-97

ANTON, A., YAMAGISHI, N., ZANGWILL, L., SAMPLE, P. & WEINREB, R. 1998. Mapping structural to functional damage in glaucoma with standard automated perimetry and confocal scanning laser ophthalmoscopy. *Am J Ophthalmol,* 125(4)**,** 436-446.

ARTES, P. H. & CHAUHAN, B. C. 2005. Longitudinal changes in the visual field and optic disc in glaucoma. *Progress in Retinal and Eye Research,* 24**,** 333-354.

ARTES, P. H., IWASE, A., OHNO, Y., KITAZAWA, Y. & CHAUHAN, B. C. 2002. Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Invest Ophthalmol Vis Sci,* 43**,** 2654-2659.

BAGGA, H., GREENFIELD, D. S. & FEUER, W. J. 2005. Quantitative assessment of atypical birefringence images using scanning laser polarimetry with variable corneal compensation. *Am J Ophthalmol,* 139**,** 437-46.

BARTZ-SCHMIDT, K. U., THUMANN, G., JONESCU-CUYPERS, C. P. & KRIEGLSTEIN, G. K. 1999. Quantitative morphologic and functional evaluation of the optic nerve head in chronic open-angle glaucoma. *Surv Ophthalmol,* 44 Suppl 1**,** S41-53.

BENGTSSON, B., BIZIOS, D. & HEIJL, A. 2005. Effects of Input Data on the Performance of a Neural Network in Distinguishing Normal and Glaucomatous Visual Fields. *Invest Ophthalmol Vis Sci,* 46**,** 3730-3736.

BENGTSSON, B. & HEIJL, A. 1998. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand,* 76**,** 268-272.

BENGTSSON, B. & HEIJL, A. 1999a. Comparing significance and magnitude of glaucomatous visual field defects using the SITA and Full Threshold strategies. *Acta Ophthalmol Scand,* 77**,** 143-146.

BENGTSSON, B. & HEIJL, A. 1999b. Inter-subject variability and normal limits of the SITA Standard, SITA Fast, and the Humphrey Full Threshold computerized perimetry strategies, SITA STATPAC. *Acta Ophthalmol Scand,* 77**,** 125-129.

BENGTSSON, B., OLSSON, J., HEIJL, A. & ROOTZEN, H. 1997. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand,* 75**,** 368-375.

BENNETT, A. G. & RABBETTS, R. B. 1998. *Bennett and Rabbetts' clinical visual optics,* Boston, Butterworth-Heinemann, Oxford.

BISHOP, C. M. 1996. *Neural network for pattern recognition*, Oxford University Press.

BLUMENTHAL, E. Z., HORANI, A., SASIKUMAR, R., GARUDADRI, C., UDAYKUMAR, A. & THOMAS, R. 2006. Correlating structure with function in end-stage glaucoma. *Ophthalmic Surg Lasers Imaging,* 37**,** 218-223.

BLUMENTHAL, E. Z., SAMPLE, P. A., BERRY, C. C., LEE, A. C., GIRKIN, C. A., ZANGWILL, L., CAPRIOLI, J. & WEINREB, R. N. 2003. Evaluating several sources of variability for standard and SWAP visual fields in glaucoma patients, suspects, and normals. *Ophthalmology,* 110**,** 1895-1902.

BOCK, M., BRANDT, A. U., DORR, J., PFUELLER, C. F., OHLRAUN, S., ZIPP, F. & PAUL, F. 2010. Time domain and spectral domain optical coherence tomography in multiple sclerosis: a comparative cross-sectional study. *Mult Scler,* 16**,** 893-6.

BOER, J. F. D., CENSE, B., PARK, B. H., PIERCE, M. C., TEARNEY, G. J. & BOUMA, B. E. 2003. I. Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography. *Opt Lett,* 28**,** 2067-2069.

BOUMA, B. E. & TEARNEY, G. J. 2001. *Handbook of Optical Coherence Tomography,* New York, Marcer Dekker Inc.

BOURNE, R. R., MEDEIROS, F. A., BOWD, C., JAHANBAKHSH, K., ZANGWILL, L. M. & WEINREB, R. N. 2005. Comparability of Retinal Nerve Fiber Layer Thickness Measurements of Optical Coherence Tomography Instruments. *Invest Ophthalmol Vis Sci,* 46**,** 1280-1285.

BOWD, C., CHAN, K., ZANGWILL, L. M., GOLDBAUM, M. H., LEE, T.-W., SEJNOWSKI, T. J. & WEINREB, R. N. 2002. Comparing Neural Networks and Linear Discriminant Functions for Glaucoma Detection Using Confocal Scanning Laser Ophthalmoscopy of the Optic Disc. *Invest Ophthalmol Vis Sci,* 43**,** 3444-3454.

BOWD, C., ZANGWILL, L. M., MEDEIROS, F. A., TAVARES, I. M., HOFFMANN, E. M., BOURNE, R. R., SAMPLE, P. A. & WEINREB, R. N. 2006. Structure and Function in Glaucoma: The Relationship between a Functional Visual Field Map and an Anatomic Retinal Map. *Invest Ophthalmol Vis Sci,* 47**,** 2889-2895.

BRIGATTI, L. & CAPRIOLI, J. 1995. Correlation of visual field with scanning confocal laser optic disc measurements in glaucoma. *Arch Ophthalmol,* 113(9)**,** 1191-1194.

BRIGATTI, L., HOFFMAN, D. & CAPRIOLI, J. 1996. Neural networks to identify glaucoma with structural and functional measurements. *Am J Ophthalmol,* 121**,** 511-521.

BRIGATTI, L., NOURI-MAHDAVI, K., WEITZMAN, M. & CAPRIOLI, J. 1997. Automatic detection of glaucomatous visual field progression with neural networks. *Arch Ophthalmol,* 115**,** 725-728.

BUDENZ, D. L., CHANG, R. T., HUANG, X., KNIGHTON, R. W. & TIELSCH, J. M. 2005. Reproducibility of Retinal Nerve Fiber Thickness Measurements Using the Stratus OCT in Normal and Glaucomatous Eyes *Invest Ophthalmol Vis Sci,* 46**,** 2440-2443.

CAMPBELL, R. J., COUPLAND, S. G., BUHRMANN, R. R. & KERTES, P. J. 2007. Effect of Eccentric and Inconsistent Fixation on Retinal Optical Coherence Tomography Measures. *Arch Ophthalmol,* 125**,** 624-627.

CAPRIOLI, J. & MILLER, J. M. 1988. Correlation of structure and function in glaucoma. *Ophthalmology,* 95**,** 723-727.

CARPINETO, P., CIANCAGLINI, M., ZUPPARDI, E., FALCONIO, G., DORONZO, E. & MASTROPASQUA, L. 2003. Reliability of nerve fiber layer thickness measurements using optical coherence tomography in normal and glaucomatous eyes. *Ophthalmology,* 110**,** 190-195.

CHEN, P. P. & BUDENZ, D. L. 1998. The effect of cataract on the visual field of eyes with chronic open-angle glaucoma. *Am J Ophthalmol,* 125**,** 325-333.

CHIU, S. J., LI, X. T., NICHOLAS, P., TOTH, C. A., IZATT, J. A. & FARSIU, S. 2010. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. *Opt Express,* 18**,** 19413-19428.

CHO, J. W., SUNG, K. R., HONG, J. T., UM, T. W., KANG, S. Y. & KOOK, M. S. 2010. Detection of Glaucoma by Spectral Domain-scanning Laser Ophthalmoscopy/Optical Coherence Tomography (SD-SLO/OCT) and Time Domain Optical Coherence Tomography. *J Glaucoma,* Publish Ahead of Print, 10.1097/IJG.0b013e3181d1d332.

CIOS, K. J., PEDRYCZ, W., SWINIARSKI, R. W. & KURGAN, L. A. 2007. *Data Mining: A Knowledge Discovery Approach* New York, Springer.

COFFEY, M., REIDY, A., WORMALD, R., XIAN, W. X., WRIGHT, L. & COURTNEY, P. 1993. Prevalence of glaucoma in the west of Ireland. *Br J Ophthalmol,* 77**,** 17-21.

COHEN, J. 1968. Multiple Regression as a general data-analytic system. *Psychological Bulletin,* 70**,** 426-443.

COPAS, J. B. 1983. Regression, prediction and shrinkage (with discussion). *J Roy Stat Soc B Met,* 45**,** 311-354.

DEMPSTER, A., LAIRD, N. & RUBIN, D. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society,* 39(1)**,** 1–38.

DIELEMANS, I., VINGERLING, J. R., WOLFS, R. C., HOFMAN, A., GROBBEE, D. E. & JONG, P. T. 1994. The prevalence of primary open-angle glaucoma in a population-based study in The Netherlands. The Rotterdam Study. Ophthalmology, 101**,** 1851-1855.

DREHER, A. W., REITER, K. & WEINREB, R. N. 1992. Spatially resolved birefringence of the retinal nerve fiber layer assessed with a retinal laser ellipsometer. *Appl Opt,* 31**,** 3730-3735.

European Glaucoma Society 2008. Terminology and guidelines for glaucoma (3rd edition). Savona, Italy: European Glaucoma Society.

FABRITIUS, T., MAKITA, S., MIURA, M., MYLLYL˙A, R. & YASUNO, Y. 2009. Automated segmentation of the macula by optical coherence tomography. *Opt Express,* 17**,** 15659-15669.

FAYYAD, U., PIATETSKY-SHAPIRO, G. & SMYTH, P. 1996a. The KDD process for extracting useful knowledge from volumes of data. *Commun. ACM,* 39**,** 27-34.

FAYYAD, U. M., PIATETSKY-SHAPIRO, G., SMYTH, P. & UTHURUSAMY, R. 1996b. *Advances in Knowledge Discovery and Data Mining*, MIT Press.

FERCHER, A. F., HITZENBERGER, C. K., KAMP, G. & EL-ZAIAT, S. Y. 1995. Measurement of intraocular distances by backscattering spectral interferometry. *Opt Commun,* 117**,** 43-48.

FERNÁNDEZ, D. C., SALINAS, H. M. & PULIAFITO, C. A. 2005. Automated detection of retinal layer structures on optical coherence tomography images. *Opt Express,* 13**,** 10200-10216.

FOROOGHIAN, F., CUKRAS, C., MEYERLE, C. B., CHEW, E. Y. & WONG, W. T. 2008. Evaluation of Time Domain and Spectral Domain Optical Coherence Tomography in the Measurement of Diabetic Macular Edema. *Invest Ophthalmol Vis Sci,* 49**,** 4290-4296.

FOSTER, P. & JOHNSON, G. 2000. Primary angle closure: epidemiology and mechanism. *In:* HITCHINGS, R. (ed.) *Glaucoma.* London: BMJ Publishing Group.

FRANGI, A. F., NIESSEN, W. J., VINCKEN, K. L. & VIERGEVER, M. A. 1998. Multiscale vessel enhancement filtering. *Medical Image Computing and Computer-Assisted Intervention* 130-137.

GABRIELE, M. L., ISHIKAWA, H., WOLLSTEIN, G., BILONICK, R. A., TOWNSEND, K. A., KAGEMANN, L., WOJTKOWSKI, M., SRINIVASAN, V. J., FUJIMOTO, J. G., DUKER, J. S. & SCHUMAN, J. S. 2008. Optical coherence tomography scan circle location and mean retinal nerve fiber layer measurement variability. *Invest Ophthalmol Vis Sci,* 49**,** 2315-2321.

GARDINER, S. K., JOHNSON, C. A. & CIOFFI, G. A. 2005. Evaluation of the Structure-Function Relationship in Glaucoma. *Invest Ophthalmol Vis Sci,* 46**,** 3712-3717.

GARVIN, M., ABRAMOFF, M., KARDON, R., RUSSELL, S., WU, X. & SONKA, M. 2008. Intraretinal Layer Segmentation of Macular Optical Coherence Tomography Images Using Optimal 3-D Graph Search. *IEEE Trans Med Imaging,* 27**,** 1495-1505.

GARVIN, M. K., ABRAMOFF, M. D., WU, X., RUSSELL, S. R., BURNS, T. L. & SONKA, M. 2009. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans Med Imaging,* 28**,** 1436-47.

GARWAY-HEATH, D. F., HOLDER, G. E., FITZKE, F. W. & HITCHINGS, R. A. 2002. Relationship between Electrophysiological, Psychophysical, and Anatomical Measurements in Glaucoma. *Invest Ophthalmol Vis Sci,* 43**,** 2213-2220.

GARWAY-HEATH, D. F., POINOOSAWMY, D., FITZKE, F. W. & HITCHINGS, R. A. 2000. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology,* 107**,** 1809-1815.

GARWAY-HEATH, D. F., VISWANATHAN, A. & WESTCOTT, M. 1999. Relationship between perimetric light sensitivity and optic disc neuroretinal rim area. *In:* WALL, M.

& WILD, J. M. (eds.) *Perimetry Update 1998/1999.* The Netherlands: Kugler Publications The Hague.

GDxVCC Instrument Manual 2004. *RNFL Analysis with GDx VCC: A Primer and Clinical Guide,* San Diego, CA, Laser Diagnostic Technologies, Inc.

GOLDBAUM, M. H., SAMPLE, P. A., WHITE, H., COLT, B., RAPHAELIAN, P., FECHTNER, R. D. & WEINREB, R. N. 1994. Interpretation of automated perimetry for glaucoma by neural network. *Invest Ophthalmol Vis Sci,* 35**,** 3362-3373.

GOLUB, G. & KAHAN, W. 1965. Calculating the singular value and pseudo-inverse of a matrix. *SIAM Numerical Analysis,* 63**,** 205-224.

GONZALEZ-HERNANDEZ, M., PABLO, L. E., ARMAS-DOMINGUE, K., VEGA, R. R., FERRERAS, A. & ROSA, M. G. 2009. Structure-function relationship depends on glaucoma severity. *Br J Ophthalmol,* 93**,** 1195-1199.

GÖTZINGER, E., PIRCHER, M., GEITZENAUER, W., AHLERS, C., BAUMANN, B., MICHELS, S., SCHMIDT-ERFURTH, U. & HITZENBERGER, C. K. 2008. Retinal pigment epithelium segmentation by polarization sensitive optical coherence tomography. *Opt Express,* 16**,** 16410-16422.

GREENFIELD, D. 2002. Optic nerve and retinal nerve fiber layer analyzers in glaucoma. *Curr Opin Ophthalmol,* 13**,** 68-76.

GUEDES, V., SCHUMAN, J. S., HERTZMARK, E., WOLLSTEIN, G., CORRENTI, A., MANCINI, R., LEDERER, D., VOSKANIAN, S., VELAZQUEZ, L., PAKTER, H. M., PEDUT-KLOIZMAN, T., FUJIMOTO, J. G. & MATTOX, C. 2003. Optical Coherence Tomography Measurement of Macular and Nerve Fiber Layer Thickness in Normal and Glaucomatous Human Eyes. *Ophthalmology,* 110**,** 117-189.

HAEKER, M., SONKA, M., KARDON, R., SHAH, V. A., WU, X. & M ABRAMOFF 2007. Automated segmentation of intraretinal layers from macular optical coherence tomography images. *Proc. the SPIE: Medical Imaging*, 6512.

HALEY, M. J. 1987. *The Field Analyzer Primer, 2nd ed,* San Leandro, CA, Allergan Humphrey.

HAN, J., KAMBER, M. & PEI, J. 2005. *Data Mining: Concepts and Techniques, 2nd edition*, Morgan Kaufmann.

HAND, D. J. 1998. Data Mining: Statistics and More? *The American Statistician,* 52**,** 112-118.

HARRINGTON, D. O. 1976. *The Visual Fields: A Textbook and Atlas of Clinical Perimetry*, St Louis.

HARWERTH, R. S., CARTER-DAWSON, L., SMITH, E. L. & CRAWFORD, M. L. 2005. Scaling the structure-function relationship for clinical perimetry. *Acta Ophthalmol Scand,* 83**,** 448-455.

HEE, M. R., IZATT, J. A., SWANSON, E. A., HUANG, D., SCHUMAN, J. S., LIN, C. P., PULIAFITO, C. A. & FUJIMOTO, J. G. 1995. Optical coherence tomography of the human retina. *Arch Ophthalmol,* 113**,** 325-332.

HEIJL, A. (ed.) 1991. *Perimetry Update 1990/1991,* Amsterdam: Kugler&Ghedini.

HEIJL, A., LESKE, M. C., BENGTSSON, B., L, L. H., BENGTSSON, B. & HUSSEIN, M. 2002. Reduction of intraocular pressure and glaucoma progression: results from the early manifest glaucoma trial. *Arch Ophthalmol,* 120**,** 1268-1279.

HEIJL, A., LINDGREN, G. & LINDGREN, A. 1991. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol,* 105**,** 1544-1549.

HEIJL, A., LINDGREN, G. & OLSSON, J. 1987. A package for the statistical analysis of visual fields. *Doc Ophthalmol Proc Ser,* 49**,** 153-168.

HEIJL, A., LINDGREN, G. & OLSSON, J. 1989. Test retest variability in glaucomatous visual fields. *Am J Ophthalmol,* 108**,** 130-135.

HENSON, D. B. 2000. *Visual Fields, 2nd ed*, Oxford University Press.

HENSON, D. B., CHAUDRY, S., ARTES, P. H., FARAGHER, E. B. & ANSONS, A. 2000. Response Variability in the Visual Field: Comparison of Optic Neuritis, Glaucoma, Ocular Hypertension, and Normal Eyes *Invest Ophthalmol Vis Sci,* 41**,** 417-421.

HENSON, D. B., EVANS, J., CHAUHAN, B. C. & LANE, C. 1996. Influence of fixation accuracy on threshold variability in patients with open angle glaucoma. *Invest Ophthalmol Vis Sci,* 37**,** 444-450.

HENSON, D. B., EVANS, J. & LANE, C. Fixation accuracy of patients with glaucoma during full threshold perimetry. *In:* MILLS, R. P. & WALL, M., eds. *Perimetry Update 1995*, 1995 Amsterdam. Kugler&Ghedini, 249-255.

HITCHINGS, R. 2000. *Glaucoma,* London, BMJ Publishing Group.

HO, J., SULL, A. C., VUONG, L. N., CHEN, Y., LIU, J., FUJIMOTO, J. G., SCHUMAN, J. S. & DUKER, J. S. 2009. Assessment of Artifacts and Reproducibility across Spectral- and Time-Domain Optical Coherence Tomography Devices. *Ophthalmology,* 116**,** 1960-1970.

HOCKING, R. R. 1976. The analysis and selection of variables in linear regression. *Biometrics,* 32**,** 1-49.

HOOD, D. C., ANDERSON, S. C., WALL, M. & KARDON, R. H. 2007. Structure versus Function in Glaucoma: An Application of a Linear Model. *Invest Ophthalmol Vis Sci,* 48**,** 3662-3668.

HOOD, D. C., FORTUNE, B., ARTHUR, S. N., XING, D., SALANT, J. A., RITCH, R. & LIEBMANN, J. M. 2008. Blood vessel contributions to retinal nerve fiber layer thickness profiles measured with optical coherence tomography. *J Glaucoma,* 17**,** 519-528.

HOOD, D. C. & KARDON, R. H. 2007. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Ret Eye Res,* 26**,** 688-710.

HOOD, D. C., RAZA, A. S., KAY, K. Y., SANDLER, S. F., XIN, D., RITCH, R. & LIEBMANN, J. M. 2009. A comparison of retinal nerve fiber layer (RNFL) thickness obtained with frequency and time domain optical coherence tomography (OCT). *Opt Express,* 17.

HOOVER, A. D., KOUZNETSOVA, V. & GOLDBAUM, M. 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *Medical Imaging, IEEE Transactions on,* 19**,** 203-210.

HUANG, D., SWANSON, E. A., LIN, C. P., SCHUMAN, J. S., STINSON, W. G., CHANG, W., HEE, M. R., FLOTTE, T., GREGORY, K., PULIAFITO, C. A. & FUJIMOTO, J. G. 1991. Optical coherence tomography. *Science,* 254**,** 1178-1181.

HUDSON, C., WILD, J. M. & O'NEILL, E. C. 1994. Fatigue effects during a single session of automated static threshold perimetry. *Invest Ophthalmol Vis Sci,* 35**,** 268-280.

HYMAN, L., HEIJL, A., LESKE, M. C., BENGTSSON, B., YANG, Z. & FOR THE EARLY MANIFEST GLAUCOMA TRIAL GROUP 2010. Natural History of Intraocular Pressure in the Early Manifest Glaucoma Trial: A 6-Year Follow-up. *Arch Ophthalmol,* 128**,** 601-607.

IESTER, M., MIKELBERG, F. S., COURTRIGHT, P. & DRANCE, S. M. 1997. Correlation between the visual field indices and Heidelberg retina tomograph parameters. *J Glaucoma,* 6(2)**,** 78-82.

ISHIKAWA, H., STEIN, D. M., WOLLSTEIN, G., BEATON, S., FUJIMOTO, J. G. & SCHUMAN, J. S. 2005. Macular Segmentation with Optical Coherence Tomography. *Invest Ophthalmol Vis Sci,* 46**,** 2012-2017.

JIAO, S., KNIGHTON, R., HUANG, X., GREGORI, G. & PULIAFITO, C. 2005. Simultaneous acquisition of sectional and fundus ophthalmic images with spectral-domain optical coherence tomography. *Opt Express,* 13**,** 444-452.

JOHNSON, C. A., SAMPLE, P. A., ZANGWILL, L. M., VASILE, C. G., CIOFFI, G. A., LIEBMANN, J. R. & WEINREB, R. N. 2003. Structure and function evaluation (SAFE): II. Comparison of optic disk and visual field characteristics. *Am J Ophthalmol,* 135**,** 148-154.

JOHNSON, C. A. & SAMUELS, S. J. 1997. Screening for glaucomatous visual field loss with frequency-doubling perimetry. *Invest Ophthalmol Vis Sci,* 38**,** 413-425.

JOLLIFFE, I. T. 1986. *Principal Component Analysis*, Springer Press.

JONAS, J. B. & GRUNDLER, A. E. 1997. Correlation between mean visual field loss and morphometric optic disk variables in the open-angle glaucomas. *Am J Ophthalmol,* 124**,** 488-497.

JONAS, J. B., GUSEK, G. C. & NAUMANN, G. O. 1988. Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes. *Invest Ophthalmol Vis Sci,* 29, 1151-1158.

KATZ, J., TIELSCH, J. M., QUIGLEY, H. A. & SOMMER, A. 1995. Automated perimetry detects visual field loss before manual Goldmann perimetry. *Ophthalmology,* 102, 21-6.

KELTNER, J. L., JOHNSON, C. A., CELLO, K. E., EDWARDS, M. A., BANDERMANN, S. E., KASS, M. A. & GORDON, M. O. 2003. Classification of visual field abnormalities in the ocular hypertension treatment study. *Arch Ophthalmol,* 121, 643-650.

KIM, J. S., ISHIKAWA, H., GABRIELE, M. L., WOLLSTEIN, G., BILONICK, R. A., KAGEMANN, L., FUJIMOTO, J. G. & SCHUMAN, J. S. 2010. Retinal Nerve Fiber Layer Thickness Measurement Comparability between Time Domain Optical Coherence Tomography (OCT) and Spectral Domain OCT. *Invest Ophthalmol Vis Sci,* 51, 896-902.

KIM, J. S., ISHIKAWA, H., SUNG, K. R., XU, J., WOLLSTEIN, G., BILONICK, R. A., GABRIELE, M. L., KAGEMANN, L., DUKER, J. S., FUJIMOTO, J. G. & SCHUMAN, J. S. 2009. Retinal nerve fibre layer thickness measurement reproducibility improved with spectral domain optical coherence tomography. *Br J Ophthalmol,* 93, 1057-1063.

KLEIN, B. E., KLEIN, R., SPONSEL, W. E., FRANKE, T., CANTOR, L. B., MARTONE, J. & MENAGE, M. J. 1992. Prevalence of glaucoma. The Beaver Dam Eye Study. *Ophthalmology,* 99, 1499-1504.

KNIGHTON, R. & HUANG, X. R. 2002a. Analytical methods for scanning laser polarimetry. *Opt Express,* 10, 1179-1189.

KNIGHTON, R. W. & HUANG, X. R. 2002b. Linear birefringence of the central human cornea. *Invest Ophthalmol Vis Sci,* 43, 82-86.

KOOZEKANANI, D., BOYER, K. & ROBERTS, C. 2001. Retinal thickness measurements from optical coherence tomography using a Markov boundary model. *IEEE Trans Med Imaging,* 20, 900-916.

LACHENMAYR, B. J. & VIVELL, P. M. O. 1993. *Perimetry and its Clinical Correlations*, Stuttgart: Georg Thieme Verlag.

LAM, B. S. Y., YONGSHENG, G. & LIEW, A. W. C. 2010. General Retinal Vessel Segmentation Using Regularization-Based Multiconcavity Modeling. *Medical Imaging, IEEE Transactions on,* 29, 1369-1381.

LEE, K., NIEMEIJER, M., GARVIN, M. K., KWON, Y. H., SONKA, M. & ABRAMOFF, M. D. 2010. Segmentation of the optic disc in 3-D OCT scans of the optic nerve head. *IEEE Trans Med Imaging,* 29, 159-68.

LEITGEB, R., HITZENBERGER, C. K. & FERCHER, A. F. 2003. Performance of fourier domain vs. time domain optical coherence tomography. *Opt Express,* 11, 889-894.

LESKE, M. C. 1983. The epidemiology of open-angle glaucoma: a review. *Am J Epidemiol,* 118, 166-191.

LESKE, M. C., CONNELL, A. M., SCHACHAT, A. P. & HYMAN, L. 1994. The Barbados Eye Study. Prevalence of open angle glaucoma. *Arch Ophthalmol,* 112, 821-829.

LESKE, M. C., HEIJL, A., HUSSEIN, M., BENGTSSON, B., HYMAN, L., KOMAROFF, E. & FOR THE EARLY MANIFEST GLAUCOMA TRIAL GROUP 2003. Factors for Glaucoma Progression and the Effect of Treatment: The Early Manifest Glaucoma Trial. *Arch Ophthalmol,* 121, 48-56.

LEUNG, C. K.-S., CHEUNG, C. Y. L., WEINREB, R. N., QIU, K., LIU, S., LI, H., XU, G., FAN, N., PANG, C. P., TSE, K. K. & LAM, D. S. C. 2010. Evaluation of Retinal Nerve Fiber Layer Progression in Glaucoma: A Study on Optical Coherence Tomography Guided Progression Analysis. *Investigative Ophthalmology & Visual Science,* 51, 217-222.

LEUNG, C. K., CHAN, W. M., CHONG, K. K., YUNG, W. H., TANG, K. T., WOO, J. & TSE, K. K. 2005. Comparative study of retinal nerve fiber layer measurement by StratusOCT and GDx VCC, I: correlation analysis in glaucoma. *Invest Ophthalmol Vis Sci,* 46, 3214-20.

LEUNG, C. K., CHEUNG, C. Y., WEINREB, R. N., QIU, Q., LIU, S., LI, H., XU, G., FAN, N., HUANG, L., PANG, C. P. & LAM, D. S. 2009. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. *Ophthalmology,* 116, 1257-1263.e2.

LI, D., WINFIELD, D. & PARKHURST, D. J. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. IEEE Vision for Human-Computer Interaction Workshop at CVPR, 2005. 1-8.

LI, K., WU, X., CHEN, D. Z. & SONKA, M. 2006. Optimal Surface Segmentation in Volumetric Images—A Graph-Theoretic Approach. *IEEE Trans Pattern Anal Machine Intell,* 28, 119-134.

LIBBY, R. T., GOULD, D. B., ANDERSON, M. G. & JOHN, S. W. 2005. Complex Genetics of Glaucoma Susceptibility. *Annu Rev Genomics Hum Genet,* 6, 15-44.

MACKINNON, M. J. & GLICK, N. 1999. Applications: Data Mining and Knowledge Discovery in Databases – An Overview. *Australian & New Zealand Journal of Statistics,* 41, 255-275.

MAI, T. A., REUS, N. J. & LEMIJ, H. G. 2007. Structure-Function Relationship Is Stronger with Enhanced Corneal Compensation than with Variable Corneal Compensation in Scanning Laser Polarimetry. *Invest Ophthalmol Vis Sci,* 48, 1651-1658.

MCNAUGHT, A. I., CRABB, D. P., FITZKE, F. W. & HITCHINGS, R. A. 1995. Modelling series of visual fields to detect progression in normal-tension glaucoma. *Graefes Arch Clin Exp Ophthalmol,* 233, 750-755.

MILLER, A. J. 1984. Selection of subsets of regression variables (with discussion*). J Roy Stat Soc A Sta,* 147, 389-425.

MINCKLER, D. S. 1980. The organization of nerve fiber bundles in the primate optic nerve head. *Arch Ophthalmol,* 98, 1630-1636.

MISHRA, A., WONG, A., BIZHEVA, K. & CLAUSI, D. A. 2009. Intra-retinal layer segmentation in optical coherence tomography images. *Opt Express,* 17, 23719-23728.

MITCHELL, P., SMITH, W., ATTEBO, K. & HEALEY, P. R. 1996. Prevalence of openangle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology,* 103, 1661-1669.

MITCHELL, T. M. 1997. *Machine Learning,* New York, McGraw-Hill.

MUJAT, M., CHAN, R., CENSE, B., PARK, B., JOO, C., AKKIN, T., CHEN, T. & BOER, J. D. 2005. Retinal nerve fiber layer thickness map determined from optical coherence tomography images. *Opt Express,* 13, 9480-9491.

MWANZA, J. C., CHANG, R. T., BUDENZ, D. L., DURBIN, M. K., GENDY, M. G., SHI, W. & FEUER, W. J. 2010. Reproducibility of Peripapillary Retinal Nerve Fiber Layer Thickness and Optic Nerve Head Parameters Measured with CirrusTM HD-OCT in Glaucomatous Eyes. *Invest. Ophthalmol. Vis. Sci.*, Publish Ahead of Print.

NASSIF, N., CENSE, B., PARK, B. H., YUN, S. H., CHEN, T. C., BOUMA, B. E., TEARNEY, G. J. & BOER, J. F. D. 2004. In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography. *Opt Lett,* 29, 480-482.

NIEMEIJER, M., GARVIN, M., GINNEKEN, B. V., SONKA, M. & ABRAMOFF, M. 2008. Vessel segmentation in 3D spectral OCT scans of the retina. *Proc. SPIE,* 6914.

ORR, M. J. L. 1996. Introduction to Radial Basis Function Networks.

PARK, S., PARK, M. & KANG, M. 2003. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine,* 20, 21-36.

PATEL, P. J., CHEN, F. K., CRUZ, L. & TUFAIL, A. 2009. Segmentation Error in Stratus Optical Coherence Tomography for Neovascular Age-Related Macular Degeneration. *Invest Ophthalmol Vis Sci,* 50, 399-404.

PAUNESCU, L. A., SCHUMAN, J. S., PRICE, L. L., STARK, P. C., BEATON, S., ISHIKAWA, H., WOLLSTEIN, G. & FUJIMOTO, J. G. 2004. Reproducibility of nerve fiber thickness, macular thickness, and optic nerve head measurements using StratusOCT. *Invest Ophthalmol Vis Sci,* 45, 1716-1724.

POGGIO, T. & GIROSI, F. 1990. Networks for approximation and learning. *Proceedings of the IEEE,* 78, 1481-1497.

POHJANPELTO, P. E. & PALVA, J. 1974. Ocular hypertension and glaucomatous optic nerve damage. *Acta Ophthalmol,* 52**,** 194-200.

POLAK, E. 1971. *Computational methods in optimization: a unified approach,* New York, Academic Press.

POWELL, M. J. D. Radial basis functions for multivariate interpolation: A review. *In:* MASON, J. C. & COX, M. G., eds. Algorithms for the approximation of functions and data, 1987. Clarendon Press, 143-167.

QIN, J. & LEUNG, D. H. 2005. A semiparametric two-component "compound" mixture model and its application to estimating malaria attributable fractions. *Biometrics,* 61**,** 456-64.

QUELLEC, G., LEE, K., DOLEJSI, M., GARVIN, M. K., ABRAMOFF, M. D. & SONKA, M. 2010. Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula. *IEEE Trans Med Imaging,* 29**,** 1321-30.

QUIGLEY, H. A., ADDICKS, E. M. & GREEN, W. R. 1982. Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papill-edema, and toxic neuropathy. *Arch Ophthalmol,* 100**,** 135–146.

QUIGLEY, H. A. & BROMAN, A. T. 2006. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology,* 90**,** 262-267.

RADIUS, R. L. & ANDERSON, D. R. 1979. The course of axons through the retina and optic nerve head. *Arch Ophthalmol,* 97**,** 1154-1158.

REUS, N. J., COLEN, T. P. & LEMIJ, H. G. 2003. Visualization of localized retinal nerve fiber layer defects with the GDx with individualized and with fixed compensation of anterior segment birefringence. *Ophthalmology,* 110**,** 1512-1516.

REUS, N. J. & LEMIJ, H. G. 2005. Relationships between Standard Automated Perimetry, HRT Confocal Scanning Laser Ophthalmoscopy, and GDx VCC Scanning Laser Polarimetry. *Invest Ophthalmol Vis Sci,* 46**,** 4182-4188.

SAMPSON, D. D. & HILLMAN, T. R. Optical coherence tomography. *In:* PALUMBO, G. & PRATESI, R., eds. Lasers and Current Optical Techniques in Biology, 2004 Cambridge, UK. 481-571.

SCHLOTTMANN, P. G., CILLA, S. D., GREENFIELD, D. S., CAPRIOLI, J. & GARWAY-HEATH, D. F. 2004. Relationship between Visual Field Sensitivity and Retinal Nerve Fiber Layer Thickness as Measured by Scanning Laser Polarimetry. *Invest Ophthalmol Vis Sci,* 45**,** 1823-1829.

SCHUMAN, J. S., HEE, M. R., ARYA, A. V., PEDUT-KLOIZMAN, T., APULIAFITO, C., FUJIMOTO, J. G. & SWANSON, E. A. 2005. Optical coherence tomography: a new tool for glaucoma diagnosis. *Curr Opin Ophthalmol,* 6**,** 89-95.

SCHUMAN, J. S., PEDUT-KLOIZMAN, T., PAKTER, H., WANG, N., GUEDES, V., HUANG, L., PIEROTH, L., SCOTT, W., HEE, M. R., FUJIMOTO, J. G., ISHIKAWA, H., BILONICK, R. A., KAGEMANN, L. & WOLLSTEIN, G. 2007. Optical Coherence Tomography and Histologic Measurements of Nerve Fiber Layer Thickness in Normal and Glaucomatous Monkey Eyes. *Invest Ophthalmol Vis Sci,* 48**,** 3645-3654.

SEHI, M., GREWAL, D. S., SHEETS, C. W. & GREENFIELD, D. S. 2009. Diagnostic ability of Fourier-domain vs time-domain optical coherence tomography for glaucoma detection. *Am J Ophthalmol,* 148**,** 597-605.

SEHI, M., GREWAL, D. S., ZHU, H., FEUER, W. J. & GREENFIELD, D. S. 2010. Quantification of change in axonal birefringence following surgical reduction in IOP. *Ophthalmic Surg Lasers Imaging,* 42**,** 45-52.

SHAARAWY, T., SHERWOOD, M. B., HITCHINGS, R. A. & CROWSTON, J. G. (eds.) 2009. *Glaucoma,* Edinburgh: Elsevier.

SHARMA, A. K., GOLDBERG, I., GRAHAM, S. L. & MOHSIN, M. 2000. Comparison of the Humphrey Swedish interactive thresholding algorithm (SITA) and full threshold strategies. *J Glaucoma,* 9**,** 20-27.

SHIOSE, Y., KITAZAWA, Y., TSUKAHARA, S., AKAMATSU, T., MIZOKAMI, K., FUTA, R., KATSUSHIMA, H. & KOSAKI, H. 1991. Epidemiology of glaucoma in Japan-a nationwide glaucoma survey. *Jpn J Ophthalmol,* 35**,** 133-55.

SHIRATO, S., INOUE, R., FUKUSHIMA, K. & SUZUKI, Y. 1999. Clinical evaluation of SITA: a new family of perimetric testing strategies. *Graefe's Archive for Clinical and Experimental Ophthalmology,* 237**,** 29-34.

SMITH, S. D., KATZ, J. & QUIGLEY, H. A. 1997. Effect of cataract extraction on the results of automated perimetry in glaucoma. *Arch Ophthalmol,* 115**,** 1515-1519.

SOARES, J. V. B., LEANDRO, J. J. G., CESAR, R. M., JELINEK, H. F. & CREE, M. J. 2006. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *Medical Imaging, IEEE Transactions on,* 25**,** 1214-1222.

SOMMER, A., TIELSCH, J. M., KATZ, J., QUIGLEY, H. A., GOTTSCH, J. D., JAVITT, J. & SINGH, K. 1991. Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey. *Arch Ophthalmol,* 109**,** 1090-1095.

SPENCELEY, S. E., HENSON, D. B. & BULL, D. R. 2007. Visual field analysis using artificial neural networks. *Ophthalmic and Physiological Optics,* 14**,** 239-248.

SPRY, P. G. D., JOHNSON, C. A., BATES, A. B., TURPIN, A. & CHAUHAN, B. C. 2002. Spatial and temporal processing of threshold data for detection of progressive glaucomatous visual field loss. *Arch Ophthalmol,* 120**,** 173-180.

STAAL, J., ABRAMOFF, M. D., NIEMEIJER, M., VIERGEVER, M. A. & VAN GINNEKEN, B. 2004. Ridge-based vessel segmentation in color images of the retina. *Medical Imaging, IEEE Transactions on,* 23**,** 501-509.

STROUTHIDIS, N. G. & GARWAY-HEATH, D. F. 2008. New developments in Heidelberg retina tomograph for glaucoma. *Curr Opin Ophthalmol,* 19**,** 141-8.

STROUTHIDIS, N. G. & GARWAY-HEATH, D. F. 2009. Detecting Glaucoma Progression by Imaging. *In:* GREHN, F. & STAMPER, R. (eds.) *Glaucoma.* Springer Berlin Heidelberg.

SUNG, K. R., KIM, J. S., WOLLSTEIN, G., FOLIO, L., KOOK, M. S. & SCHUMAN, J. S. 2010. Imaging of the retinal nerve fibre layer with spectral domain optical coherence tomography for glaucoma diagnosis. *Br J Ophthalmol,* Publish Ahead of Print.

SWINDALE, N. V., STJEPANOVIC, G., CHIN, A. & MIKELBERG, F. S. 2000. Automated Analysis of Normal and Glaucomatous Optic Nerve Head Topography Images. *Invest Ophthalmol Vis Sci,* 41**,** 1730-1742.

TANG, L., KWON, Y. H., ALWARD, W. L. M., GREENLEE, E. C., LEE, K., GARVIN, M. K. & ABRAMOFF, M. D. 3D reconstruction of the optic nerve head using stereo fundus images for computer-aided diagnosis of glaucoma. *In:* NICO, K. & RONALD, M. S., eds., 2010. SPIE, 76243D.

TATE, G. W. & LYNN, J. R. 1977. *Principles of Quantitative Perimetry,* New York, Grune and Stratton.

TEESALU, P., VIHANNINJOKI, K., AIRAKSINEN, P., TUULONEN, A. & LAARA, E. 1997. Correlation of blue-on-yellow visual fields with scanning confocal laser optic disc measurements. *Invest Ophthalmol Vis Sci,* 38**,** 2452-2459.

THAYANANTHAN, A., NAVARATNAM, R., STENGER, B., TORR, P. H. S. & CIPOLLA, R. Multivariate Relevance Vector Machines for Tracking. Proceedings of the 9th European Conference on Computer Vision (ECCV '06), 2006 Graz, Austria. 124-138.

THYLEFORS, B., NEGREL, A. D., PARARAJASEGARAM, R. & DADZIE, K. Y. 1995. Global data on blindness. *Bull World Health Organ,* 73**,** 115-21.

TIELSCH, J. M., KATZ, J., SOMMER, A., QUIGLEY, H. A. & JAVITT, J. C. 1995. Hypertension, perfusion pressure, and primary open-angle glaucoma. A population-based assessment. *Arch Ophthalmol,* 113**,** 216-221.

TIELSCH, J. M., SOMMER, A., KATZ, J., ROYALL, R. M., QUIGLEY, H. A. & JAVITT, J. 1991. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *The journal of the American Medical Association,* 266**,** 369-374.

TIPPING, M. E. The Relevance Vector Machine. *In:* SOLLA, S. A., LEEN, T. K. & MULLER, K. R., eds. Advances in Neural Information Processing Systems, 2000. MIT Press, 652-658.

TÓTH, M. & HOLLÓ, G. 2005. Enhanced corneal compensation for scanning laser polarimetry on eyes with atypical polarisation pattern. *British Journal of Ophthalmology,* 89**,** 1139-1142.

TRAQUAIR, H. M. 1931. Perimetry in the study of glaucoma. *Trans Ophthalmol Soc UK,* 51**,** 585-599.

TUULONEN, A. & AIRAKSINEN, P. J. 1991. Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression. *Am J Ophthalmol,* 111**,** 485-90.

UCHIDA, H., BRIGATTI, L. & CAPRIOLI, J. 1996. Detection of structural damage from glaucoma with confocal laser image analysis. *Invest Ophthalmol Vis Sci,* 37**,** 2393-2401.

VELTHOVEN, M. E. J., LINDEN, M. H., SMET, M. D., FABER, D. J. & VERBRAAK, F. D. 2006. Influence of cataract on optical coherence tomography image quality and retinal thickness. *Br J Ophthalmol,* 90**,** 1259-1262.

VELTHOVEN, M. E. V., FABER, D. J., VERBRAAK, F. D., LEEUWEN, T. G. V. & SMET, M. D. D. 2007. Recent developments in optical coherence tomography for imaging the retina. *Prog. Retin. Eye Res.,* 6**,** 57-77.

VILLAIN, M. A. & GREENFIELD, D. S. 2003. Peripapillary nerve fiber layer thickness measurement reproducibility using optical coherence tomography. *Ophthalmic Surg Lasers Imaging,* 34**,** 33-37.

VIZZERI, G., BOWD, C., MEDEIROS, F. A., WEINREB, R. N. & ZANGWILL, L. M. 2008. Effect of improper scan alignment on retinal nerve fiber layer thickness measurements using Stratus optical coherence tomograph. *J Glaucoma,* 17**,** 341-349.

VIZZERI, G., BOWD, C., MEDEIROS, F. A., WEINREB, R. N. & ZANGWILL, L. M. 2009a. Effect of signal strength and improper alignment on the variability of stratus optical coherence tomography retinal nerve fiber layer thickness measurements. *Am J Ophthalmol,* 148**,** 249-255 e1.

VIZZERI, G., WEINREB, R. N., GONZALEZ-GARCIA, A. O., BOWD, C., MEDEIROS, F. A., SAMPLE, P. A. & ZANGWILL, L. M. 2009b. Agreement between spectral-domain and time-domain OCT for measuring RNFL thickness. *British Journal of Ophthalmology,* 93**,** 775-781.

WALL, M., PUNKE, S. G., STICKNEY, T. L., BRITO, C. F., WITHROW, K. R. & KARDON, R. H. 2001. SITA Standard in Optic Neuropathies and Hemianopias: A Comparison with Full Threshold Testing. *Invest Ophthalmol Vis Sci,* 42**,** 528-537.

WANG, X. Y., HUYNH, S. C., BURLUTSKY, G., IP, J., STAPLETON, F. & MITCHELL, P. 2007. Reproducibility of and effect of magnification on optical coherence tomography measurements in children. *Am J Ophthalmol,* 143**,** 484-488.

WEBER, J., DANNHEIM, F. & DANNHEIM, D. 1990. The topographical relationship between optic disc and visual field in glaucoma. *Acta Ophthalmol,* 68**,** 568-574.

WEINREB, R. N., DREHER, A. W., COLEMAN, A., QUIGLEY, H., SHAW, B. & REITER, K. 1990. Histopathologic Validation of Fourier-Ellipsometry Measurements of Retinal Nerve Fiber Layer Thickness. *Arch Ophthalmol,* 108**,** 557-560.

WEINREB, R. N. & KHAW, P. T. 2004. Primary open-angle glaucoma. *Lancet,* 363**,** 1711-1720.

WEINREB, R. N., SHAKIBA, S., SAMPLE, P. A., SHAHROKNI, S., S, V. H., GARDEN, V. S., ASAWAPHUREEKORN, S. & ZANGWILL, L. 1995. Association between quantitative nerve fiber layer measurement and visual field loss in glaucoma. *Am J Ophthalmol,* 120**,** 732-738.

WERNER, E. B. 1991. *Manual of visual fields,* New York, Churchill Livingstone.

WILD, J. M., MOSS, I. D., WHITAKER, D. & O'NEILL, E. C. 1995. The statistical interpretation of blue-on-yellow visual field loss. *Invest Ophthalmol Vis Sci,* 36**,** 1398-1410.

WILD, J. M., PACEY, I. E., HANCOCK, S. A. & CUNLIFFE, I. A. 1999. Between-algorithm, between-individual differences in normal perimetric sensitivity: full threshold, FASTPAC, and SITA: Swedish Interactive Threshold Algorithm. *Invest Ophthalmol Vis Sci,* 40**,** 1152-1161.

WILLIAMS, C. K. I. Regression with Gaussian Processes. *In:* ELLACOTT, S. W., MASON, J. C. & ANDERSON, I. J., eds. Mathematics of Neural Networks: Models, Algorithms and Applications, 1995 Oxford, UK. Kluwer.

WOJTKOWSKI, M., LEITGEB, R., KOWALCZYK, A., BAJRASZEWSKI, T. & FERCHER, A. F. 2002. In vivo human retinal imaging by Fourier domain optical coherence tomography. *J Biomed Opt,* 7**,** 457-463.

WOOD, J. M., WILD, J. M., HUSSEY, M. K. & CREWS, S. J. 1987. Serial examination of the normal visual field using Octopus automated projection perimetry evidence for a learning effect. *Acta Ophthalmol,* 65**,** 326-333.

WYATT, H. J., DUL, M. W. & SWANSON, W. H. 2007. Variability of visual field measurements is correlated with the gradient of visual sensitivity. *Vision Research,* 47**,** 925-936.

XIAOYI, J. & MOJON, D. 2003. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* 25**,** 131-137.

XU, J. & CHUTATAPE, O. 2004. 3-D optic disk reconstruction from low-resolution stereo images based on combined registrations and sub-pixel matching. *Conf Proc IEEE Eng Med Biol Soc,* 3**,** 1825-8.

YANAGISAWA, M., TOMIDOKORO, A., SAITO, H., MAYAMA, C., AIHARA, M., TOMITA, G., SHOJI, N. & ARAIE, M. 2009. Atypical retardation pattern in measurements of scanning laser polarimetry and its relating factors. *Eye,* 23**,** 1796–1801.

ZANA, F. & KLEIN, J. C. 2001. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *Image Processing, IEEE Transactions on,* 10**,** 1010-1019.

ZHOU, Q. & KNIGHTON, R. W. 1997. Light scattering and form birefringence of parallel cylindrical arrays that represent cellular organelles of the retinal nerve fiber layer. *Appl Opt,* 36**,** 2273-2285.

ZHOU, Q., REED, J., BETTS, R. W., TROST, P. K., LO, P., WALLACE, C., BIENIAS, R. H., LI, G., WINNICK, R., PAPWORTH, W. A. & SINAI, M. Detection of glaucomatous retinal nerve fiber layer damage by scanning laser polarimetry with variable corneal compensation. *In:* MANNS, F., SÖDERBERG, P. G. & HO, A., eds. Ophthalmic Technologies XIII, 2003. 32-41.

ZHU, H., CRABB, D. P., FREDETTE, M.-J., ANDERSON, D. R. & GARWAY-HEATH, D. F. 2011a. Quantifying Discordance Between Structure and Function Measurements in the Clinical Assessment of Glaucoma. *Arch Ophthalmol***,** Accepted. In press.

ZHU, H., CRABB, D. P. & GARWAY-HEATH, D. F. A Bayesian Radial Basis Function Model to Link Retinal Structure and Visual Function in Glaucoma. The 3rd International Conference of Bioinformatics and Biomedical Engineering, June 2009 Beijing, China. IEEE, 1-4.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G. & GARWAY-HEATH, D. F. 2007. Aligning Sequential Stratus OCT RNFL Scans - Solving the Problem. *Invest Ophthalmol Vis Sci,* 48**,** 516.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., HO, T. & GARWAY-HEATH, D. F. 2010a. *FloatingCanvas*: quantification of 3D retinal structures from spectral-domain optical coherence tomography. *Opt Express,* 18**,** 24595-24610.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., LEMIJ, H., REUS, N. J., HEALEY, P. R., MITCHELL, P., HO, T. & GARWAY-HEATH, D. F. 2010b. Predicting Visual Function from the Measurements of Retinal Nerve Fibre Layer Structure. *Invest Ophthalmol Vis Sci,* 51**,** 5657-5666.

ZHU, H., CRABB, D. P., SCHLOTTMANN, P. G., WOLLSTEIN, G. & GARWAY-HEATH, D. F. 2011b. Aligning Scan Acquisition Circles in Optical Coherence Tomography Images of The Retinal Nerve Fibre Layer. *IEEE Trans Med Imaging***,** Accepted. In Press. DOI 10.1109/TMI.2011.2109962.

ZIO, M. D., GUARNERA, U. & ROCCI, R. 2007. A mixture of mixture models for a classification problem: The unity measure error. *Computational Statistics & Data Analysis,* 51**,** 2573-2585.