ANALYSIS OF RETINAL IMAGES IN GLAUCOMA

ANDREW JAMES PATTERSON

A thesis submitted in partial fulfilment of the requirements of The Nottingham Trent University for the degree of Doctor of Philosophy

Collaborative Institute: Glaucoma Research Unit, Moorfields Eye Hospital, London

March 2006

Declaration

This thesis has been completed solely by the candidate, Andrew James Patterson. The work contained within was done by the candidate.

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

All sources of information have been acknowledged and references given.

Title: Analysis of Retinal Images in Glaucoma

Author: Andrew James Patterson (The Nottingham Trent University)

This thesis is submitted for the degree of Doctor of Philosophy (PhD)

Abstract. Glaucoma is a leading cause of visual disability. Confocal scanning laser tomography (CSLT) yields reproducible three-dimensional images of the optic nerve head and is widely used in the assessment of the disease. The real promise of this technology may be in evaluating progressive structural deterioration in the optic nerve head (ONH) associated with glaucoma over a patient's follow-up. This might be possible as the measurements from the technology have been shown to be sufficiently reproducible. The purpose of this thesis is twofold: to investigate statistical techniques for detecting progressive structural glaucomatous damage; and to investigate techniques which improve the repeatability of images obtained from the technology. Proven quantitative techniques, collectively referred to as *statistic image mapping* (SIM) are widely used in neuro-imaging. In this thesis some of these techniques are adapted and applied to series of ONH images. The pixel by pixel analysis of topographic height over time yields a 'change map' flagging areas and intensity of active change in series of ONH images. The technique is compared to the Topographic Change Analysis (TCA superpixel analysis) and to change in summary measures of the three-dimensional ONH ('stereometric parameters'). The comparisons are made using a novel computer simulation developed in this thesis and further tested on clinical data. A false-positive rate was recorded using test-retest data obtained from 74 patients with ocular hypertension (OHT) or glaucoma. A true-positive rate was estimated using a longitudinal dataset of 52 OHT patients classified as having progressed by visual fields during follow-up. Maximum Likelihood (ML) deconvolution is an image processing technique which estimates the original scene from a degraded image using maximum likelihood probability. This technique has been used in other confocal applications to remove 'out-of-focus' haze and noise in 3D confocal data. In this thesis the approach is applied to test-retest series to evaluate if the technique improves the repeatability of image series. Computer simulation indicated that SIM has better diagnostic precision than TCA in

detecting change. The stereometric parameter analyses have prohibitively high false-positive rates as compared to SIM. In the longitudinal data SIM detected change significantly earlier than the stereometric parameters ($P<0.001$). ML Deconvolution produced an improvement in both intra- and inter-scan repeatability with particular gains in scans that exhibit poor image quality. The techniques developed in this thesis may prove to have real clinical utility in managing patients with glaucoma.

Contents

List of Figures and Tables

List of Abbreviation

API	Application Programming Interface
C	Cortical
CAG	Closed Angle Glaucoma
CI	Confidence Interval
CSLT	Confocal Scanning Laser Tomography
CSM	Cup Shape Measure
DLL	Dynamic Link Library
fMRI	Functional Magnetic Resonance Imaging
GUI	Graphical User Interface
HRT	Heidelberg Retina Tomograph
HVC	Height Variation Contour
IOP	Intraocular Pressure
LOCS	Lens Opacity Classification System
MFC	Microsoft Foundation Classes
ML	Maximum Likelihood
MPHSD	Mean Pixel Height Standard Deviation
MRI	Magnetic Resonance Imaging
NC	Nuclear Colour
NO	Nuclear Opalescence
NTG	Normal Tension Glaucoma
OHT	Ocular Hypertension
ONH	Optic Nerve Head
PC	Posterior Subcapsular
PET	Positron Emission Tomography
PSF	Point Spread Function
POAG	Primary Open Angle Glaucoma
RA	Rim Area
RGC	Retinal Ganglion Cell
RNFL	Retinal Nerve Fibre Layer
RV	Rim Volume
SIM	Statistic Image Mapping
TCA	Topographic Change Analysis
VF	Visual Field
VTK	Visualization Toolkit

Acknowledgements

I would like to thank my supervisory team which has lead to preparation of this thesis: I am grateful to David Crabb (City University) and David Garway-Heath (Moorfields Eye Hospital (MEH), London), for their support and direction. I would like to thank Nick Strouthidis for his help with the clinical data used in this thesis. I would like to acknowledge that my PhD benefited from being registered as a visiting research fellow at the glaucoma research unit. I am grateful for the direction and support I received at quarterly research meetings in the glaucoma research unit MEH. I would like to thank MEH special trustees for funding a generous bursary which facilitated attending international conferences and publication costs over a three year period. I would also like to acknowledge Paul Artes and Balwantray Chauhan from the Department of Ophthalmology, Halifax University, Nova Scotia, Canada for funding an internship during my PhD and for the hospitality I received during my visit.

I would also especially like to thank my parents for their encouragement and support.

Outline

This thesis will primarily be of interest to a researcher in retinal imaging or clinicians using the technology. This thesis will be of interest to any researcher interested in statistical techniques for detecting change in longitudinal series of imaging data. This thesis applies statistical and image processing techniques, but knowledge of these subject areas is not assumed.

Chapter 1 sets the scene: it introduces glaucoma, its risk factors, prevalence and treatment of the disease. Methods used for evaluating glaucoma are discussed, chiefly measuring pressure within the eye, measuring changes to visual function (visual field) and assessing the optic nerve head. Confocal scanning laser tomography (CSLT), the focus of this thesis, is described; an exemplar of a technique for assessing the ONH and posterior segment of the eye and its role in measuring structural damage synonymous with glaucoma. Current literature describing the use of the technology for measuring glaucoma progression is reviewed. Finally the objectives of this thesis are presented.

Chapter 2 describes a computer simulation which is developed here to test quantitative methods used to detect progressive damage to the optic nerve head. The simulation is designed to mimic 'stable' and 'unstable' patient series. The simulation is used in subsequent chapters 3 and 4. The chapter explains how noise and change is simulated and the assumptions made within the model.

Chapter 3 describes *statistic image mapping* (SIM), a proven technique used in neuroimaging to flag significant areas of activity in three-dimensional images of the brain. SIM provides a 'change map' which identifies areas of activity within the image, and a global probability value of the extent of change. This chapter adapts and applies SIM to longitudinal series of ONH images. SIM is evaluated by comparing it to the Topographic Change Analysis (TCA), a method on the

Heidelberg Retina Tomograph (HRT) software (a commercially available CSLT). The comparison uses extensive computer simulation. It concludes that SIM has a better diagnostic precision in separating 'stable' and 'unstable' patient series than the TCA.

Chapter 4 investigates the global probability value output by SIM to infer significant change. Thus far, the method used to infer significant change makes assumption as regards the nature of glaucomatous damage. This chapter challenges this assumption and introduces 'combining functions', a simple mathematical tool that provides a more flexible mechanism to infer significant change. The assumptions are tested using computer simulation. This chapter concludes that combining functions are better at separating 'stable' and 'unstable' patient series.

Chapter 5 evaluates a range of summary measures of the ONH, called 'stereometric parameters'. The chapter examines these parameters using an 'event analysis' available on the current HRT software; the analysis flags a patient series as changing by making comparison between baseline and the follow-up examinations. The performance of SIM and the 'stereometric parameters' is compared using a 'time to event' analysis. Also, SIM and the parameter analysis are compared in some case studies. This chapter concludes that using summary measures to detect change results in loss of sensitivity. It also shows that 'change maps' as output by SIM provide important information on intensity, spatial extent and location of change which is clinically meaningful.

Chapter 6 applies ML deconvolution to confocal scanning laser tomography. This image processing technique estimates the original scene from a degraded one. This technique has been used in other applications of confocal scanning laser microscopy to remove 'out-of-focus' haze and noise. In this chapter the technique is applied to a test-retest series. This chapter concludes that deconvolution significantly improves the repeatability of ONH images.

Chapter 7 sums up the work in the thesis, noting the novel contributions to the field of work and gives suggestions for future work.

Appendix A – In this thesis quantitative techniques to detect changes in images of the ONH have been developed, evaluated and optimized (chapters 2, 3, 4 and 5). An outcome from this thesis has been the development of a windows based program entitled *SIM*. This appendix is a self-contained tutorial written to enable a user to install SIM, export ONH images from the Heidelberg Eye Explorer software, setup a patient record and import images into SIM. It describes SIM's functionality to check images for alignment, magnification error before processing patients using the automated batch file generator. Export functions are described which output parameters from the combining functions and partial tests (chapter 4).

Appendix B – The objective of this section is to describe in sufficient detail the SIM program to enable a researcher to append or modify the source code. The SIM software was developed using the C++ programming language and benefits from the use of application programming interfaces (API) and libraries which C++ can utilize. The section explains how the SIM source code exploits numerical methods, image processing, visualization and graphical user interfaces, API's and libraries. It documents the files which SIM needs to compile and link, as well as the dynamic link libraries (DLL) required by SIM at run-time. It provides a description of the inputs and outputs used in SIM. The section finally documents limitations and suggest improvements for the current code.

1. Background and aims

This chapter gives an introduction to glaucoma and the clinical need to detect and monitor the disease. Confocal Scanning Laser Tomography, the subject of this thesis, is introduced as a technology which can help detect and monitor the disease. The aims of this thesis are then described.

1.1 Glaucoma

Glaucoma is a group of progressive optic neuropathies that have in common a slow progressive degeneration of retinal ganglion cells (RGC) and their axons, resulting in a distinct appearance to the optic nerve head (ONH), often called 'cupping' (Weinreb and Khaw, 2004). Glaucoma is the third leading cause of blindness, yet amongst those with the disease it is relatively rare to be registered blind according to World Health Organization criteria, as central vision is often preserved until late in the disease despite disabling loss of peripheral vision. This damage is often linked with elevated intraocular pressure. This damage to the ONH causes partial to full loss of the *visual field*, which is the portion of space in which objects are simultaneously visible in the steadily fixating eye (Harrington, 1976). Damage to the visual field is irreversible; however, the loss can be transitory in the early stages of glaucoma. If the condition is untreated the damage to the affected visual field usually worsens and spreads until eventually complete loss of vision can occur.

To understand glaucoma it is important to consider *aqueous humor*, the clear watery fluid that circulates through the anterior chamber. This fluid is not related to tears, or to the dense jelly-like substance called vitreous humor that is contained in the rear chamber. The function of aqueous humor is to nourish the area around the iris and the cornea and it exerts pressure to maintain the shape of the eye. The fluid is continuously produced causing pressure known as *intraocular pressure* (IOP). To maintain an intraocular pressure this inflow is offset by an outflow by drainage between the iris and cornea, primarily (80-90%) through a sponge like substance known as the *trabecular meshwork*, the remaining fluid drains independently

through the uveoscleral pathway. Previously it was believed that glaucoma was always a result of elevated intraocular pressure and definitions for glaucoma historically included this. Ocular hypertension (OHT) is a condition in which IOP is greater than 21 mmHg, which is 2 standard deviations above the mean IOP 15.5 mmHg (Colton and Ederer, 1980). However, Sommer and colleagues (Sommer, Tielsch et al, 1991) showed that only 10% of patients with OHT developed glaucomatous visual field damage, but it did show an increased prevalence of glaucoma with increased IOP. It is now understood that glaucoma can occur in eyes with normal intraocular pressure (<21 mmHg). Thus it is best to understand that IOP is a risk factor for glaucomatous damage and that some eyes are more susceptible to the effects of IOP and sustain damage at a lower level. Thus reducing the IOP remains the focus of glaucoma treatment.

Glaucoma can be broadly categorized as primary open angle glaucoma (POAG), closed angle glaucoma (CAG) or congenital. Glaucoma can also be defined as secondary in which the glaucoma is a result of some other condition perhaps an ocular or orbital disease. Normal tension glaucoma (NTG) is a subdivision of POAG. The definition of NTG is an IOP below 21 mmHg. Outside Japan, 30%+ of newly diagnosed cases are NTG. However, the condition may be underdiagnosed in Western countries because of the nature of case-finding for glaucoma. In Japan NTG is the most prevalent form of Glaucoma (Hitchings, 2000). In CAG the iris is pushed against the trabecular meshwork, sometimes sticking to it, closing off the drainage angle. It may occur suddenly resulting in an immediate rise in pressure (‘acute angle closure’). CAG may account for up to 50% of glaucoma worldwide as it has a higher prevalence amongst Asians. Congenital glaucoma is a rare sub-group of glaucoma typically characterized by malformation of the aqueous drainage route. This thesis has focused on the assessment of POAG.

POAG is the most common form of glaucoma in European and North American populations. To summarise recent studies the prevalence was reported at 1.5-2.4% in Caucasians and 6-8% in Afro-Caribbean’s (Tielsch, Sommer et al, 1991; Klein,

Klein et al, 1992; Coffey, Reidy et al, 1993; Dielemans, Vingerling et al, 1994; Leske, Connell et al, 1994; Mitchell, Smith et al, 1996).

To understand glaucoma first consider how the eye functions: the eye gathers and converts light information into neuronal signals, when light enters the eye it travels to the retina and stimulates cells called cones and rods allowing vision to operate over an enormous range of brightness levels. The rods become active at low levels of luminance while cones are active at high levels. RGCs process signals from the cones and rods before relaying them to the brain via their axons which exit the eye via the ONH. In humans there are over a million RGC. In mammals these cells are guided to the ONH during embryonic development, however in fish and frogs axons continue to develop during adulthood (Oster, Deiner et al, 2004). The centre of the retina (macula) has a higher concentration of RGCs, where vision resolution is better (Bennett and Rabbetts, 1998). The axons of RGCs comprise the innermost layer of the retinal nerve fibre layer. These axons converge on the ONH and exit the eye after traversing the lamina cribrosa (a series of perforated connective tissue layers). The convergence of the axons forms a rim of neural tissue and central depression in the optic disc, known as the cup (Weinreb and Khaw, 2004).

POAG is an optic nerve neuropathy which is characterised by changes to the ONH and the visual field; these changes might be associated with elevated IOP. Typically the IOP is elevated because the tiny channels in the trabecular meshwork become clogged, and the subsequent increase in outflow resistance leads to the need for a higher intraocular pressure to maintain fluid flow through the anterior chamber. While the pathophysiology of glaucomatous nuerodegeneration is not fully understood, it is thought that when IOP exerts a force onto the back of the eye, the IOP indirectly leads to the apoptosis of RGCs. This *mechanical theory* hypothesizes that over time, this force causes ‘strangulation’ of the RGC axons at the ONH. Axons leave the eye at the ONH through a sieve-like structure called the lamina cribrosa. The IOP produces shear forces in the tissues of the lamina cribrosa, obstructing to-and-fro transportation of neurotrophic factors and leading to the death

of the cell (Crawford, Harwerth et al, 2000). Animal models of short-term pressure rise show that as IOP increases, pressure gradients across the lamina cribrosa increase. Histology has shown the structure does not return to its original state when the pressure is reduced (plastic deformation) and the structure becomes less rigid (hypercompliant deformation) when pressure is reapplied (Bellezza, Rintalan et al, 2003). An alternative *vascular theory* hypothesizes that changes within the microcirculation of the ONH capillaries are the cause of the damage. The concurrence of glaucoma and splinter haemorrhages at the ONH supports this theory, as a primary reason or co-factor for increased susceptibility to glaucomatous damage (Bathija, 2000). The result of RGC apoptosis and loss of axons, together with deformation of the lamina cribrosa, is a change in the surface topography of the ONH. The neuroretinal rim decreases in size with concurrent enlargement of the cup. The neuroretinal rim is of specific interest in the evaluation of disease state. However, optic disc and rim size are known to have large inter-individual variation. This physiological variability makes glaucoma identification difficult. Understanding the features of the neuroretinal rim is important for assessment: the rim is usually broadest in the **i**nferior disc region, followed by the **s**uperior, then **n**asal and finally **t**emporal (Jonas and Garway-Heath, 2000). This follows the so-called 'ISNT' rule. Glaucomatous damage to the rim has regional preferences depending on the stage of the disease. It typically starts with loss in the inferotemporal and superotemporal regions, then followed by temporal and lastly in the nasal region (Airaksinen and Drance, 1985; Jonas, Budde et al, 1999). Optic disc haemorrhages and retinal nerve fibre layer defects are also associated with the disease (Jonas and Garway-Heath, 2000). Therefore examination of the ONH has always been of importance in both diagnosis and detection of progressive damage.

Several risk factors predispose an individual to POAG. The risk factors reported highlight POAG to be multi-factorial in nature and it is likely that a combination of factors increase an individual's risk. Early work showed that the higher the presenting IOP, the greater the percentage of patients with optic nerve head damage (Pohjanpelto and Palva, 1974). Leske (Leske, 1983) suggests that overall risk of

developing POAG is five times higher with IOP > 21 mmHg. In a recent large population study of OHT patients, baseline IOP remains a leading risk factor (Gordon, Beiser et al, 2002). Population studies have shown that age is one of the most important risk factors (Tielsch, Sommer et al, 1991; Klein, Klein et al, 1992; Coffey, Reidy et al, 1993; Dielemans, Vingerling et al, 1994; Leske, Connell et al, 1994; Mitchell, Smith et al, 1996). These studies reported prevalence rates four to ten times higher in the oldest age groups compared to the baseline (usually subjects in their forties). Studies into racial risk factors show that being of African, African American or African Caribbean origin put one at a four fold increased risk of developing POAG (Tielsch, Sommer et al, 1991; Klein, Klein et al, 1992; Coffey, Reidy et al, 1993; Dielemans, Vingerling et al, 1994; Leske, Connell et al, 1994; Mitchell, Smith et al, 1996). Little data are available regarding POAG in other racial groups, such as those from the Indian sub-continent, Eastern Europeans or from Hispanic origin. There is little doubt that a positive family history of the disease puts an individual at greater risk, however POAG does not usually exhibit *Mendelian inheritance*. The disease appears to be multifactorial and POAG may represent a collection of clinically indistinguishable disorders (Weinreb and Khaw, 2004). Recent advances in genetics have lead to the mapping of glaucoma genes, however, these genes only account for a small portion of diagnosed glaucoma. A mutation in one of these genes, *myocilin*, is found in 3% of late-onset POAG and in greater proportion of juvenile open angle glaucoma cases (Libby, Gould et al, 2005). Other risk factors for POAG include diabetes (Leske, 1983), while another study reports a relationship between elevated blood pressure and elevated IOP (Tielsch, Katz et al, 1995). A risk factor for NTG, a sub-group of POAG, includes being of Japanese descent (Shiose, Kitazawa et al, 1991). Risk factors for CAG include being of Asian or African descent and the disease has a higher prevalence in women than men (Foster and Johnson, 2000). A complete review of risk factors in glaucoma can be found in (Hitchings, 2000).

Most treatments for glaucoma, including NTG, are aimed at reducing IOP and its fluctuation. Treatment can be categorized as medical, laser and fistulising surgery.

Various medical treatments are available, and their pharmacological actions vary. Beta-blockers inhibit aqueous secretion, cholinergic agents cause ciliary muscle contraction which stretch the trabecular meshwork (Krieglstein, 2000), alpha agonists and topical carbonic anhydrase inhibitors also inhibit aqueous production, and prostaglandin analogues increase outflow through the uveoscleral pathway (Hitchings, 2000). As the actions of the various groups of drugs are different, combinations of these agents can be applied. These treatments have side effects (local to the eye and systemic) and the side effects of the agents apply to any combination. Side effects vary in severity, in one study treatment was linked to respiratory impairment (Diggory and Franks, 1997). Argon laser trabeculoplasty reduces IOP by improving aqueous humor outflow. The treatment applies laser burns, usually to one hemisphere of the circumference of the trabecular meshwork at a time. The treatment is simple and cost-effective but long-term follow-up has shown IOP tends to rise over time in many patients (Schwartz, Love et al, 1985). Fistulising surgery, typified by trabeculectomy, is a standard practice for surgery in adults. Studies have shown it to be more efficient than medical and laser treatments at lowering IOP and in preserving visual function in the long-term (Jay and Allan, 1989; Migdal, Gregory et al, 1994). The procedure creates a passageway for aqueous to escape. The escape route is not directly to the external surface of the eye as this would have the potential for infection. The drainage of fluid from inside the anterior chamber of the eye to a “pocket” created between the conjunctiva, the outermost covering of the eye, and the sclera, the underlying fibrous wall of the eye. In the last decade some clinical trials have reported on the affect of treatment over long term follow-up. The early manifest glaucoma trial (Heijl, Leske et al, 2002) compared the effects of lowering IOP verses no treatment or later treatment, where treatment involved trabeculoplasty plus medication. The study showed treatment significantly delays visual field progression. The advanced glaucoma intervention study (AGIS, 2000) studied the association of visual field deterioration and control of IOP by surgical intervention by both argon laser trabeculoplasty and trabeculectomy. After 6 years of patient follow-up, the study showed lowering pressure reduces progression rates in the visual field. Weinreb and Khaw (Weinreb

and Khaw, 2004) provide a review of methods and outcomes from other clinical trials, including those of OHT patients. These studies support the view that lowering IOP reduces the rates of progression in visual fields and damage to the optic disc.

Diagnostics

Diagnosis of glaucoma no longer relies on the presence of elevated IOP alone, but requires the additional assessment of the visual field and the ONH. However, elevated IOP (together with subject age) remains the most important single prognostic risk. Physicians determine the IOP using *tonometry*. This technology measures the force necessary to applanate or indent the cornea. The instrument can be categorized as contact or non-contact. While non-contact remains the most popular and is often called the 'air-puff' test; contact tonometry has been shown to have better inter-observer agreement (Tonnu, Ho et al, 2005). In normal individuals diurnal variation of IOP varies from 3-6 mmHg. A diurnal variation of >10 mmHg is suggestive of glaucoma, while some glaucomatous eye have reported diurnal IOP fluctuation of >30 mmHg (Newell and Krill, 1964). Cornea thickness may be an important source of error in tonometry (Tonnu, Ho et al, 2005). Thick corneas are linked to high measured IOP and patients with thin corneas may therefore have higher IOP than that reported by tonometry (Yagci, Eksioglu et al, 2005).

Perimetry is a diagnostic technique for measuring the visual field (light sensitivity at various retinal locations). It remains central to monitoring visual function in glaucoma, as it determines what the patient can actually see (Hitchings, 2000). The technique can therefore help addresses important issues such as quality-of-life and fitness-to-drive. Normally automated perimetry measures the central 25-30º visual field and has become a clinical standard. Perimetry, typified by the commercially available Humphrey Field Analyzer [Carl Zeiss Ophthalmic Systems, Dublin, CA], measures the light-difference sensitivity across the visual field. A stimulus of a certain size and intensity is presented at set test positions and the patient indicates when they have seen the stimulus (Haley, 1987; Werner, 1991). Using a full-threshold algorithm, the stimulus intensity is stepped in fixed increments until a final

sensitivity value is recorded. This process is repeated at each test position. An alternative testing strategy, known as SITA, has been introduced to reduce examination time (Bengtsson, Olsson et al, 1997). Perimetry technology detects fixation errors, false positive and false negative events which all give a measure of the reliability of the test. The output from the machine includes a map of the visual field and 'global indices' (summary values) indicating if the field has a low sensitivity or deviates from an age matched normal visual field. Known sources of variability include changes in pupil size, refractive error, ocular media opacities, subject learning, fatigue effects and fixation errors (Henson, 2000).

The ability to detect changes in the ONH morphology in follow-up assessments depends on the reproducibility of the method employed; if the method is highly reproducible then small changes in the disc can be detected. Optic disc drawing from stereoscopic examination has been an important part in examination. Serial drawing of the ONH have been able to demonstrate progressive changes over time. However, few patients are followed by a single clinician and it is known that variation between different observers is large (Garway-Heath, 2000).

Optic disc photography provides a high-resolution permanent record of ONH appearance. Flicker-chronoscopy and stereochronoscopy allows detection of very small changes between two photographs, but a false-impression can be generated by magnification error and parallax (Garway-Heath, 2000). It has been demonstrated that simultaneous and sequential stereoscopic ONH photography is capable of detecting progressive glaucomatous changes (Linner and Stromberg, 1967; Sommer, Pollack et al, 1979; Pederson and Anderson, 1980; Odberg and Riise, 1985). *Planimetry* is the term given to measurements made from photographic images. Some software programs allow viewing of digitised ONH photographs. This therefore enables quantitative assessment of the ONH but is limited by subjective interpretations of the boundaries of the ONH and neuroretinal rim (Garway-Heath, 2000). Scanning laser polarimetry and confocal scanning laser tomography are imaging technologies capable of measuring the posterior segment of the eye.

Scanning laser polarimetry uses a polarized infrared laser with the aim of measuring the thickness of the retinal nerve fibre layer (RNFL). The birefringence properties of RNFL cause a change of state to the polarization of reflected light (known as retardation). However, a light beam emerging from a living eye contains information on the polarization properties of all ocular structures (i.e. cornea, lens, humours and RNFL); the cornea in particular has birefringence properties (Bueno, 2004). To compensate for this, a variable corneal compensator has been developed. This technology shows promise in separating normal and glaucomatous eyes (Reus and Lemij, 2004; Da Pozzo, Iacono et al, 2005). CSLT obtains 3D topographic images of the ONH or other posterior segments of the eye. This technology is introduced in section 1.2.

1.2 Confocal Scanning Laser Tomography

Scanning laser ophthalmoscopy, introduced in the 1980's, developed as an alternative method to image ocular features such as the retina, macular and optic nerve head (Webb and Hughes, 1981). Since its launch, hosts of new applications have spawned, such as scanning laser polarimetry, scanning laser Doppler flowmetry, scanning laser fluorescein angiography, scanning laser corneal microscopy and confocal scanning laser tomography (CSLT). Ciulla and colleagues (Ciulla, Regillo et al, 2003) provide a review of these technologies and their application in ophthalmology. CSLT, typified by the commercially available Heidelberg Retina Tomograph (HRT) [Heidelberg Engineering, Heidelberg, Germany], is the subject of this thesis. What follows is a description of how the HRT acquires images and reviews how the technology is typically used for diagnosis of glaucoma and for detecting progression.

One distinction between digital fundus photography and confocal scanning laser tomography is that the *scanning* illumination system samples the retina point by point rather than capturing the image as a whole. This illumination set-up enables

the device to image eyes through undilated pupils. A low-energy laser is focused on a point on the retina which reflects light back to a detector. A deflector mirror then moves the laser beam horizontally so an adjacent point can be imaged. When one line has been acquired a second deflector mirror moves the beam vertically before acquiring another horizontal line. A 2D image is built up in this raster-like fashion in approximate 32 milliseconds (a total of 256 × 256 pixels), a second CSLT called the HRT II has been developed which acquires a total of 384 × 384 pixels. The image acquisition is based on *confocal optics*, a system in which a pinhole in front of the detector is optically conjugate to the focal plane of the object being imaged (see figure 1(a)). This ensures that only light from the imaged focal plane reaches the image sensor. Reflected light from in front, or behind, the focal plane is prevented from reaching the detector by the pinhole. The position of the focal plane can be moved and during acquisition it is moved from anterior to posterior to obtain a total of 32 two-dimensional images. This 3D (256 × 256 × 32 voxel) image obtained is typically called a confocal stack (see figure 1(b & c)). On the HRT II the number of scans acquired is automated and depends on the depth of the ONH, the number of scans acquired varies from 16 to 64. Images can be prone to artefacts and noise for a number of reasons. The eye tends to lose fixation during the scan, this results in translation and rotation shifts between images in the stack; the proprietary HRT software image alignment algorithms attempt to compensate for this effect. Chauhan and McCormick (Chauhan and McCormick, 1995) studied the effect of the cardiac cycle on images and found it to be a source which increases variability. Zangwill and colleagues (Zangwill, Irak et al, 1997) studied the effect of pupil size and cataract on the reproducibility of the technology. They showed, in the presence of cataract, pupil dilation improved reproducibility, a step typically not performed in a clinical setting, and also reported a significant correlation between both subjective and objective grading of cataract and image quality. Sihota and colleagues (Sihota, Gulati et al, 2002) reported that uncorrected astigmatism and poor visual acuity resulted in images with a higher variability. Orgul and colleagues (Orgul, Croffi et al, 1997) reported that the proprietary HRT inter-image alignment software reduced

the reproducibility of the technology. Chapter 6 examines the utility of the technology in imaging patients with cataracts.

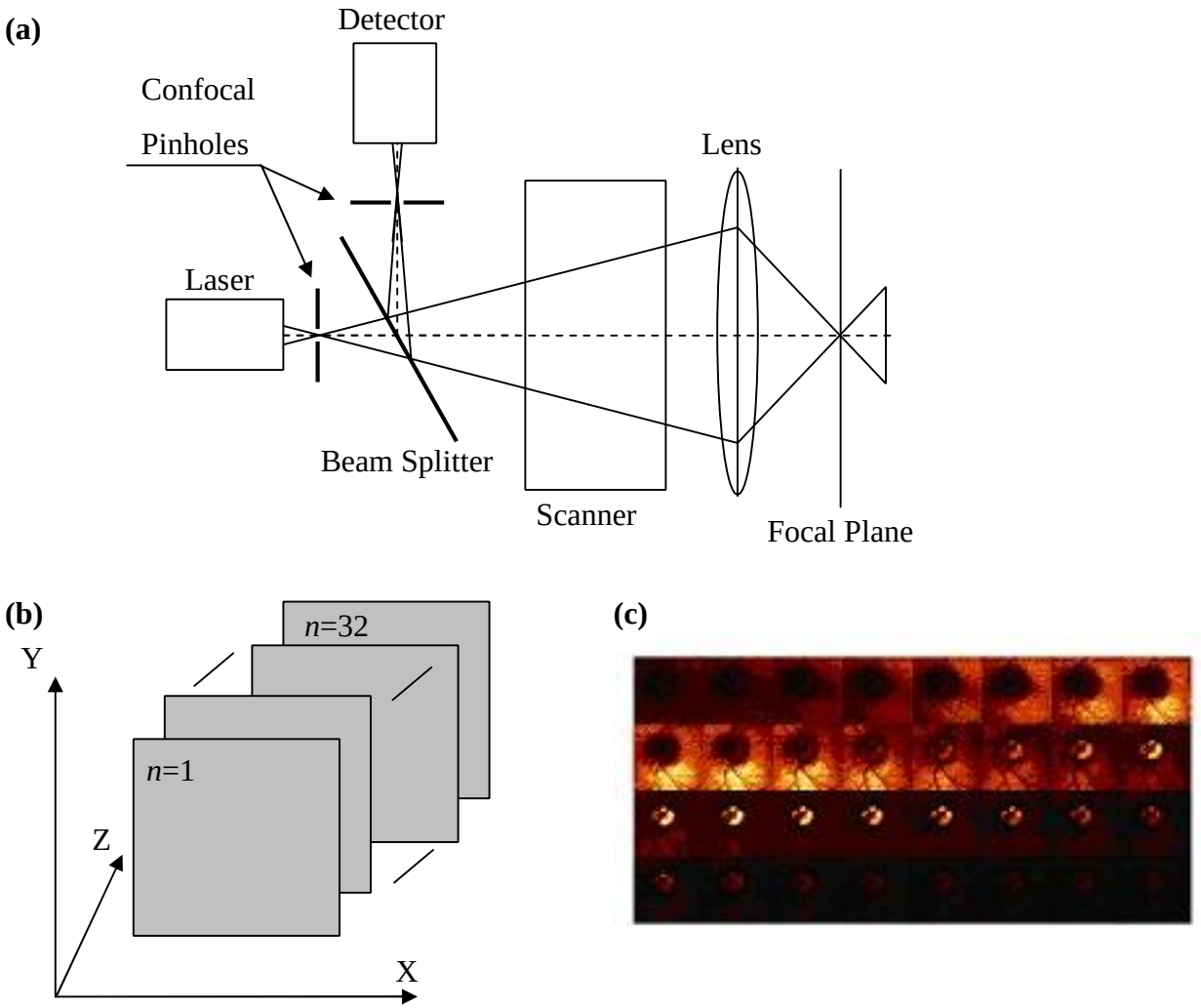


Figure 1 (a) Confocal optical setup (b) A schematic diagram illustrating the 3D confocal stack obtained from a scanning laser tomograph. (c) The 3D confocal stack of an optic nerve head illustrated as an 8 × 4 grid of 2D images going in sequence from top left (*n*=1) to bottom right (*n*=32). Each 2D optical section represents a different focal plane (Courtesy of Heidelberg Engineering, reproduced from the 'HRT tutorial', available at www.heidelbergengineering.com)

A 2D topography image is formed by calculating the position of maximum reflectivity at each z-profile, a two dimensional profile of 32 signal intensity values parallel to the optical axis (see figure 2(a)). The topography image represents the

surface height of the optic nerve head and surrounding papillary retina (see figure 2(b)). Therefore, the intensity at each pixel within the image represents a height in microns. Typically three topography images are acquired at each visit (and this is automated with the HRT II) and in most applications averaged to calculate a mean topography; this became the convention after an early study into the effect of repetitive imaging (Weinreb, Lusky et al, 1993). Image registration algorithms within the HRT software align the topography images for the within visit and between visit differences in scan positions. A further description of the technology is provided by Zinser, Wijnaendts-van-Resandt et al, 1989; Chauhan, 1996. The technology has been shown to obtain reproducible topography images of the ONH (Chauhan, LeBlanc et al, 1994; Rohrschneider, Burk et al, 1994).

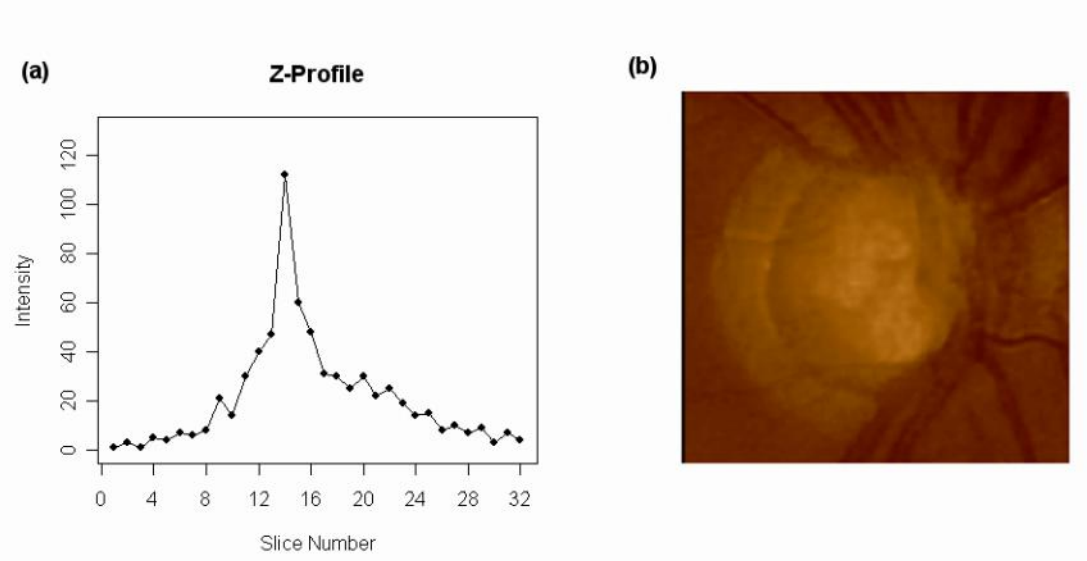


Figure 2 (a) The distribution of light intensity at a signal pixel location (*x*,*y*), referred to as a confocal z-profile. (b) The topography image which consists of 256 × 256 height measurements produced by calculating the position of the reflective surface at each pixel location (*x*,*y*) in the 3D confocal image stack

The repeatability of topography images is typically quantified using mean pixel height standard deviation (MPHSD). This metric is a gauge of the variability of each pixel height measurement across the three topographies used to make up the mean topography (Dreher, Tso et al, 1991). It is calculated from the standard deviations at

each pixel across the mean topographic image, i.e. the MPHSD is the mean of 256 × 256 pixel height standard deviations in HRT I images.

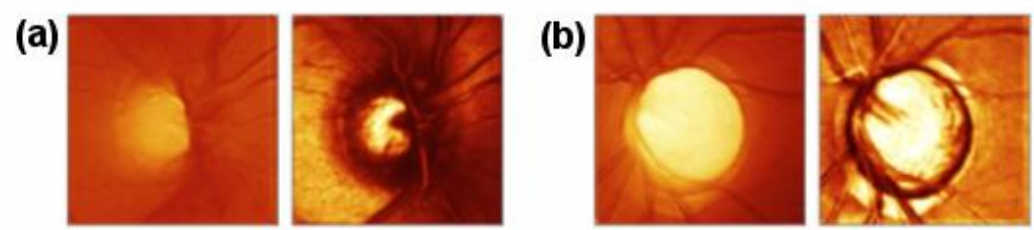


Figure 3 Pair of topography and reflectance images for a normal (a) and glaucomatous (b) eye. (Courtesy of Heidelberg Engineering)

In Figure 3(a) a topography image (left) and reflectance image (right) are shown for a normal subject's right eye. Figure 3(b) shows a topography and reflectance image for a glaucoma patient's left eye. In this example the difference in size and shape of the ONH cup is obvious.

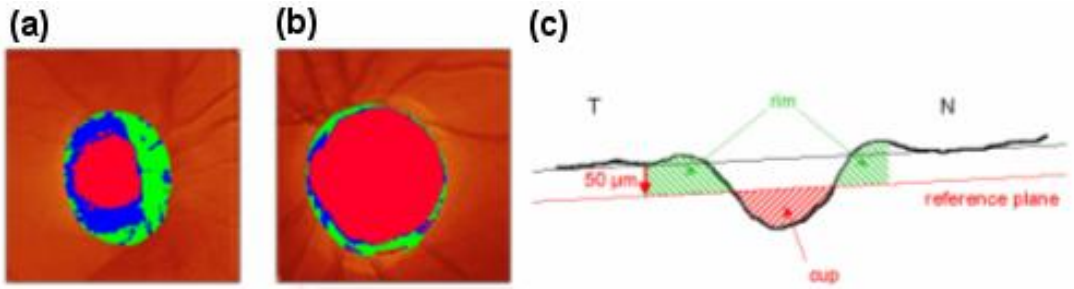


Figure 4 HRT output showing the rim and cup for a normal (a) and glaucomatous (b) eye. The red colour represents cup, while the green and blue represent rim. (c) Shows a two-dimensional section through a topography image. Anything below the reference plane is cup (marked as red), while anything above the reference plane is rim (marked and green and blue). Courtesy of Heidelberg Engineering

The HRT software quantifies structural features of the ONH in glaucoma by calculating a number of 3D descriptive parameters known as *stereometric parameters*. A contour line is first defined. This is a closed elliptical shape drawn manually using a subjective assessment of the location of the boundary of the optic

disc (inner margin of Elschnig's ring). Note that although this input is subjective, its position has been shown to have good inter-operator agreement (Hatch, Flanagan et al, 1999). A "reference plane" (see Figure 4c) is calculated. This is a plane parallel to, and set below, the retinal surface and is used to divide the optic disc into rim and cup (Burk, Vihanninjoki et al, 2000). RA measurements from stereophotographs have been shown to correlate with visual function (Balazsi, Drance et al, 1984) CSLT, RA measures the area bound within the contour line which is above the reference plane. In Figure 4(a&b) RA is simply the sum of the green and blue area. In this example there is a clear difference in RA between the normal and glaucomatous eye. Rim Volume (RV) is a measure of the volume bound within the contour line and above the reference plane (Hatch, Flanagan et al, 1997). Cup Shape Measure (CSM) is a measure of the three-dimensional shape of the cup, also called the third moment. In mathematical terms, a second moment represents variance, the square root of the second moment as known as the standard deviation, and the third moment, here called CSM, represents the skewness of the distribution. A deeply cupped disc will have many outliers and a flat cup will have fewer outliers (Burk, Rohrschneider et al, 1990). Height Variation Contour (HVC) is the retinal surface height variation around the disk contour line (Hatch, Flanagan et al, 1997). Retinal Nerve Fiber Layer (RNFL) thickness, as calculated on the HRT, is the mean distance between the reference plane and the retinal surface (Iester and Mermoud, 2005). As well as yielding a global stereometric parameter, when applicable, the HRT provides values in six predefined segments: temporal, temporal superior, temporal inferior, nasal, nasal superior and nasal inferior.

Stereometric parameters have been used with some success to discriminate between normal optic discs and those with glaucoma. It is worth noting that the known large inter-subject variability of optic disc size, rim size and depth of cupping makes this task non-trivial. A variety of statistical and quantitative techniques applied to the stereometric parameters have been used for this task. Wollstein and colleagues (Wollstein, Garway-Heath et al, 1998) examined the best parameters that separated patients with early glaucoma from normal subjects. They reported the highest

separation using the 99% prediction interval from linear regression between the optic disc area and the log of the RA. Uchida and colleagues (Uchida, Brigatti et al, 1996) applied neural networks to CSM. Linear discriminant analysis combines parameters to achieve separation. Studies have typically used CSM, HVC and RV (Mikelberg, Pafitt et al, 1995; Iester, Mikelberg et al, 1997; Bathija, Zangwill et al, 1998). Another approach divides RA into sectors and computes "ranked sector distribution curves" (Asawaphureekorn, Zangwill et al, 1996; Gundersen and Asman, 2000).

The main subject of this thesis has been to develop statistical techniques to detect glaucomatous progression. As HRT measurements have been shown to be repeatable, it is hoped that progressive glaucomatous damage can be identified by repeated scanning of an individual patient over years of follow-up. In this thesis techniques were developed and tested on HRT I images (but are equally applicable to HRT II images).

1.3 Progression

Why measure optic disc progression? In the management of glaucomatous patients preservation of vision is the principal objective. At the mid-point of the 20th century, the amount of psychological and physical damage to the health of Americans caused by treatment of glaucoma was probably greater than the damage done by glaucoma itself (Hitchings, 2000). This was because treatment was based on a definition of glaucoma based solely on an elevated IOP. It was not until the natural history of OHT had been studied that it became recognized that only 5-10% of patients with ocular hypertension developed visual field loss (Linner and Stromberg, 1967). Therefore, it is considered ideal in the follow-up of OHT and glaucoma patients to combine tonometry, visual field tests and assess damage to ONH. However, merely performing these tests may not be enough. For example, it has been shown that subjective assessment of follow-up series of visual fields by experts has poor agreement in detecting progressive visual field loss (Werner, Bishop et al, 1988; Viswanathan, Crabb et al, 2003). Statistical techniques have been developed to

improve this agreement: a trend analysis known as PROGRESSOR (Fitzke, Hitchings et al, 1996; Viswanathan, Fitzke et al, 1997), and an event analysis known as Statpac 2 (Heijl, Lindergren et al, 1991). Clearly, it would also be worthwhile to develop similar techniques to help clinicians assess ONH damage occurring during follow-up examination. This is the primary objective of this thesis.

The thesis develops statistical techniques to quantify change across a whole image space; to do this some statistical problems need to be accounted for. If, for example, a patient has five HRT images acquired at his/her first visit (called the baseline visit) and after a period of follow-up another five topography images are acquired, the goal is to compare these images and investigate if statistically significant change occurs to the ONH during the time elapsed. To do this, a statistic could be calculated at each pixel location; this would give an image made up of statistic values (known as the univariate statistic image). In this image we would expect that the values with highest intensity would occur at the locations where the greatest structural change has occurred to the ONH. The question is to determine if statistically significant structural change has occurred to the ONH, thereby inferring if significant change occurred across the whole image. To answer this question it is necessary to account for the many thousands of statistic values in the univariate statistic image. For now lets imagine that we are only interested in testing to determine whether change has occurred at one specific pixel location. Does the change at this pixel location provide convincing evidence of structural change? To answer this question, a two-sample t-test could be performed to determine if the average height of the five topographic values from the follow-up visit is significantly less than the average of the five baseline values. Lets say the t-test returns a test statistic of 2.9. The t value is then tested against the null hypothesis, which is the hypothesis that there is no effect. To do this the t value is compared against the *null distribution* of t statistics, which is the distribution of t statistics that would be expected if there was no difference. This mathematically derived distribution tells us that the probability of observing a value greater than 2.9 is 0.01 if there is no effect. In this case we reject the null hypothesis with a 1% risk of *type I* error, which

is the likelihood that the result has in fact arisen when there is in fact no change. The situation is unfortunately more complicated: there are approximately 65 thousand pixels locations in a topography image. If a t-test was performed at each pixel location, is there any evidence to suggest that change has occurred across the whole image? Simply put, this means that approximately 650 t statistics in the image are likely to be greater than 2.9 by chance (0.01 × 65,000 = 650). This is known as the *multiple comparison problem*. To solve this statistical problem, a new threshold is needed so that in an image of 65000 statistics there is only a 1% probability of there being one or more test statistic values above this threshold. One method for solving this problem is to use the *Bonferroni correction*. A Bonferroni correction can be applied to keep the type I error rate of 1% from before. To do this the required probability at a single pixel location will need to be equal to or more extreme than (0.01/65000) 0.00000015. The corresponding test statistic for this probability value is 15.5. If any pixel's test statistic is equal to or greater 15.5, then it can be concluded that this statistic has only a 1% chance of have arisen anywhere in the image by chance. Unfortunately, in many cases in imaging data using the Bonferroni correction for calculating Type I error rates is too conservative. This is because images have some degree of *spatial correlation*; at each pixel there is a correlation between a pixel and its neighbouring pixel values. This means, in the case of topography images, there are fewer independent values than there are pixels in the image. Some degree of spatial correlation is almost universally present in imaging data. This phenomenon is called the point spread function (explained in more detail in chapter 6). Also, spatial smoothing often used to improve the signal to noise ratio by definition increase spatial correlation. Developing a statistical technique which accounts for the multiple comparison and spatial correlation problems is the main subject of this thesis. Brett and colleagues (Brett, Penny et al, 2003) provides a detailed review of the issues of applying statistical techniques to imaging data. What follows in this section is a review of current methodologies for detecting glaucomatous progression using the HRT.

To date, few statistical tools have been developed which use stereometric parameters to detect glaucomatous progression. Studies have measured the test-retest variability of each parameter (Mikelberg, Wijsman et al, 1993; Rohrschneider, Burk et al, 1994); if change exceeds this variability, it is proposed that this represents true morphological change (Kamal, Viswanathan et al, 1999; Kamal, Garway-Heath et al, 2000). Tan and Hitchings (Tan and Hitchings, 2003) developed a technique using 30 degrees segments of rim area and an experimental reference plane. Strouthidis and colleagues (Strouthidis, White et al, 2005) showed rim area to be the most repeatable parameter, both with inter-visit and inter-operator variability. Stereometric parameters are summary measures, i.e. they are quantified by averaging data over parts of the topography image. This is a highly data reductive process and any statistical technique may, by definition, have reduced sensitivity in detecting localised change. However, statistical techniques which use summary measures do not need to account for the spatial correlation and multiple comparison problems which arise when quantifying change at pixel by pixel level.

Chauhan and co-workers were the first to develop a statistical technique which looked for change within the image (Chauhan, Blanchard et al, 2000). The Topographic Change Analysis (TCA), now included in the HRT software, divides the image into a 64 × 64 superpixel array (each superpixel contains 4×4, or 16 pixels). An ANOVA is conducted to measure the extent of a constant shift in the topographic height over all 16 pixels within each superpixel from one set of images (3 replicates at baseline) to another (3 replicates at the follow-up visit, see Figure 5). The significance of change at each superpixel is evaluated using the F-distribution where the degrees of freedom are adjusted via a correction to account for spatial correlation within each superpixel. It is worth highlighting that this correction is used *within* a superpixel and does not correct for the spatial correlation across the whole image. The authors established an empirical criterion for 'significant' change of 20 or more statistically significant superpixels bound within the ONH contour line, and when this topographic change occurred at the same superpixels in 3 consecutive sets of follow-up images, compared with baseline (Chauhan, McCormick et al,

2001). This criterion for change was introduced to set the specificity at a particular level, the limits being derived from empirical data (longitudinal data from normal subjects). In a subsequent publication, three criteria for change where presented to demonstrate the variation in overlap of patients with VF and ONH progression at different levels of criterion stringency: least conservative (6% area of depressed significant superpixels within the contour line), intermediate (10%) and most conservative (18%) (Artes and Chauhan, 2005). The technique therefore has no intrinsic mechanism to account for the multiple comparison problem. The TCA result is also based solely on comparing the three most recent follow-up images with the baseline image, thus detecting change is highly dependant on the quality of the baseline image.

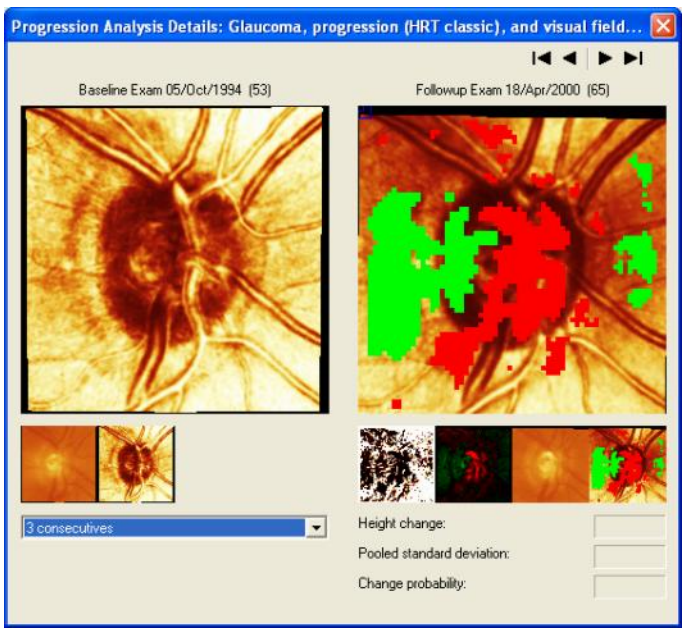


Figure 5 TCA output from the HRT Eye Explorer software (version 1.4.1.0). The red (green) clusters overlaid on the image represent statistically significant depressed (elevated) superpixels which were confirmed as significant in three consecutive visits after comparing the baseline visit with the follow-up visits. (Courtesy of Heidelberg Engineering)

This thesis develops and tests techniques to address the issues of applying standard statistical technique across an image space for detecting glaucomatous damage.

1.4 Objectives

The objectives of this thesis were to improve the analysis of ONH images from the CLST. In particular, this thesis aims to:

- Develop a novel technique for detecting structural change in ONH images.
- Develop a simulation for evaluating the specificity and sensitivity of techniques aimed at detecting glaucoma progression.
- Validate glaucoma progression techniques on simulated and clinical datasets.
- Investigate techniques for improving the repeatability of CSLT ONH images.

2. Simulation of serial optic nerve head (ONH) images

The major aim of this thesis is to evaluate and implement new methodologies for detecting change in the ONH associated with the morphological changes that occur due to glaucoma. The aim is to apply these techniques to serial images of the ONH acquired using confocal scanning laser tomography. In this chapter we discuss the issues of measuring the sensitivity and specificity of techniques aimed at measuring these changes. These issues lead to the necessity of developing a fair, reliable and unbiased testing procedure for comparing methods for detecting change to the ONH.

2.1 Previous work

One 'traditionally' held viewpoint infers that optic disc changes (structural changes) precede the onset of defects in the visual field (functional changes). There is a body of work which supports this theory. Histological results report a reduction in the number of RGC at retinal locations known to have visual field defects (Quigley, Dunkelberger et al, 1989; Kerrigan-Baumrind, Quigley et al, 2000). Quigley and colleagues (Quigley, Dunkelberger et al, 1989) have shown 20% and 40% of RGCs were gone in locations with 5 dB and 10 dB visual field sensitivity loss, respectively. Kerrigan-Baumrind and colleagues (Kerrigan-Baumrind, Quigley et al, 2000) report at least 25% to 35% of RGC loss is associated with statistical significant abnormalities in automated visual field testing. Another study measures RNFL defects from photographs of 1344 eyes using two masked observers (Sommer, Katz et al, 1991). In the sub-sample with visual field defects the most sensitive observer of the two observers identified RNFL defects in 88% of the photographs at the time in which a visual field defect had first occurred. The same observer reports RNFL defects in 11% of normal eyes and 26% of OHT eyes.

With the advent of confocal scanning laser tomography, studies compared structural changes measured at the ONH against changes to the visual field that occur during patient follow-up. Chauhan and colleagues (Chauhan, McCormick et al, 2001) report on a population of 77 glaucoma patients examined at regular intervals. This study

shows 40% progress with ONH change only, while 4% progress with visual field only. However, of the 29% who progressed with both techniques, 45% first progressed with ONH changes and 41% first progressed with visual field changes. Artes and Chauhan (Artes and Chauhan, 2005) later reported that perimetry techniques and scanning laser tomography provide largely independent measures of progression. This work has contested the 'traditionally' held viewpoint and implies that it cannot be assumed that though a patient's visual field has changed to a diagnosis of glaucoma over the course of the patient's follow-up, that the corresponding topography images of the ONH will not necessarily have developed a structural glaucomatous defect. This means that there is no surrogate measure to confirm glaucomatous damage may have occurred to the ONH. This issue is compounded by the wide variability in appearance of the normal optic disc. Medeiros and colleagues (Medeiros, Zangwill et al, 2005) have suggested using traditional examination of the optic disc by expert assessment of serial photographs and drawings as a gold standard to evaluate the diagnostic accuracy of imaging instruments for detecting glaucoma progression. However, previous studies have shown poor inter-expert agreement in diagnosing progression based on drawing or photographs (Coleman, Sommer et al, 1996).

To resolve these issues it is necessary for this thesis to develop a simulation mimicking serial images obtained from 'stable' and 'unstable' patients. To date there has been no published research in this area. However, in perimetry computer simulations were used to mimic series of VFs (Gardiner, 2003). Also an optimized VF testing strategy has been implemented after initial validation on computer simulation (Bengtsson, Olsson et al, 1997). Simulations of ONH images have been used to test the ability of new techniques to separate glaucomatous from normal ONH (Swindale, Stjepanovic et al, 2000; Adler, Hothorn et al, 2004). In this chapter a novel method is presented which simulates both stable and unstable series of topographic images.

2.2 Simulation of ONH images

The simulation models 'stable' and 'unstable' patients. Each longitudinal virtual patient series comprised of 30 topography images. This mimics a patient being scanned over 10 visits with 3 images obtained at each visit. To model both stable and unstable patients it was also necessary to model between-image and within-image topographic variability characteristics of a patient's image series.

Each stable patient series is simulated by simply first creating 30 identical copies of a HRT image. Then noise is applied to each image. Simulating 'unstable' patients proceeded by applying either gradual or episodic topographic change to longitudinal series of images. Unstable patient series' with gradual change (linear) were simulated by creating 30 identical images and then applying change to a cluster of pixels at the cup border in the shape of an irregular polygon. For each image change is applied $(v-1)\times {}^{d}\!/_{n-1}$ in depth, where v is the visit, d is total depressed change and n is the total number of visits. Progressing patient series' with sudden (episodic) change were also simulated by introducing a height decay of $d \times n$ applied to the entire cluster at a randomly selected visit between visit 2 and visit 10 inclusive.

Between-image variability had two elements: 'misalignment noise' and background noise. 'Misalignment noise' was simulated by applying a series of transformations to each image in translations (x', y', z') and rotations about each axis (σ_x', σ_y', σ_z') as illustrated in Figure 6.

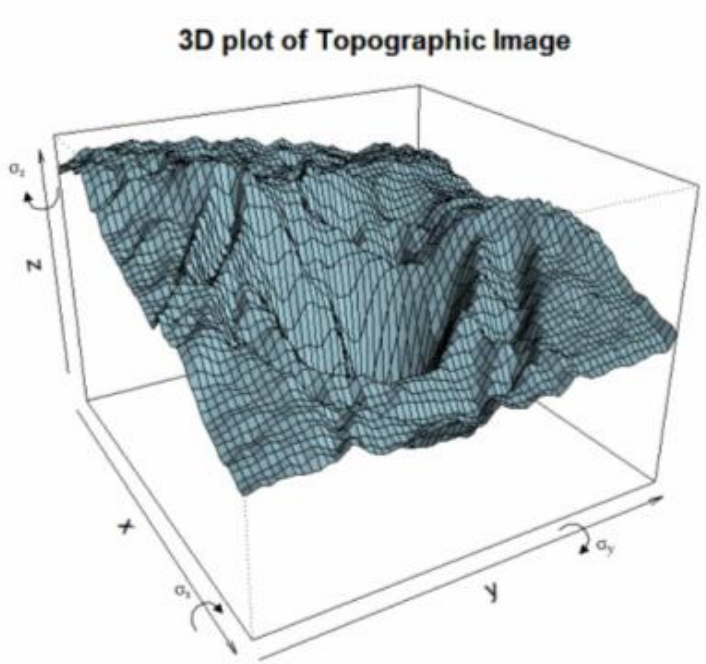


Figure 6 A 3D plot of a topography image showing the transformations *x', y,' z'* and σ_x', σ_y', σ_z'

Geometric transformations to the images are applied at a subpixel level using bicubic interpolation algorithms. Lehmann and colleagues (Lehmann, Gonner et al, 1999) provide an extensive review of image processing interpolation algorithms. The magnitude of each of the 6 transformations was made unique in each simulation by using a random number sampled from a Normal distribution where the mean of the size of the transformation is set at zero and a variance fixed for each transformation. To mimic background noise, Gaussian noise is added to each pixel with variance *v* and mean zero. A proven non biased random number generator was used to sample from a normal distribution (Press, 2002). Figure 7 shows the computational sequence used to mimic a patient series.

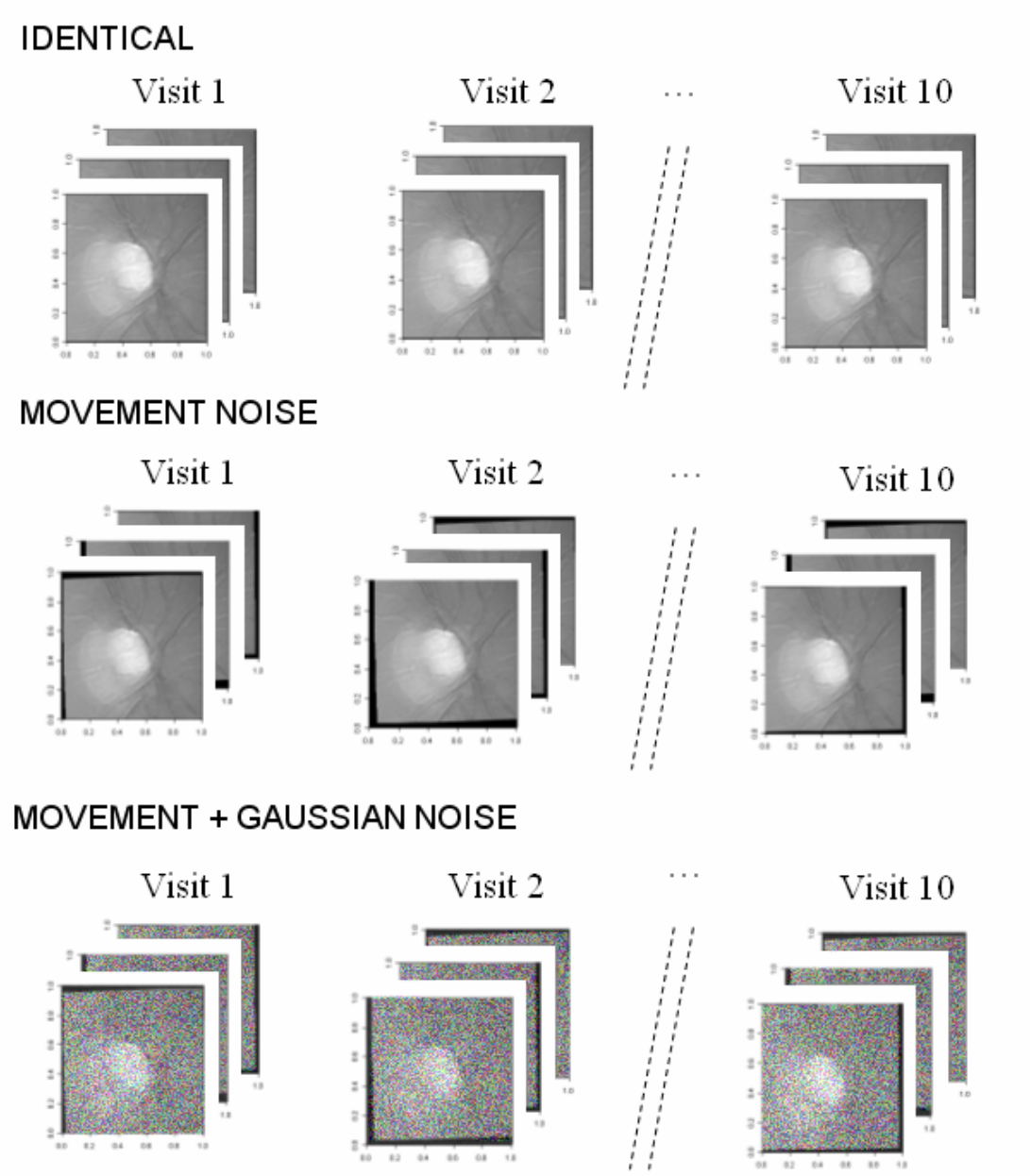


Figure 7 Computer simulation of a patient's image series. A topography image is replicated 30 times to represent 10 visits with 3 scans acquired at each visit. Then 'movement' and Gaussian noise are added

The MPHSD was mimicked using a trial and error approach by varying the spread of the normal distribution used to simulate each of the movement parameters (x', y' z', σ^x, σ^y, σ^z) and the Gaussian noise.

The resultant between-image variability demonstrated characteristic noise patterns previously reported in literature (see Figure 8), with higher variation in areas of high gradient change, such as across blood vessels and the cup border (Brigatti, Weitzman et al, 1995; Chauhan and McCormick, 1995).

Figure 8 The result of calculating standard deviations of the topographic height at each *pixel(i,j)* in an image series of a simulated stable patient. The darker pixels (seen along blood vessels) indicate areas of high variability; this pattern would be expected in a real series

There are many solutions to finding an acceptable noise pattern. This thesis (chapters 3 & 4) reports the results from computer experiments which use this simulation to test new techniques ability to distinguish between 'stable' and 'unstable' patient series. The reader should however be aware of the assumptions made by the simulation as any simulation might not be truly representative of the true data. The spread of the normal distribution which is randomly sampled to apply translation and rotation to mimic movement noise was estimated by trial and error. This is because it is impossible to measure the remaining movement error that remains after the HRT alignment algorithm are applied to image series. Also, the simulation randomly samples from a normal distribution which results in the within and between visit variability being consistent. Chauhan and MacDonalds (Chauhan and MacDonald, 1995) results support this assumption in finding no statistically

significant difference between scans taken on the same day and those acquired on separate visits. The simulation assumes no auto-correlation in time. This means that a single pixel location will be represented by 30 topographic height values (10 visits with 3 scans per visit), the simulation assumes that these values are independent of each other (knowing one value does not provide you with any information of another value). Auto-correlation has however been reported in RA measurements obtained from the HRT (Crabb, Owen et al, 2006).

In the proceeding chapters (3 and 4) computer experiments use this simulation to replicate the repeatability of topographic height measurements typical of real patient data.

3. SIM: a new technique for detecting change in series of ONH images

The real promise of the HRT may lie in objectively measuring progressive structural damage, or stability, in patients being followed over time. This is possible because the local height measurements at each of the pixels of a topography image are sufficiently reproducible (Chauhan, LeBlanc et al, 1994; Rohrschneider, Burk et al, 1994). However, to date research in this area has used stereometric parameters, summary measures of the ONH, to detect change. Alternately, the TCA (see section 1.3) developed by Chauhan and colleagues and available on the current HRT software identifies areas of change at a superpixel level (4 × 4 array of pixels) on the ONH. As discussed previously this technique fails to handle the multiple comparison problem, spatial correlation across the image and has no specific mechanism to determine if an area of damage is statistically significant.

In neuroimaging, PET (positron emission tomography) or MRI (magnetic resonance imaging) scans yield a sequence of three dimensional images of the subject's brain from which the temporal and spatial characteristics of neuronal activity can be deduced. In the case of fMRI (functional MRI) for example, this is done by measuring changes in cerebral blood oxygenation related to brain activity. The images are complex and high-dimensional, typically containing as many as 100,000 measured volume elements or voxels (3 dimensional pixels). Consequently the neuroimaging research community has developed an extensive suite of techniques to register, align, process and analyse arrays of imaging data (Frackowiak Richard, 1997). This thesis exploits this work, specifically the techniques collectively referred to as *Statistic Image Mapping* (SIM) sometimes referred to as (Statistical Parametric Mapping - SPM) which are used for determining areas of activity and change in MRI and PET type images, and by applying them to series of retinal and ONH topography images. In particular, we use a non-parametric version of these techniques; these are intuitive to understand and assessment of change in the image

is based solely on the subject's own data and within-subject image variability, rather than any *a priori* information or patient population characteristics.

This chapter describes and applies SIM techniques to HRT images. The performance of this new statistical approach is evaluated by comparing it to the TCA method. This is done with an extensive series of computer simulation experiments. SIM is then applied to longitudinal sets of real HRT data from patients and normal subjects to make comparisons with the TCA method.

The work in this chapter has formed a paper published in *Investigative Ophthalmology and Visual Science* (Patterson, Garway-Heath et al, 2005). It was also presented at the International Perimetric Society conference in Barcelona, Spain on June 29 - July 2, 2004; at Moorfields Eye Hospital Bicentenary meeting, London, UK on 16-19 March 2005; and at Image Morphometry and Glaucoma in Europe Meeting, Milan, Italy on April 4-5, 2005. The results in this chapter have also been described in part at the American Academy of Ophthalmology meetings New Orleans, USA, on Oct 23-26, 2004 and Chicago, USA, on Oct 15-18, 2005.

3.1 Methods

As described in Chapter 1 a known characteristic of progressing glaucoma patients is increasing optic nerve head (ONH) excavation and nerve fibre layer thinning with time, often called structural progression (Drance, 1975; Schwartz, 1976; Spaeth, Hitchings et al, 1976). The ideal clinical tool for assessing a longitudinal set of these HRT images would highlight this structural progression as localised areas of the ONH that are 'changing' beyond the natural 'within test' and 'between test' variation in the images. What follows is the description of a quantitative statistical techniques developed for such a purpose. Here it was applied to series of HRT images, but these methods are applicable to several of the retina imaging modalities.

Statistic Image Mapping

The methods take advantage of proven statistical techniques that have been developed to analyse series of MRI and PET images. These analyses usually proceed by computing a statistic at the pixel level (or 3-D pixel, voxel, in the case of MRI and PET images), indicating evidence for the observed effect of interest, and resulting in an 'image of statistics', or a *statistic image map*. Each pixel in this statistic image is simply a number (or statistic) indicating the level of change. The entire statistic image must be assessed for significant effects, using a method that accounts for the inherent multiplicity involved in testing all the pixels at once (see section 1.3 for a description of the statistical problems encountered when applying standard statistical methods to imaging data). This analysis can be accomplished in a classical parametric statistical framework (Worsley, Evans et al, 1992; Friston, Holmes et al, 1995), but we use an alternative that is based on *permutation testing*. The latter is conceptually simple, does not rely on any theoretical probability models that may or may not be appropriate for the HRT data, deals with the multiple comparison problem of testing a vast image space and, critically, derives significance limits for change based only on the individual patient's series of data. These specific techniques and the mathematics that underpin them, as applied to PET and MRI data, are extensively described elsewhere (Arndt, Cizadlo et al, 1996; Holmes, Blair et al, 1996; Everitt and Dunn, 1998; Bullmore, Suckling et al, 1999; Nichols and Holmes, 2002; Hayasaka and Nichols, 2003; Hayasaka, Phan et al, 2004). What follows is a description of three component parts of this approach, and how we apply them to a series of HRT data.

Permutation testing at individual pixels

Consider that three HRT images, at each patient visit, are acquired at regular intervals over a clinical follow-up. After registration of the image series, the topographic height at each individual pixel is considered in turn. Visually this could be done by plotting the topographic height at each pixel as a time series (see Figure 9). Next a suitable statistic is derived for summarising the change, or stability, of the topographic height at that pixel over time: the line of best fit (slope) derived from

ordinary least squares regression. The standard error of this slope gives an indication of how well the data fits the linear trend, with relatively high values indicating a poor fit or a noisy series of observation. Our test statistic at each pixel is simply the absolute value of the slope divided by the standard error: a relatively large test statistic would be evidence of clear linear change of topographic height at that pixel. This process is performed at all of the pixels, and the patient's series of data is reduced to a *statistic image*; this is no longer a physiological image but a 256 × 256 map of statistics 'summarising' change within the image (see Figure 10).

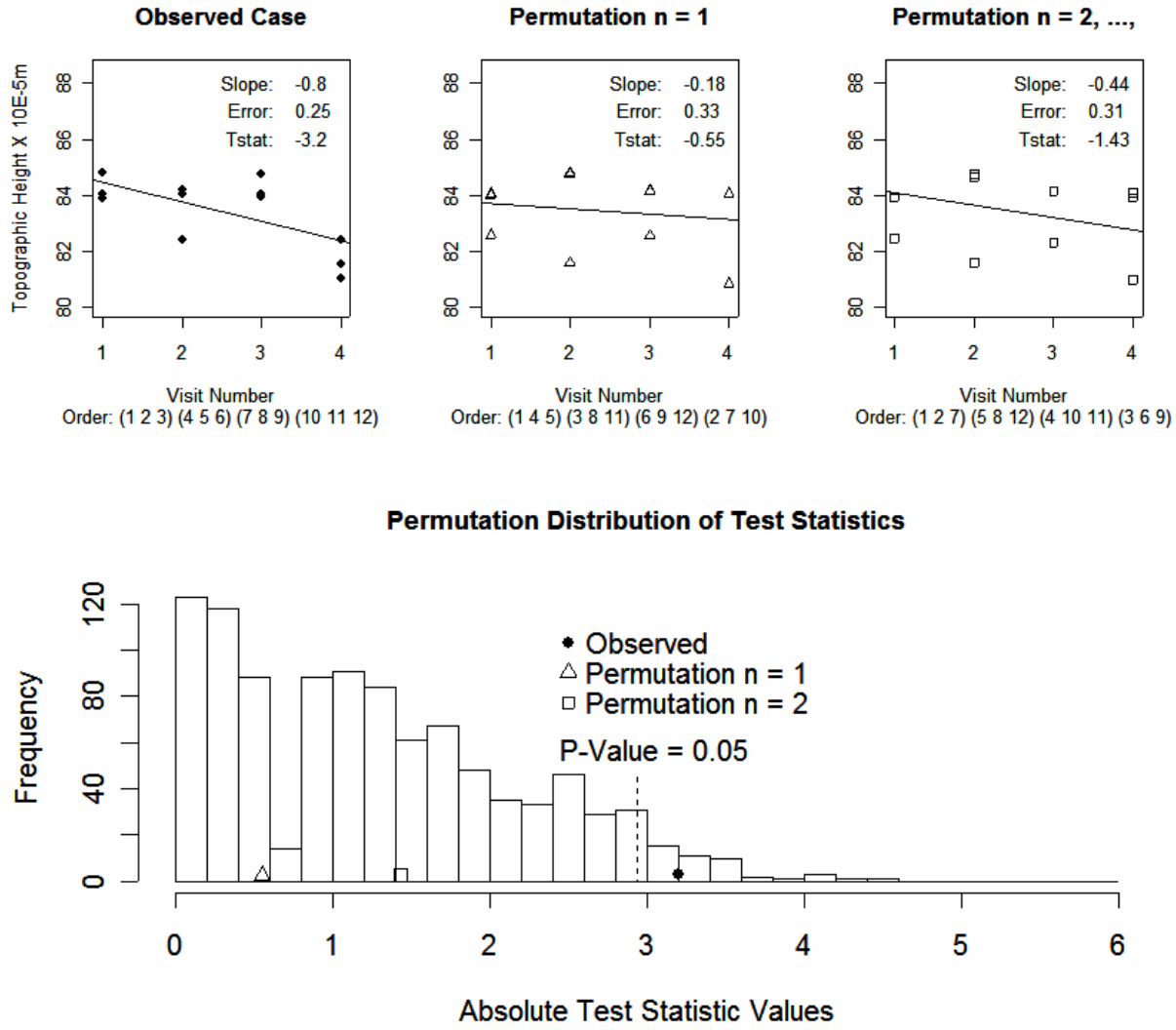


Figure 9 The permutation distribution of test statistics at *pixel(i,j)* is calculated by generating 1000 unique permutations, see the computation paradigm in section 3.2 for further details. The observed (●) and the first two unique permutations (□, Δ) are marked on the distribution. The probability that *pixel(i,j)* is statistically significant is defined as a value which exceeds the 95th percentile in the permutation distribution (marked by the dashed line). As the observed test statistic is very unusual (a P-value less than 0.05) *pixel(i,j)* is marked as statistically significant on the statistic image

The next step is to determine whether the observed test statistic at each pixel is 'unusual', or more extreme, than would be expected by chance. This testing of the

significance of the test statistic is not completed in the conventional manner, by considering the observed test statistic as a random variable from a probability model, but uses a *permutation test.* We randomly 'shuffle', or re-label, the order of the observed data and recalculate the test statistic for all possible permutations of the order of images. If we let N denote the number of all possible labellings, t_i the statistic corresponding to labelling i, then the set of t_i for all possible re-labelling constitutes the *permutation distribution.* For example, there would be 369600 ($12!/(3! \times 3! \times 3! \times 3!)$) of these in a series of four clinical visits with three scans at each visit (see the computational paradigm in section 3.2 for more details of this calculation). We then assume that all of the t_i are equally likely and determine the significance of the observed test statistic by counting the proportion of the permutation distribution as or more extreme than our observed value, giving us our p-value; if this is, for example, less than 5% we label the pixel as being active or 'changing'. We therefore assume that images acquired at the same visit are no more correlated than images acquired between visits; previous work on the influence of time separation on inter-image topographic variability support the intuition behind this approach (Chauhan and MacDonald, 1995). This *permutation test* is done 'pixel by pixel' and the statistic image becomes 'thresholded' at the 5% level with pixels flagged if they are 'significant' or not (see Figure 9). In practice, a sample of 1000 randomisations (drawn without replacement from all the possible labellings) are used to generate the permutations distribution (Manly, 1991; Good, 2000). This eases the computation burden but still allows for a statistically exact test at standard levels of significance testing (Manly, 1991). It is worth noting, however, that larger samples would be needed to accurately model the tail of the probability distribution for smaller p-values such as 0.1%.

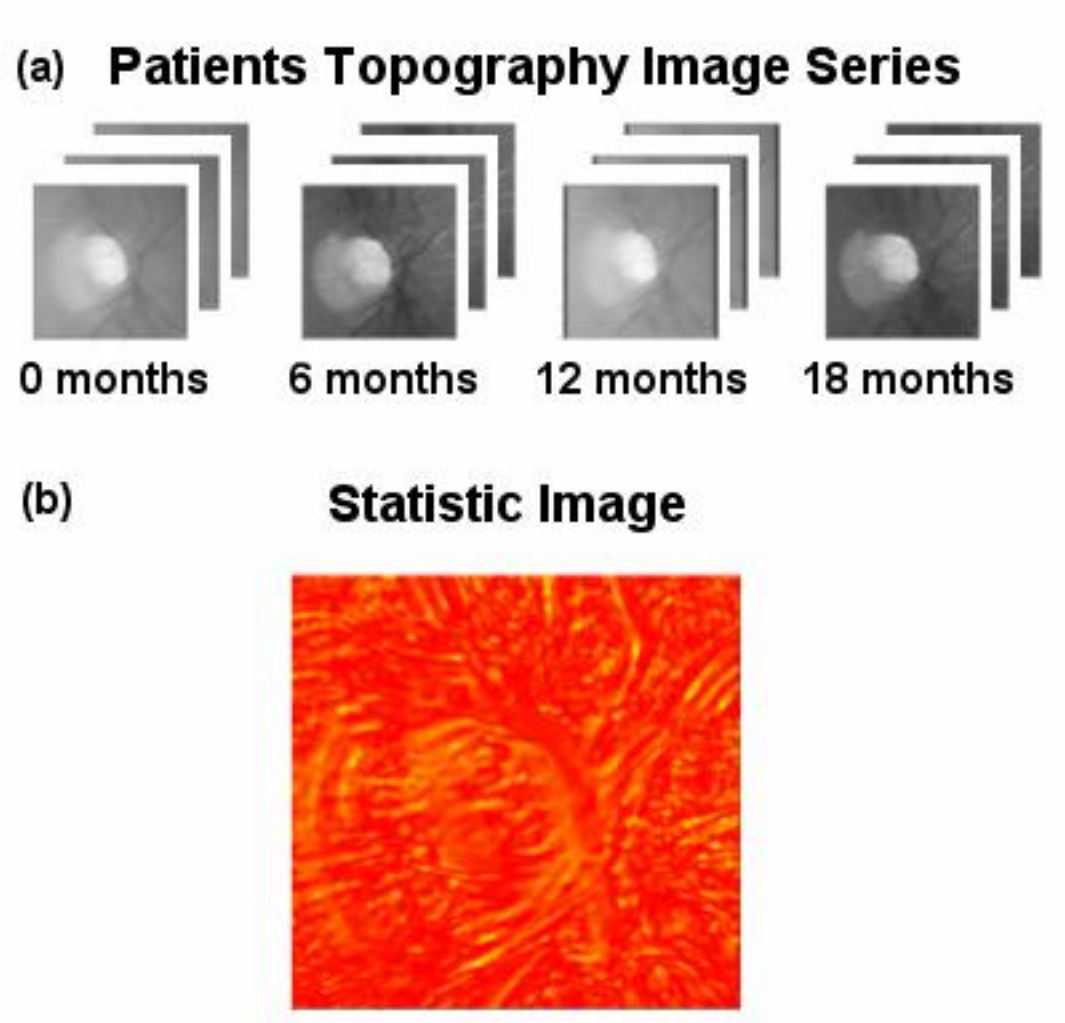


Figure 10 (a) An example of a typical patients topographic image series. Three images are typically acquired at each visit. (b) A statistic image is generated by calculating a statistic at each pixel location. In this case linear regression is performed, each statistic is comprised of a slope divided by the standard. For display purposes the statistics are represented in a colour coded form, red represent a small statistic through to yellow representing a larger statistic

Permutation testing to threshold clusters

Thus far we have considered a separate analysis at each of the 65356 pixels within the HRT image, with no attempt to take account of the multiplicity of testing. Statisticians refer to this as the *multiple comparisons problem* and the construction of a corrective analysis for high dimensional MRI and PET data has occupied many authors, with ideas ranging from the simple use of Bonferroni adjustments to other

mathematical solutions (Nichols and Holmes, 2002). In this work we again exploit an intuitive approach, once more using a permutation test, which has been successfully applied to sequences of MRI and PET images, and out-performs other approaches when there are few images involved (or experiments with low degrees of freedom). Once we have thresholded the statistic image pixel by pixel (see Figure 11), we are left with an image that will contain clusters of contiguous, significant or active pixels. An active pixel is defined as part of a cluster if one or more of the pixels nearest neighbours are active; there are 8 pixels in each pixels neighbourhood and a pixel is therefore part of a cluster if any of the 2 nearest horizontal pixels, 2 vertical pixels or 4 diagonal pixels are active. The size of the largest cluster in the observed image is recorded. To ascertain whether the spatial extent of the clusters in the observed image is unusually large by chance alone, we set about 'shuffling' the images again, re-compute the statistic image, calculate the cluster sizes and record the size of the largest cluster. (In fact the 'shuffling' for the pixel by pixel analysis and the cluster testing is all done in one 'sweep' in the computational algorithm). This is repeated to generate a permutation distribution of the maximum cluster size (see Figure 11); hence, we assess the significance of the observed result by considering only the patient data itself, and no knowledge of the probabilistic behaviour of the topographic heights at image pixels is required. This is particularly useful because of the *spatial correlation* that exists within the image (i.e. the topographic height of neighbouring pixels will be more similar in some part of the image than others) and, in part, this cluster testing accounts for this. The threshold value generated to determine progression will be unique for each patient and will vary depending on the patients signal to noise ratio. The criteria for progression included only depressed clusters (a continuous set of active pixels whose slopes are negative) bound within the contour line for the optic disc. Section 3.2 contains the computational details.

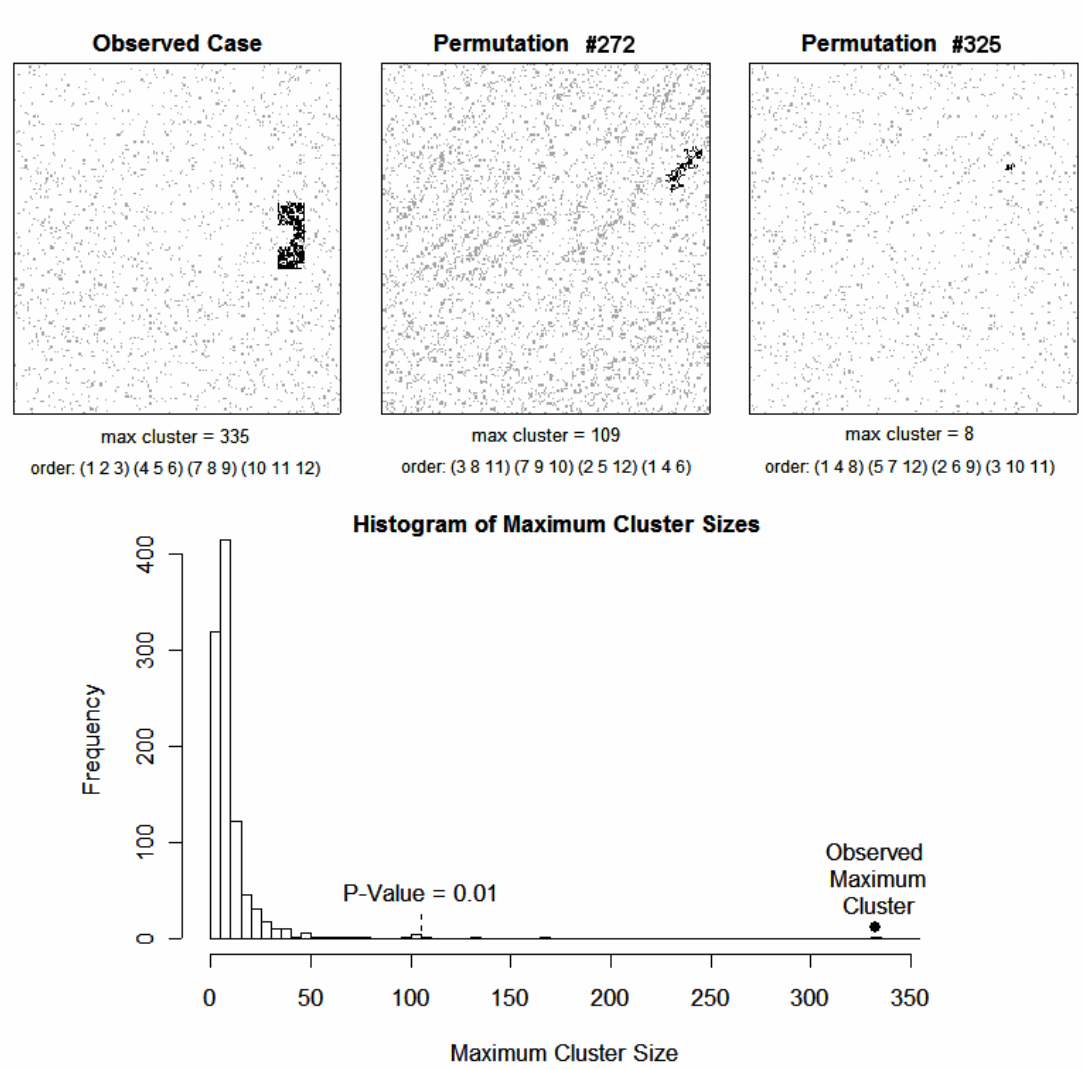


Figure 11 Simulated change: active (changing) pixels whose slopes are negative are shown in grey, with the largest cluster highlighted in black. We show the observed statistic image and two of the 1000 permutations. The distribution of maximum cluster sizes is created by recording the largest cluster of active pixels in the statistical image for each unique permutation. In this case one cluster in the observed statistic image (●), generated by simulating a progressing patient, is very unusual (P-value smaller than 0.01), therefore the virtual patient is classed as progressing

Pre-processing: the pseudo test-statistic

A prerequisite for any pixel by pixel analysis of a series of images is that any given pixel represents precisely the same anatomical region across the series. Even after the HRT software alignment procedures this is a considerable leap of faith. A

proven solution to this problem is available and this involves the generation of a *pseudo test statistic*. Rather than divide the 256 × 256 matrix of individual slope values by the 256 × 256 matrix of individual standard error values to yield the test statistic, the slope values are divided by a spatially filtered standard error value (see Figure 12). The latter is the matrix of standard error values smoothed with a weighted, Gaussian kernel. Thus a pseudo test-statistic image is formed by dividing the slope matrix by the smoothed standard error matrix. Hence all the analyses, including the permutation cluster testing, proceeded with these pseudo-test statistics.

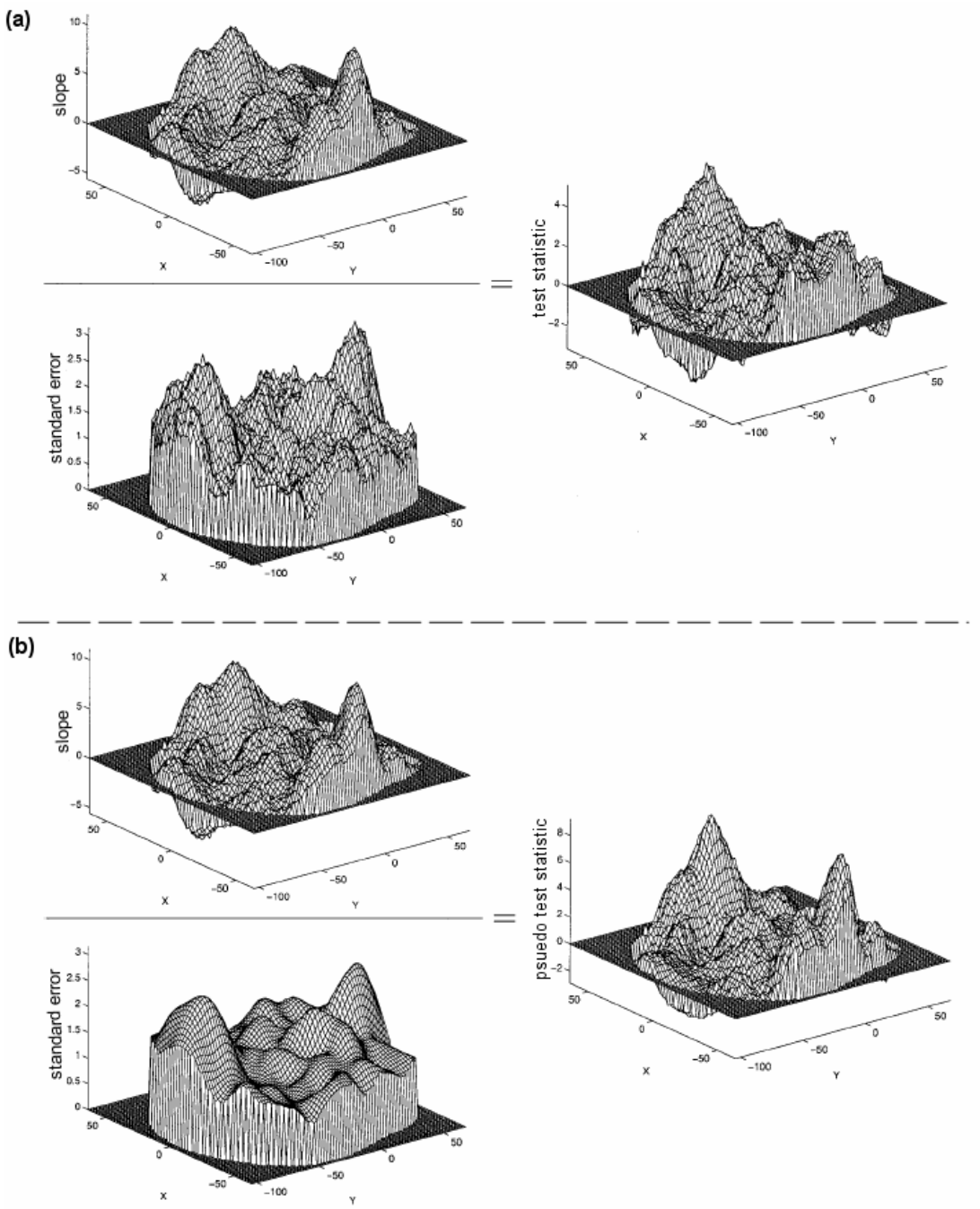


Figure 12 Illustrates the computation of the pseudo test statistic on an fMRI image. (a) Shows the slope and standard error at each pixel location. The test statistic plot is a result of dividing the slope by the standard error terms. In this example the test statistic image appear highly variable. (b) To calculate the pseudo test statistic the standard error term are first spatially smoothed. The resulting pseudo test statistic plot appears less variable. (Courtesy of Dr Holmes: permission sought to use these figures through private communication)

In essence the noise from the variance image (the matrix of standard error values) has been smoothed but not the signal. Statistic image maps constructed with smoothed variance estimates have been shown to substantially improve the power of the approach and can only be used in the non-parametric or permutation setting outlined here (Arndt, Cizadlo et al, 1996; Holmes, Blair et al, 1996; Bullmore, Suckling et al, 1999; Nichols and Holmes, 2002).

3.2 Computational paradigm

The aim of this section is to allow the reader to directly replicate the computational paradigm behind SIM. The order of the three components of SIM has been changed to reflect the computational sequence. Figure 13 provides a schematic of the entire SIM paradigm.

Permutation testing at individual pixels

A feature of permutation tests is the number of combinations required for testing a probability limit. In practice a sample of 1000 randomisations (drawn without replacement from all the possible labellings) are used to accurately model the tail of the permutations distribution at P=0.05 (Manly, 1991).

The number of possible unique permutations is expressed as:

$\frac{(s \times n)!}{(s!)^n}$ where s = number of scans per visit and n = number of visits

For example, with 4 visits and 3 scans per visit there are:

$\frac{(3 \times 4)!}{(3!)^4} = 369600$ unique permutations.

The following steps represent the computational paradigm to compute a permutation distribution and test statistical significance

1. At each *pixel(i,j)* calculate by least squares linear regression the slope *b(i,j)*, standard error *se(i,j)* and absolute test statistic *t(i,j)* of time (dependant variable) against topographic height (independent variable).

2. Shuffle the order of the dependant variable (time) to generate a unique permutation and recalculate *b*, *se* and *t*.
3. Perform step 2 above, in total 1000 times calculating a unique permutation each time. As each permutation must be unique the algorithm must perform sampling without replacement.
4. We reject the null hypothesis at a significant level of $P<0.05$; thus, for the mechanics of the permutation distribution we reject the null hypothesis if the observed test statistic is greater than or equal to the 95%ile of the permutation distribution. Therefore, sort the array of test statistic *t* produced at each *pixel(i,j)* into ascending order and test if the absolute observed test statistic is ≥ to 950^{th} (0.95×1000) value of *t*. Define this value as $t_{critical}$. Note that we retain the sign of the observed test statistic to indicate the direction of change i.e. a negative sign indicates a depression in topographic height values over time, whereas a positive sign indicates an elevation in topographic height values over time.

In practice a computer program was developed to calculate a lookup table of unique permutations, a separate lookup table was then created for each visit. Therefore an example of a unique permutation at visit 4 could be (2 10 11)(1 3 12)(4 5 8)(6 7 9) and at visit 5 could be (2 10 13)(1 3 12)(4 5 8)(6 7 9)(11 14 15). Generating these lookup tables and saving them to disk saves considerable computer time, as each time the SIM paradigm is executed it does not require evaluating 1000 unique permutations at each visit.

Pre-processing: Pseudo test-statistic

The pseudo test statistic *tstat(i,j)* is calculated by dividing slope *b(i,j)* with a spatially smoothed standard error *se(i.j)*. The smoothed standard error is calculated by convolving the standard error *se(i,j)* with a Gaussian kernel. We used a square Gaussian kernel of symmetrical full-width half maximum of 11 and size 17 by 17 to smooth the standard error *se(i,j)*. Chapter 3 in (Gonzalez and Woods, 2002) provides details on the computational method required to generate Gaussian kernels

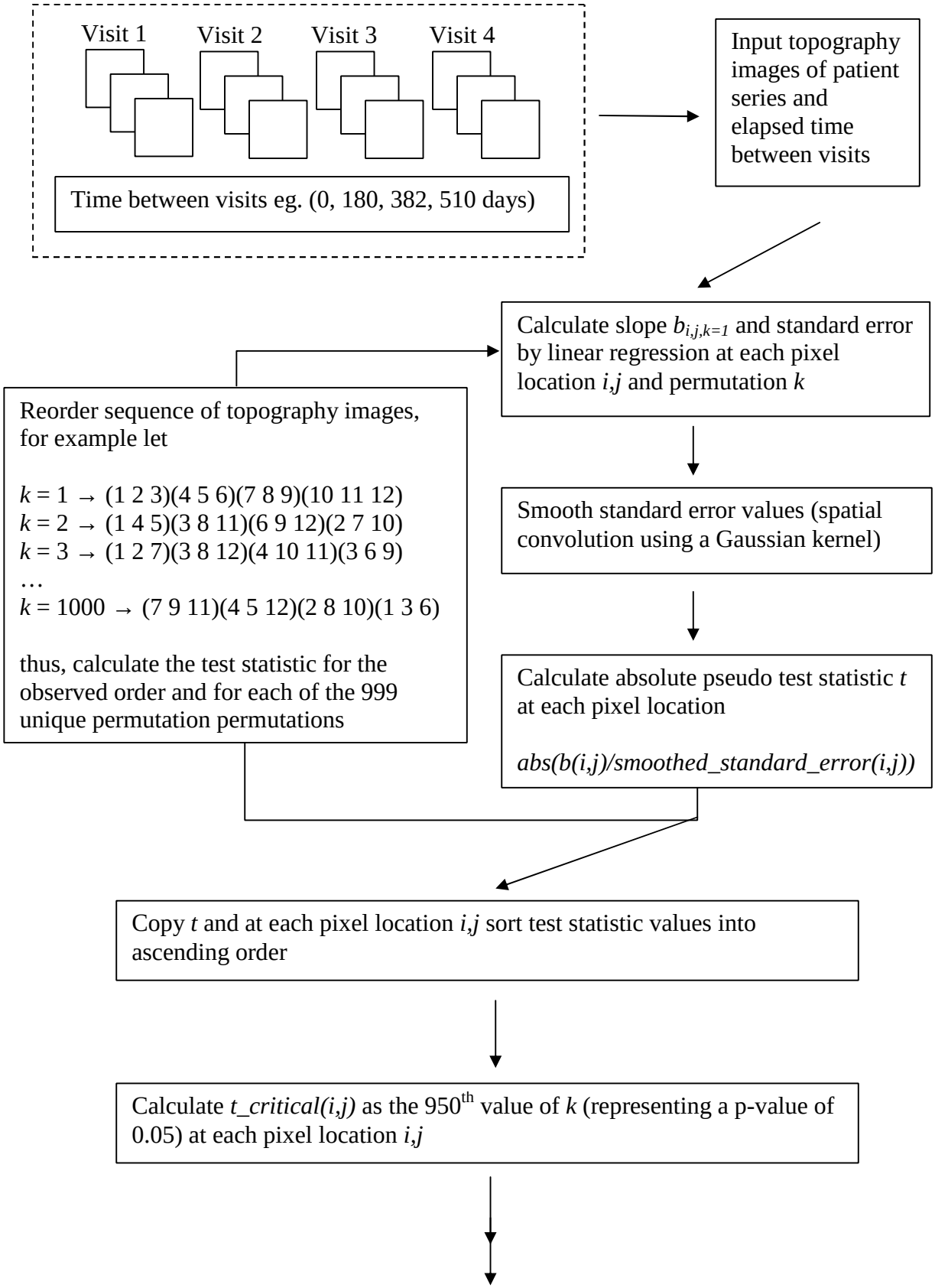
and perform spatial convolution. The pseudo test statistic is calculated for the observed case and for each unique permutation.

Permutation testing to threshold clusters

The following paradigm is a computational method for thresholding clusters:-

1. Compute the observed pseudo test statistic.
2. Compute the pseudo test statistic for each unique permutation.
3. Compute an observed statistical image *s(i,j)* by setting *s(i,j)* = *active_depressed* or *active_elevated* if the observed absolute pseudo test statistic is ≥ the 95^{th} percentile of the permutation distribution of absolute pseudo test statistic at *pixel(i,j)*. Record the size of the maximum depressed and elevated clusters within the observed statistic image, bound within the contour line (defined as the 'area-of-interest'). An active pixel within a statistical image *s(i,j)* is part of a continuous cluster if one of the eight pixels within its neighbourhood is also active i.e. 8 - connectivity.
4. Compute a statistic image at each of the 1000 unique permutations. Record the size of the maximum depressed and elevated clusters for each unique permutation, bound within the contour line.
5. Sort the array of maximum clusters into ascending order.
6. A depressed or elevated cluster (or clusters) within the observed statistic image is statistically significant if it (or they) are larger than the 99%ile of the maximum depressed and elevated cluster distributions. Progression is defined if a depressed cluster is larger than the 99%ile of the maximum depressed distribution.

An 'area-of-interest' was set as mentioned in step 3 above as the area bound within the contour line. To determine if a pixel is bound within the contour line draw either a horizontal (or vertical) line in a fixed direction from the pixel, a pixel is bound within the contour line if it passes the boundary of the contour line an odd number of times (e.g. 1 or 3) and is outside the contour line if it passes zero or an even number of times.


Visit 1
Visit 2
Visit 3
Visit 4
Time between visits eg. (0, 180, 382, 510 days)
Input topography images of patient series and elapsed time between visits
Calculate slope $b_{i,j,k=1}$ and standard error by linear regression at each pixel location i,j and permutation k
Reorder sequence of topography images, for example let
$k = 1 \rightarrow$ (1 2 3)(4 5 6)(7 8 9)(10 11 12)
$k = 2 \rightarrow$ (1 4 5)(3 8 11)(6 9 12)(2 7 10)
$k = 3 \rightarrow$ (1 2 7)(3 8 12)(4 10 11)(3 6 9)
…
$k = 1000 \rightarrow$ (7 9 11)(4 5 12)(2 8 10)(1 3 6)
thus, calculate the test statistic for the observed order and for each of the 999 unique permutation permutations
Smooth standard error values (spatial convolution using a Gaussian kernel)
Calculate absolute pseudo test statistic t at each pixel location
$abs(b(i,j)/smoothed\_standard\_error(i,j))$
Copy t and at each pixel location i,j sort test statistic values into ascending order
Calculate $t\_critical(i,j)$ as the 950^{th} value of k (representing a p-value of 0.05) at each pixel location i,j

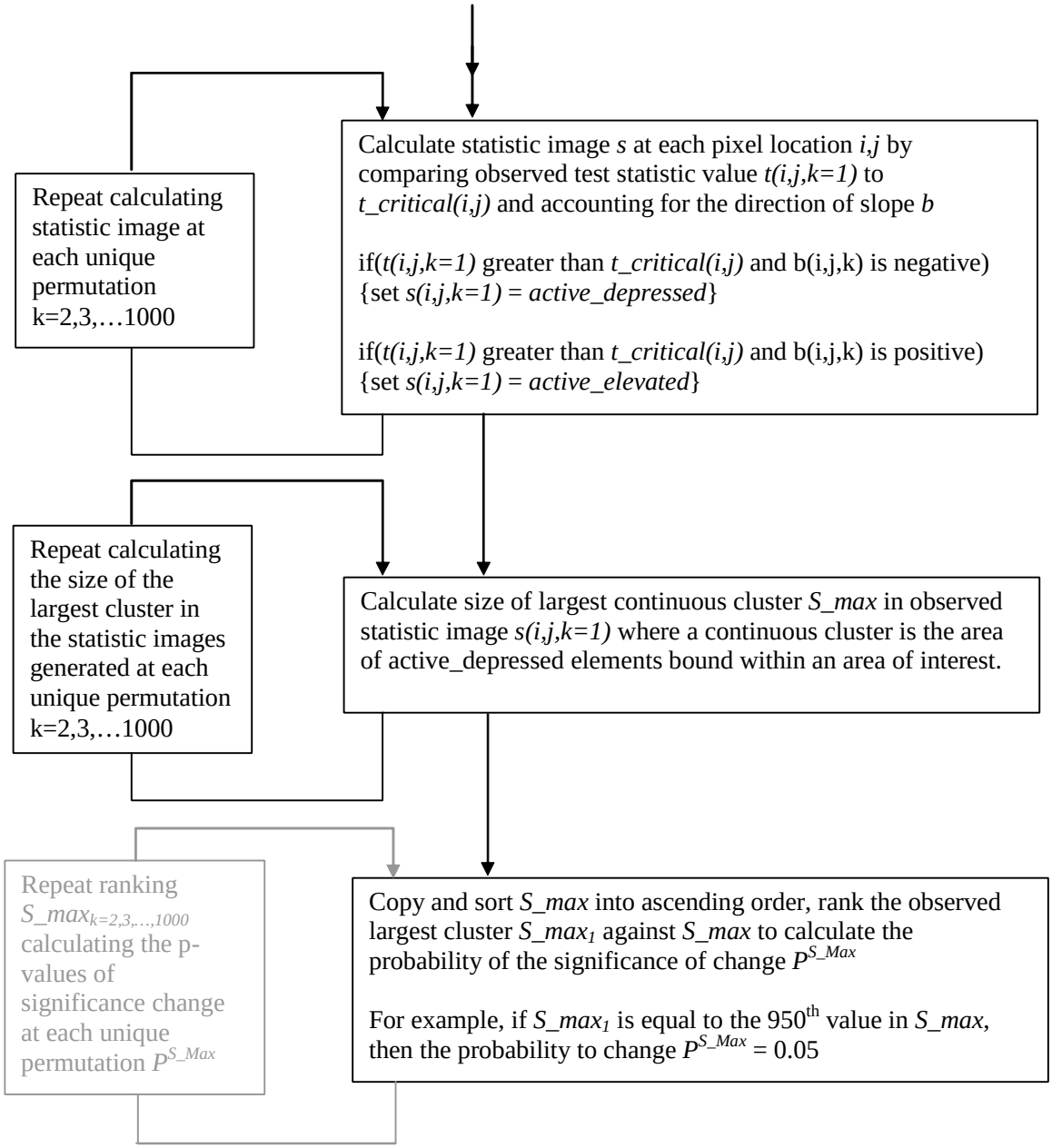


Figure 13 Schematic of the SIM computational paradigm. The details shown in grey will be referenced to in chapter 4

3.3 Testing the new approach

The performance of SIM was tested against the TCA method in a computer 'virtual patient' simulation described in chapter 2. The TCA method was replicated in consultation with the authors of the technique (David Hamilton, Department of

Mathematics and Statistics, Dalhousie University, Canada, private communication, 2004). To do this C++ software was written to replicate the TCA algorithm and visualise the TCA output. (The TCA technique is described in section 1.3.) In this experiment we used a criteria for change implemented in a previous publication (Chauhan, McCormick et al, 2001): any 'virtual patient' who showed a cluster of 20 or more significant superpixels bound within the contour line for the optic disc, where the topographic change compared with baseline occurred in 3 consecutive sets of follow-up images, was considered to have confirmed progression.

Progressing patients-series' with gradual change (linear) were simulated by applying a cumulative decay of 5µm per visit to a cluster of 480 pixels to the neuro-retinal rim. Progressing patients-series' with episodic change (sudden) were simulated by applying a height decay of 50µm to the cluster at a randomly selected visit between visit 2 and visit 10, inclusive. To replicate the repeatability of topographic height measurements seen in clinical data, groups of virtual subjects were simulated having a mean pixel height standard deviation (MPHSD) of 15, 25 and 35 µm. Chapter 2 provides full details of the simulation. Each simulated series was stored to computer disk allowing the specificity and sensitivity of both techniques to be evaluated on identical image series.

Computer experiments

Specificity was examined in our first set of experiments by generating 300 stable 'virtual patient' series. Three groups of 100 virtual patients were generated with a MPHSD of 15, 25 and 35µm respectively. We then applied our new SIM technique to these data, using the criteria for change specified in section 3.2, recording for each patient series the visit at which (false-positive) change was first detected. We then applied the TCA method to the same dataset, again recording for each patient series the visit at which (false-positive) change was first detected.

The sensitivity of the techniques was tested in 6 separate experiments: for gradual (linear) change and sudden (episodic) change; with change applied to a cluster of an

area of 480 pixels; and with a MPHSD of 15, 25 and 35µm. The same progression criteria were used as for the specificity experiment. The follow-up visit at which change was first detected was recorded for both the SIM and the TCA analysis.

The SIM technique, the replicated TCA method, the simulations and the computer experiments were all developed in purpose written software using C++.

Real longitudinal HRT series

The techniques were further tested on a selective sample of clinical data: OHT patients which were selected from the OHT clinic at Moorfields Eye Hospital (London) and normal age-matched controls. The patient group were selected from a subset of OHT patients who had developed reproducible visual field loss while under observation. The control group were typically spouses, friends and family of the OHT patients. The study groups are described in detail elsewhere (Kamal, Viswanathan et al, 1999; Kamal, Garway-Heath et al, 2000); these adhered to the declaration of Helsinki and local ethical committee approval was obtained. In short, OHT patients had an intraocular pressure (IOP) of ≥ 22 mmHg on two or more occasions, two initial reliable visual fields with AGIS score of 0 (AGIS, 1994; AGIS, 2000), absence of other significant ocular disease that would effect visual field performance and age > 35 years. The eligibility criteria for the normal subjects included IOP consistently < 21 mmHg, baseline reliable visual fields with an AGIS score of 0, no significant ocular disease, no family history of glaucoma and age > 35 years. A reliable visual field was defined as <25% fixation errors, <30% false positive errors and <30% false negative errors. The normal subjects were followed concurrently with the OHT patients.

Thirty OHT eyes that 'converted' to a clinical diagnosis of glaucoma (converters) during the follow-up and 20 eyes of 20 normal subjects were randomly selected. A converter was defined as an eye with an initial AGIS score of 0 and follow-up AGIS scores of ≥ 1 on three consecutive reliable visual fields. The reader should be aware that the AGIS scoring system may suffer from a low sensitivity in detecting visual

field deterioration as the criteria was developed to determine progression in patients with advanced glaucoma. Both groups (converters and normals) were imaged at regular intervals; the converters follow-up period ranged from 2.8 to 7.3 years and the controls ranged from 2.8 to 7.3 years. Twenty-one topography images (representing 7 visits with 3 scans per visit) were selected from each subject, taking the images from the baseline and last visit and images from 5 interim visits. Image quality was not a factor in the selection of subjects.

The topography images were extracted from the Moorfields HRT database using the scientific features of the HRT Eye-Explorer software v1.4 (Heidelberg Engineering, Heidelberg). The image data were exported as aligned for analysis by the HRT software and then subjected to SIM analysis exactly as described for the simulation experiments (using the same progression criteria at visits 4 to visit 7). TCA was performed using the HRT software.

3.4 Results

Computer Simulation

In the 300 stable 'virtual patients', under the conditions of these computer experiments, the TCA method flagged 16%, 17% and 17% at MPHSD of 15, 25 and 35µm, respectively, at some point in the follow-up series (false-positives). These values are closer to 10% in the first half of the follow-up but worsen as more visits are considered. SIM had much better specificity, with 6%, 5% and 5% flagged at the different levels of noise (see Figure 14a).

In the simulations of progressing patients, the TCA method identified progression at some point in follow-up in 95%, 31% and 28% with linear change, and 82%, 47% and 42% with episodic change, for the MPHSD of 15, 25 and 35µm conditions, respectively. SIM identified 100%, 68% and 62% with linear change, and 86%, 57% and 55% with episodic change (see Figure 14(b,c,d)). For these experiments, the TCA had slightly better or similar sensitivity as compared to SIM at detecting gradual (linear) change up to about visit 6 or visit 7, with SIM outperforming TCA

as more data became available. A similar pattern emerged when episodic loss was specified, but with equal sensitivity when the noise was low (MPHSD 15µm).

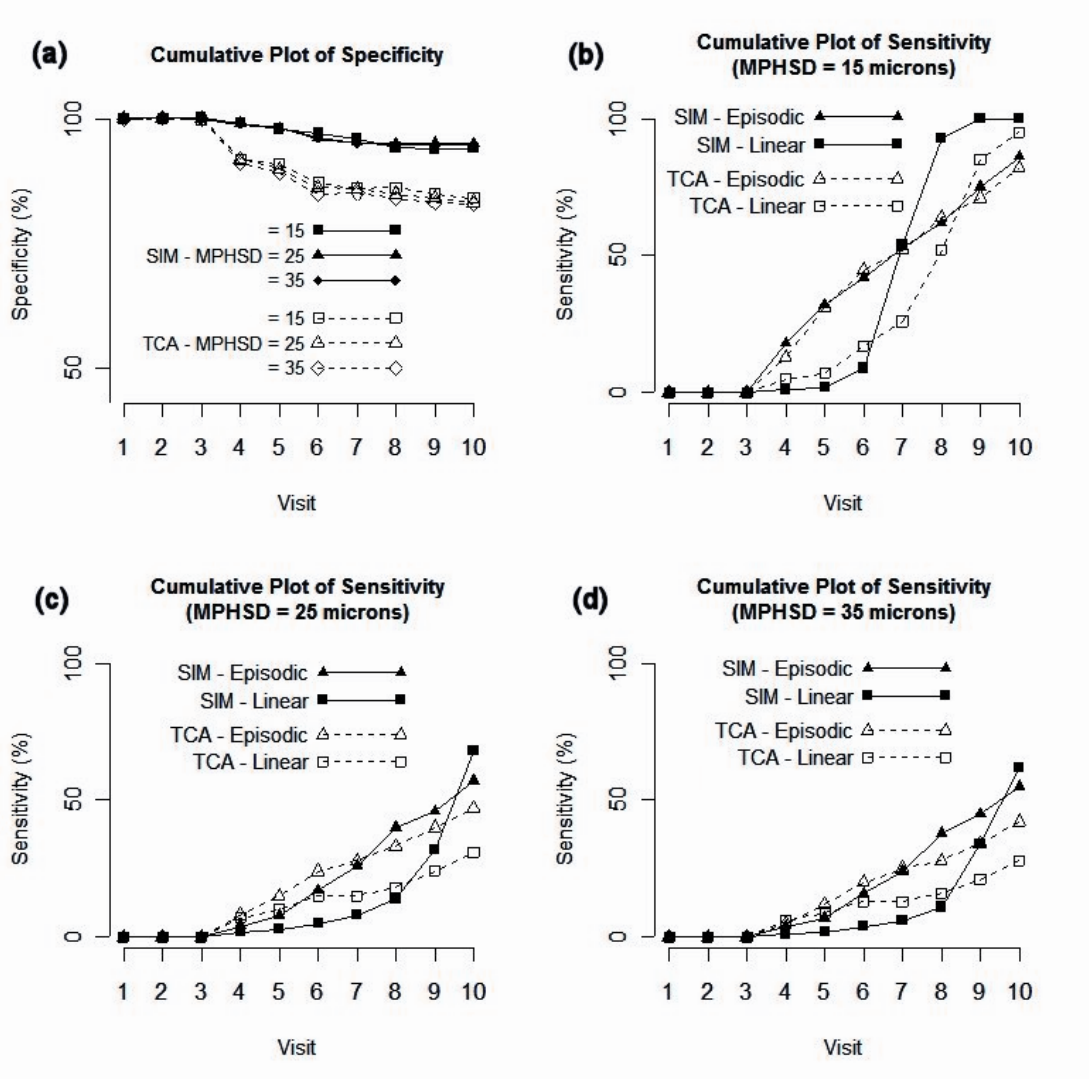


Figure 14 Computer simulation results comparing the diagnostic precision of Statistical Image Mapping (SIM) and the Topographic Change Analysis (TCA) superpixel method. (a) The specificity of SIM and TCA at MPHSDs of 15, 25 and 35µm. (b)(c)(d) The ability of SIM and TCA to detect gradual (linear) and episodic (sudden) loss at a cluster of 480 pixels to the neuro-retinal rim area at MPHSD of 15µm (b), 25µm (c) and 35µm (d)

Real longitudinal HRT Series

The results are summarised in both Figure 15 and Table 1. Examples of the similarity and differences between the SIM and TCA results are illustrated in Figure

16. Cases 1 and 2 are both OHT converters: in case 1 both SIM and TCA confirmed progression at visit 4; in case 2 SIM identified progression at visit 6 whereas the TCA did not detect progression at all.

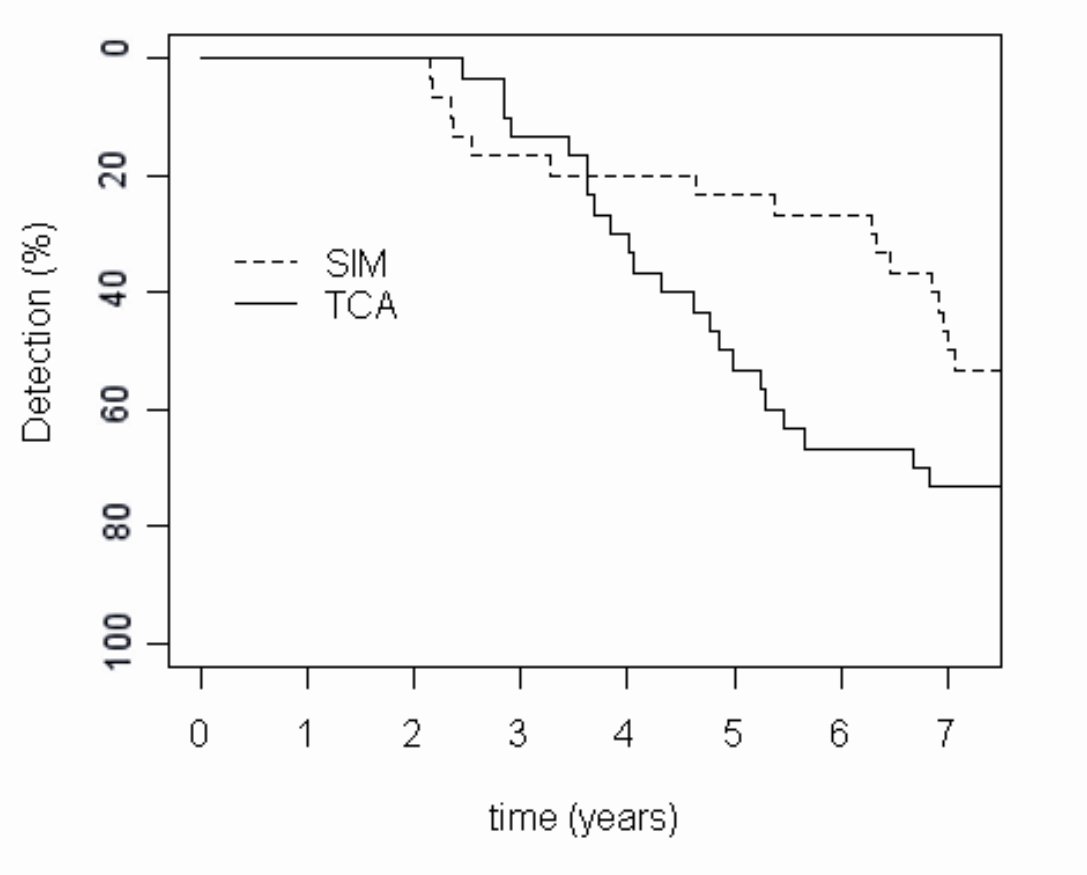


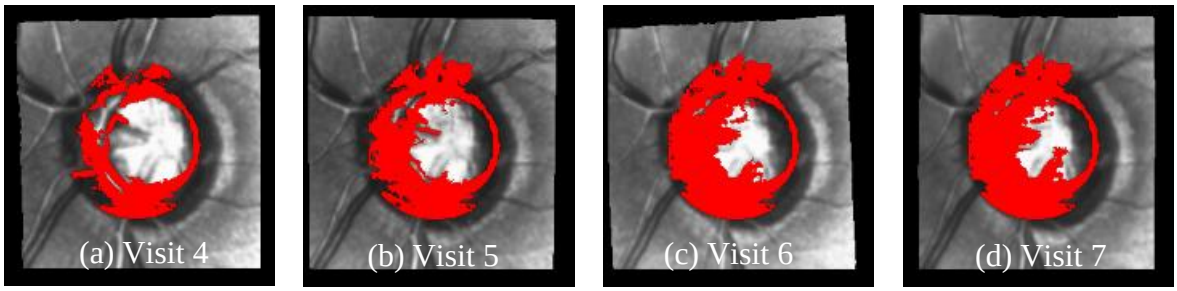
Figure 15 Detection rates of SIM and TCA on real clinical data

	SIM		TCA	
	n	% (CI)	n	% (CI)
Controls	2	10.0 (2.8-30.1)	3	15.0 (5.2-36.0)
Converters	22	73.3 (55.6-85.8)	16	53.3 (36.1-69.8)

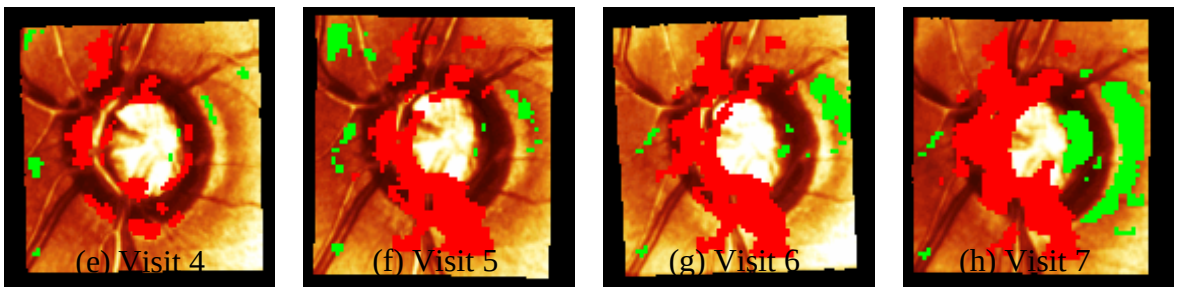
Table 1 The number of eyes determined to be progressing with Statistic Image Mapping (SIM) and Topographic Change Analysis (TCA) applied to real longitudinal HRT series: 20 normal subjects (controls) and 30 OHT patients that converted to a diagnosis of glaucoma by VF criteria (converters)

Although SIM is computationally intensive, by developing the algorithms in a low-level programming language and designing the code to reduce function calls and variable passing, the computer burden is not prohibitive. Analysis of a patient having 10 visits (30 images with 3 scans per visit) takes less than 3 minutes on a PC with a Pentium IV 3GHz processor. Shorter series take less time to analyse, but even a very long series of patient records could be handled on a standard PC during a patient visit. Further improvements to the computer code are likely to reduce this time further.

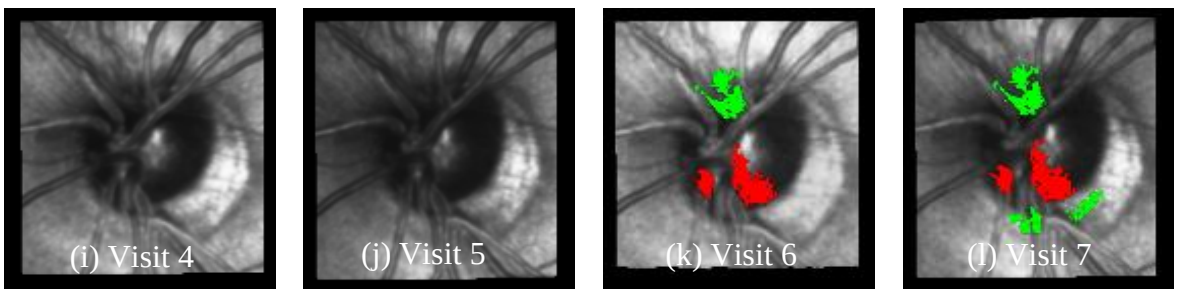
Case 1 SIM Output



Case 1 TCA Output



Case 2 SIM Output



Case 2 TCA Output

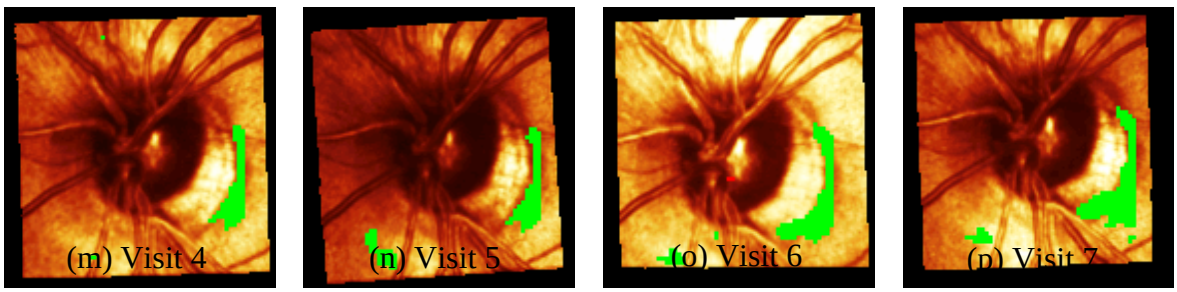


Figure 16 (a,b,c,d) Case 1 – OHT converter: the statistic image generated using SIM which has been overlaid on a mean reflectance image for visits 4 to 7 inclusive. (e, f, g, h) The TCA output (HRT Eye-Explorer software v1.4.1.0) corresponding to the same subject. Case 2 – OHT converter: SIM output (i, j, k, l) and TCA output (m,n,o,p). Note that two clusters have been flagged in the SIM analysis, since both are beyond what would be expected by chance as defined by the permutation distribution

3.5 Discussion

Reproducible scanning laser tomography images of the optic nerve head may present an objective method for measuring disease progression in glaucoma. This chapter presented and evaluated new statistical procedures for the analysis of these images. Techniques primarily developed for neuroimaging data were exploited and applied to longitudinal series of HRT images on a pixel by pixel level.

Serial analysis using trend analysis or statistical tests comparing baseline and follow-up images of stereometric parameters have been used to measure change to the ONH (section 1.2 & 1.3 provides a review). This thesis considers the hypothesis that these methods may be subject to similar inadequacies associated with using the global indices to summarize progression in VFs: chiefly loss of spatial information and poor sensitivity to identify the localized damage (Chauhan, Drance et al, 1990; Smith, Katz et al, 1996). This hypothesis is explored in detail using an array of stereometric parameters on real clinical data in chapter 5.

The computer simulation and analysis of real longitudinal HRT data provide evidence that SIM has better sensitivity at detecting localized change than the TCA method. This is achieved while reducing the number of false-positives flagged. The TCA analysis originally reported with results from computer simulation, but these were different to those reported here as they centered on a single superpixel rather than results across the whole image (Chauhan, Blanchard et al, 2000). They reported a high level of sensitivity and specificity in detecting change. When the technique was applied to real longitudinal data three confirmation tests and a requirement for a certain cluster size were required to lower the false-positive rate. A statistical adjustment (the Satterthwaite correction) is used in the TCA to correct for similarity (correlation) of the topographic height *within* a superpixel (Neter, Wasserman et al, 1985), but no real account was made for the multiplicity of testing across the whole image. The empirical solution to the problem of multiplicity of testing included the requirement for clusters of pixels to be above a certain size, based on observed series of normal subjects (Chauhan, McCormick et al, 2001). The results in this chapter

suggest that SIM is better equipped at handling false positive because it inherently corrects for the multiple comparison problem: handling this aspect of imaging data is one of the key features of the SIM approach.

SIM uses permutation testing: tailoring the analysis to the data itself without incorrectly assuming that topographic heights, across the whole image, follow the behavior of a random variable from a known probability distribution, or without reliance on some reference patient population database. Permutation methods are known to be both flexible and exact (Manly, 1991). Historic objections to the widespread application of permutation methods seem irrelevant with cheaper and faster modern computational resources.

An additional reason why SIM had a better diagnostic precision than the TCA technique in computer experiments is simply the use of the whole series of the data: the TCA method only ever uses the baseline images and three follow-up images. This may be reasonable when the follow-up is short, but when the available series of data lengthens beyond 4 visits this will result in considerable data redundancy. This is illustrated in Figure 14 when the difference between the two methods appears about half way through the potential follow-up of 10 visits. It is also interesting to note that there is no discernible difference between the power of the methods when episodic or sudden loss is specified (Figure 14). This aspect of the results is reassuring because our choice of pixel by pixel test statistic is essentially a rate (trend) parameter which might not be considered sensitive to detecting a ‘sudden change’. However, it has been previously reported, for threshold measurements in the visual field, that linear regression adequately identifies sudden change unless a series of data becomes very long (Crabb, Fitzke et al, 1999). In later chapters we show that the real advantage of using a rate parameter as this may provide clinically interpretable information once the technique has identified a significant region of change. Of course, there is no firm evidence about structural loss in glaucoma being either gradual or sudden, but it seems the new technique that we have described here will be sufficiently sensitive to both types of deterioration.

One limitation of SIM is that by definition it requires a minimum of four visits to detect progression. This may not suit some clinical circumstances: for example patients showing rapid change should not be denied therapy until such a time when sufficient data becomes available. There is therefore a need to develop event-based analyses which don't require a confirmation test to be used in cases when limited data is available. Another limitation of SIM might be in how it has been designed to detect progression. The technique detects progression based on the size of the largest cluster of active change. This may not be ideal for detecting diffuse change of low intensity or for detecting a number of small clusters. Detecting change based on the size of the largest cluster does however correspond with how glaucomatous ONH damages occurs; it is known that damage occurs with regional preferences (localised to individual sectors) depending on the stage of the disease (Airaksinen and Drance, 1985; Jonas, Budde et al, 1999). Readers should also be aware of the exchangeability assumptions made by SIM. This assumption requires the intra-test and inter-test variability to be the same.

The main value of SIM is in the output: it provides the clinician with a much needed, reliable method of visualising, quantifying and assessing rates of glaucomatous change in small localised areas in series of retinal images, rather than binary progression or stable classifications that rely on topographic summary parameters. In chapter 4 we explore assumptions made in assessing if change is significant and introduce an optimized strategy for detecting change. In chapter 5 the techniques are tested on a larger set of clinical data by comparison to summary measure (stereometric parameters) of the ONH.

4. SIM: optimizing technique for combining spatial extent and intensity of change

SIM as presented in the previous chapter provides an image of changing pixels, but also a probability value of global change, or deterioration in the image overall. This value was derived by comparing the largest cluster of active pixels to those which occurred by chance. The method measures the significance of the spatial extent of the glaucomatous damage. This provides a value for the global significance of change based on the patient's own data and the variability of the image series, while correctly accounting for the multiple comparison problem which occurs by calculating test statistics at each pixel location in the ONH. However, up till now it is assumed that glaucomatous damage is large by spatial extent. This begs the question, what if glaucomatous damage is small by extent but has high intensity (deep change)? What follows is a solution to this problem which uses a mathematical technique called a combining function. This technique is capable of detecting change which is either significant by spatial extent or intensity, or a combination of both. This chapter demonstrates using computer experiments that combining functions increase the sensitivity to detect change while maintaining the same false-positive rate.

4.1 Methods

What follows is a description of how the permutation framework can be modified to allow the technique to be more flexible in detecting change:

Combining the Intensity and Spatial Extent of Change

In this section we test a technique from functional MRI which provides a mechanism to assess both the intensity and extent of change in blood oxygenation levels (Poline, Worsley et al, 1997; Bullmore, Suckling et al, 1999). Specifically we exploit a recently developed solution which uses a permutation framework (Hayasaka and Nichols, 2004). This technique uses a simple mathematical method (combining functions) for flagging change based on either the area of damage or the intensity of

damage. The technique uses two partial tests: the cluster size statistic (chapter 3) provides a solution for flagging change which is large by spatial extent; and the maximum test statistic provides a solution for flagging change of high intensity. This maximum test statistic, referred to here as *T-max*, derives a probability value by comparing the intensity of change which occurred in the observed series, to those which occurred by chance (Nichols and Holmes, 2002). Computationally this is accomplished by comparing the maximum test statistic which occurs in the observed series of images (see Figure 10) to the distribution of maximum test statistics which occurs at each unique reordering. Figure 17(c&f) illustrate the maximum test statistic distribution. The computational details are explained in more detail in section 4.2. There are however limitations of using either method to detect change. The cluster size statistic does not account for the depth of glaucomatous damage. For example, a cluster of depressed and significant change which occurred during a patients follow-up of area 2000 microns^2 and with 200 microns excavation will by assigned the same probability value as a cluster the same size but with a deeper excavation of 1000 microns. Conversely, the maximum test statistic (T-max) does not account for the spatial extent of change. For example, a glaucomatous eye with significant change 1000 microns deep which occurred during a patients follow-up would be assigned the same probability value if the change was clustered in an area of 10 microns^2 or 500 microns^2. This point is illustrated in Figure 17; here 2 patient series are simulated using the simulation described in Chapter 2. In case 1, an 'unstable' or progression patients image series is simulated with high intensity and small spatial extent. In this case the observed intensity of change (T-max) is in the tail of the distribution (p=0.013); however, the observed largest cluster size does not appear to be significant (p=0.129). In case 2, change is simulated with low intensity and large spatial extent. Here the observed cluster size appears to be significant (p=0.028), but T-max does not (p=0.31). The figure illustrates the limitations of both techniques: the sensitivity of each technique depends on the nature of damage which has occurred.

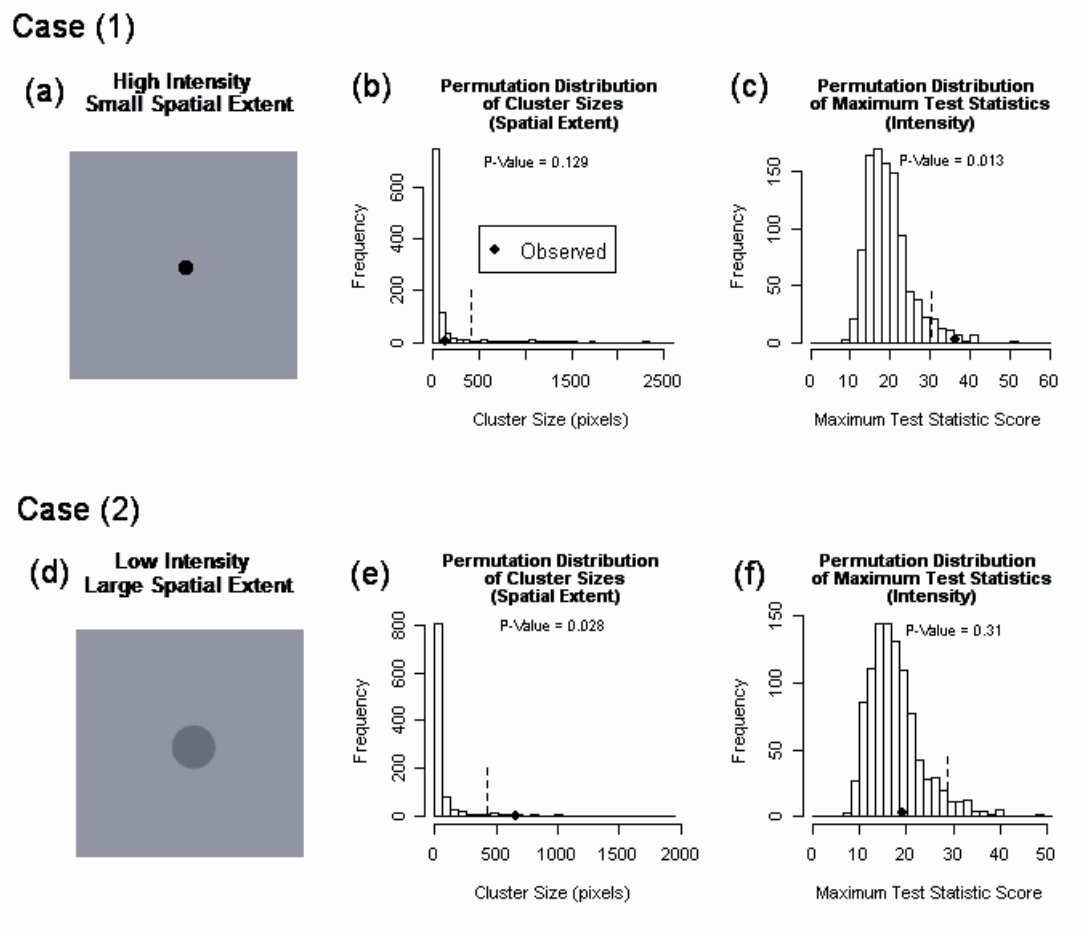


Figure 17 Detection of spatial extent and intensity of change. Longitudinal series of topography images were simulated, mimicking change over time in glaucomatous patients (see chapter 2). Two types of damage were simulated: in case 1 (a-c) damage of high intensity and small spatial extent and in case 2 (d-f) damage of low intensity and large spatial extent. Panels a & d are schematics illustrating the types of change applied. Panels b & e show the distribution of the largest cluster sizes i.e. the spatial extent of damage. Panels c & f show the distribution of maximum test statistics: this method provides a global probability value based on the depth (intensity) of topographic change. The distribution of the maximum test statistics for case 1 (c) indicates change of significant intensity (P = 0.013). Conversely in panel (e) the distribution of largest cluster sizes shows case 2 to have change of significant spatial extent (P = 0.028)

This section assumes that it is unknown if glaucomatous damage is significant by intensity or spatial extent, instead the objective is to develop a flexible solution for

detecting either type of change. This thesis describes and evaluates a solution using computer experiments (this chapter), before applying the technique to real data (chapter 5). What follows is a description of a combining function known as *Tippet.* Hayasaka and Nichols (Hayasaka and Nichols, 2004) reported that Tippet was able to detect change which is significant by either spatial extent or intensity. The Tippet function inputs the cluster size and T-max probability distributions and uses a simple mathematical function to determine which is most significant (see equation 1, page 61). Figure 18 shows the results of applying this equation to cases 1 and 2. The figure shows the new permutation distributions generated. In this instance the observed Tippet scores are in the tail of the distribution in both cases: case 1 ($p=0.012$) and case 2($p=0.004$). In section 4.3 computer experiments are devised to measure the specificity and sensitivity of this new technique.

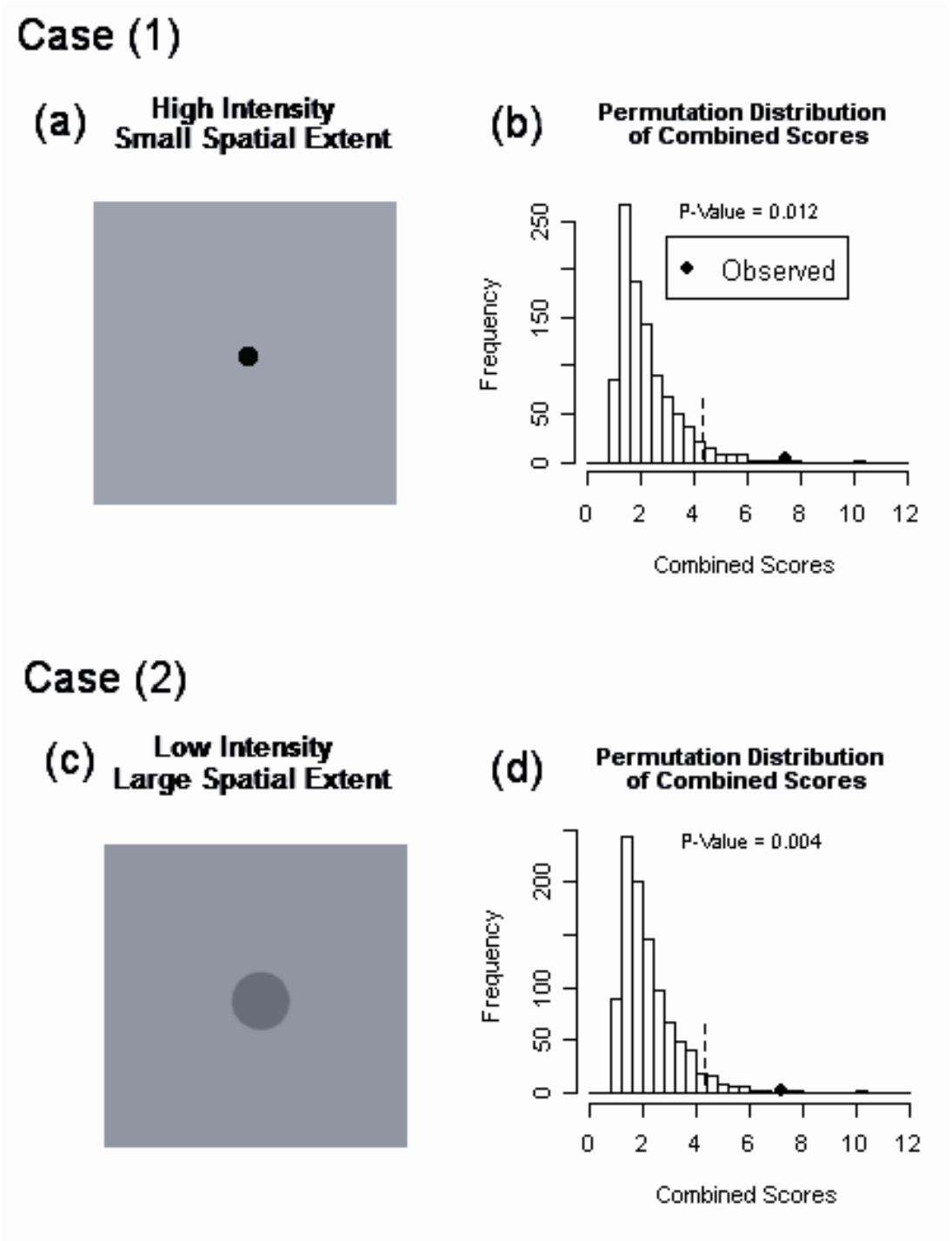


Figure 18 The Tippet combining function probability distributions. In case 1 (a-b) damage of high intensity and small spatial extent is simulated; In case 2 (c-d) damage of low intensity and large spatial extent is simulated (as shown previously in Figure 17). The observed combining functions score (b and d) show that significant change is detected for both types of change, cases 1 (P=0.012) and case 2 (P=0.004)

4.2 Computational paradigm

The aim of this section is to allow the reader to implement the computational paradigms behind the cluster size, T-max and Tippet probability distributions. This section appends the methods described in section 3.2:

Calculate the probability distribution of cluster sizes P^{S_Max}

The following paradigm calculates the distribution of largest cluster sizes (a schematic of the computational paradigm is shown in Figure 13):

1. Compute steps 1 to 4 as per section 3.2 in "permutation testing to threshold clusters" and define the distribution of maximum depressed clusters *S*. In this form S_1 represents the observed maximum depressed cluster and $S_{2\text{-}1000}$ represents the size of the observed maximum depressed clusters at each unique permutation
2. The probability value of the observed maximum cluster can be obtained by sorting a copy of *S* into ascending order and determining the rank of the observed largest cluster in the distribution. This probability value represents P^s_1
3. Repeat step 2 but instead compare the size of the largest cluster at each unique permutation $S_{2\text{-}1000}$ to the sorted copy of *S* and therefore calculate the rank and probability of each unique permutation $P^S_{2\text{-}1000}$
4. P^s_1 is the probability of the spatial extent of the observed cluster (used in chapter 3 to define change)

Calculate the probability distribution of maximum test statistic (T-max) P^{T_Max}

The following paradigm calculates the distribution of maximum test statistics (a schematic representing the computational details in shown in Figure 19).

1. Compute steps 1 and 2 as per section 3.2 in "permutation testing to threshold clusters" and define the pseudo test statistic *T_stat(i,j,k)*, where *i,j* are pixel locations within the topography image and where *k* represents each reordering. In this form *T_stat(i,j,1)* are the observed pseudo test statistic

values and $T\_stat(i,j,k_{2\rightarrow 1000})$ represents the pseudo test statistic at each reordering

2. Find the maximum pseudo test statistic in $T\_stat(i,j,1)$, whose slope was negative and let i,j be a pixel location bound within the contour line. Define this value T_max_1
3. Repeat step 2 finding the maximum pseudo test statistic at each reordering $T\_stat(i,j,k_{2\rightarrow 1000})$. Define this the distribution of maximum test statistics T_max
4. Calculate the rank and probability of the observed maximum test statistic by comparing its value to the distribution of maximum test statistics T_max
5. Repeat step 4 comparing T_max at each unique reordering calculating the rank and probability at each reordering. Define this the probability values P^t

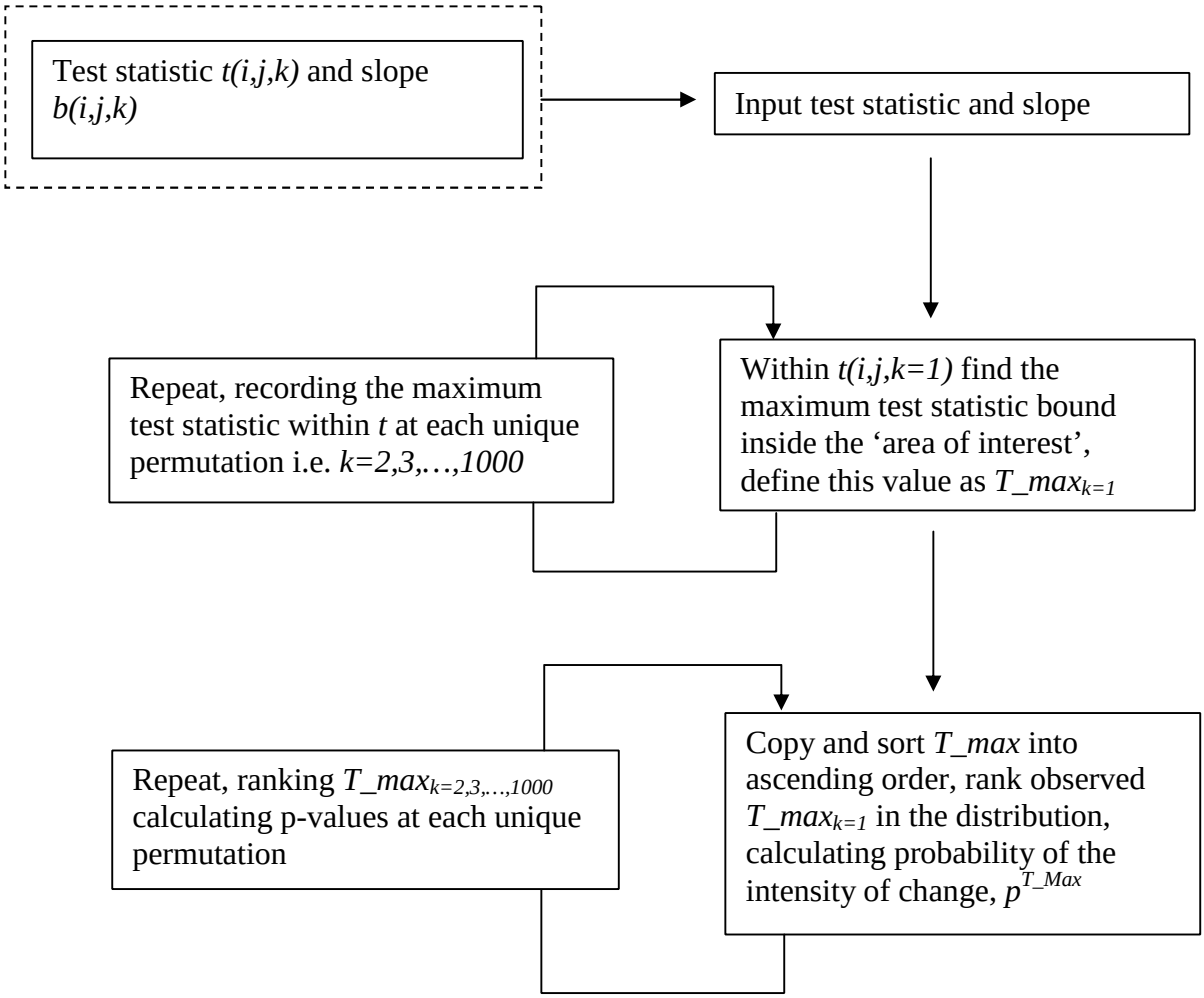


Figure 19 Schematic represents the computational details of the probability of the intensity of change *T_max*

Calculate the Tippet probability distribution P^T

The following paradigm calculates the Tippet probability distribution, a schematic showing the computational details is shown in Figure 20.

1. Apply the probability distribution of largest cluster sizes P^s and maximum test statistic P^t to the following equation, and calculate W^T for the observed case and for each unique reordering:

$$W_i^T = 1 - \min(\log P_i^t, \log P_i^s) \quad (1)$$

2. Compare the observed Tippet value W_1^T to the Tippet distribution W^T to find the rank and probability

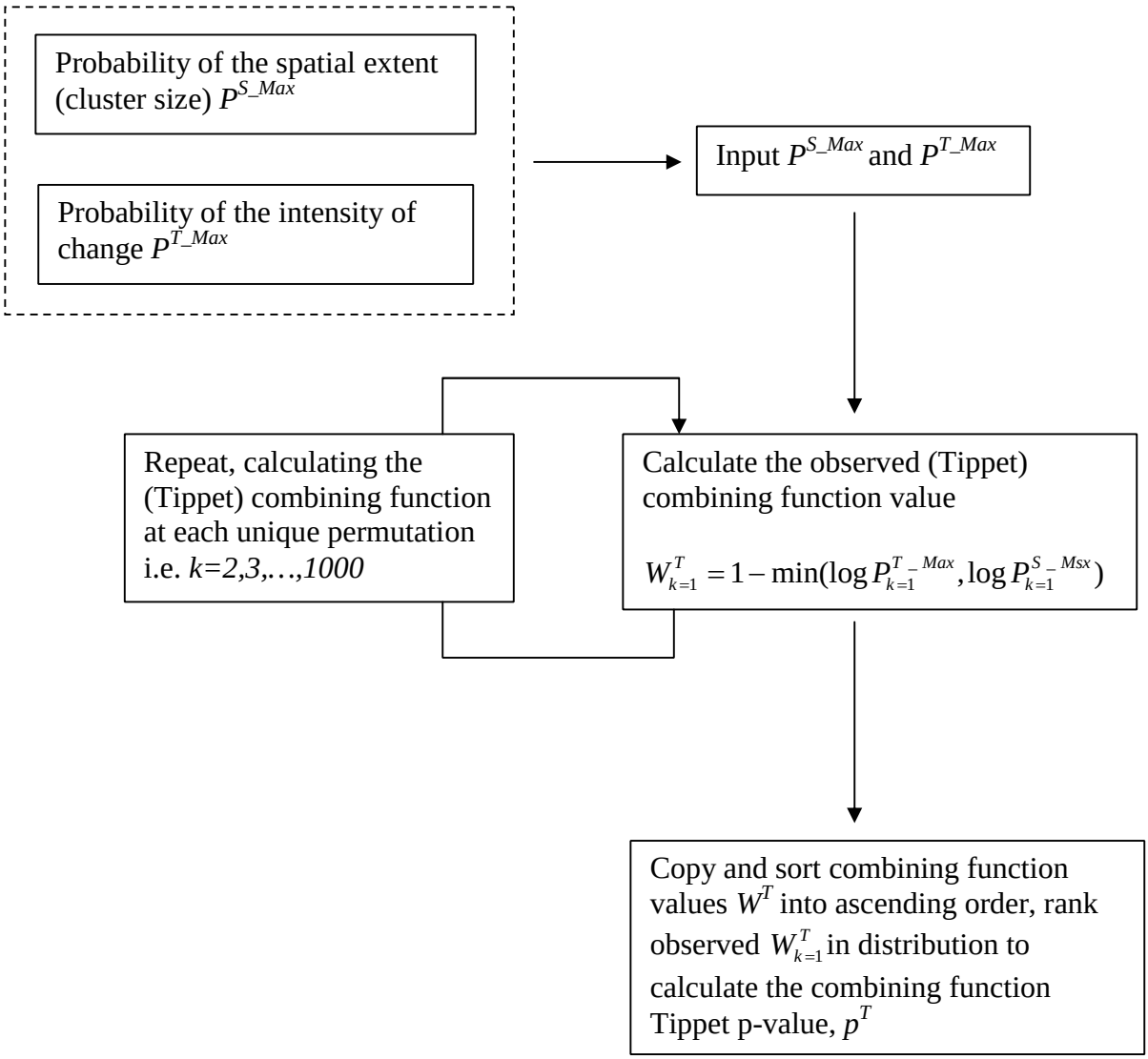


Figure 20 Schematic illustrating the computation of the Tippet combining function

4.3 Testing the combining function

The performance of the combining function is tested using the 'virtual patient' simulation described in chapter 2. The specificity is tested by generating 300 stable

virtual patients; three groups of 100 virtual patients are generated with a MPHSD of 15, 25 and 35µm respectively. The objective is to replicate the computer experiments which measured specificity in chapter 3 to compare the specificity of the combining function Tippet with the cluster size and T-max statistics. These techniques are applied as they are described in section 4.2. For each patient series the visit at which (false-positive) change is first detected is recorded.

The sensitivity of the combining function Tippet is tested using 200 unstable virtual patients; two groups of 100 virtual patients. The first group had gradual change simulated by applying a cumulative decay of 15µm per visit to a cluster of 240 pixels on the neuro-retinal rim. This group is designed to mimic change with high intensity and small spatial extent. Movement and Gaussian noise is then applied to each image series to simulate a MPHSD of 25µm. The second group had gradual change simulated by applying a slower cumulative decay of 5µm to a larger cluster of 640 pixels on the neuro-retinal rim. This simulation is designed to mimic change with low intensity and large spatial extent. Movement and Gaussian noise is again applied to simulate a MPHSD of 25µm. The same criteria for detecting change are used as in the specificity experiments. The follow-up visit at which change is first detected is recorded for each technique.

4.4 Results

In the computer experiments of 300 stable virtual patients at visit 10 in the follow-up series, the cluster size statistic detected 5% in all three groups of noise levels (MPHSD of 15, 25 and 35 µm), while T-max detected between 4% and 1%, and the combining function Tippet detected between 3% and 1%. Figure 21 shows a cumulative plot of the specificity at each noise level. The graph shows that the specificity of Tippet (solid line) is never worse than the specificity of the cluster size statistic (dotted line).

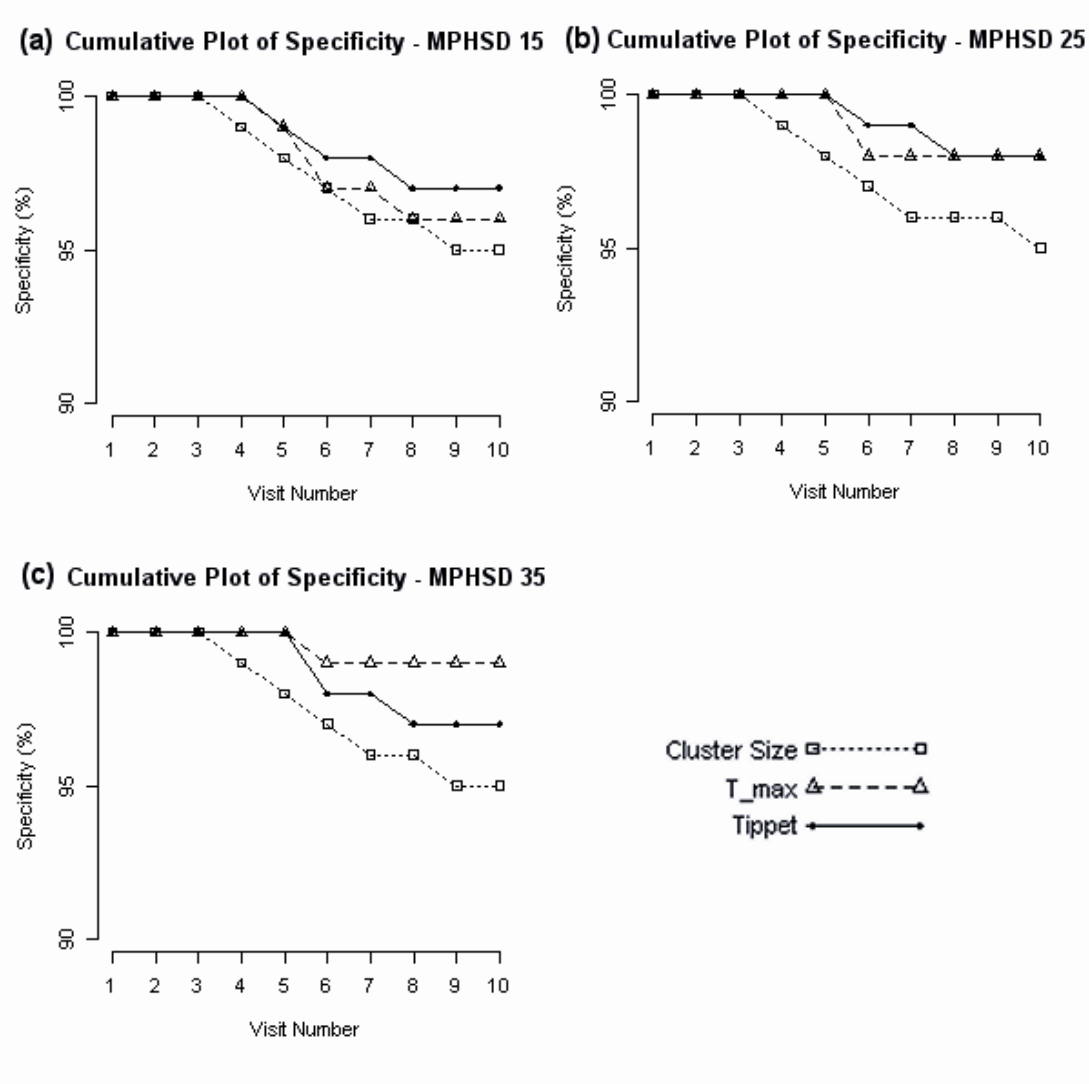


Figure 21 Computer simulation results comparing the specificity. Note that the specificity range is scaled between 90% and 100%. The specificity of cluster size, T-max and combining function Tippet are shown by simulating stable image series at different noise levels: (a) MPHSD 15, (b) MPHSD 25 and (c) MPHSD 35

In the simulation of progressing ('unstable') patients with high intensity and small spatial extent 'cluster size' detects 12% by visit 10, while 'T-max' flagged 99% of the patients. This is illustrated in Figure 22, here the cluster size statistic (dotted line) fails to detect change of high intensity, while T-max (dashed line) detects almost all of the series with high intensity change. However in the simulation of patients with low intensity and large extent this situation is reversed, 'cluster size' detects nearly half (52%), while 'T-max' detects only 11%. In these computer experiments the

combining function Tippet (solid line) detects 92% in the first sensitivity computer experiment (Figure 22a) and 46% in the second experiment (Figure 22b).

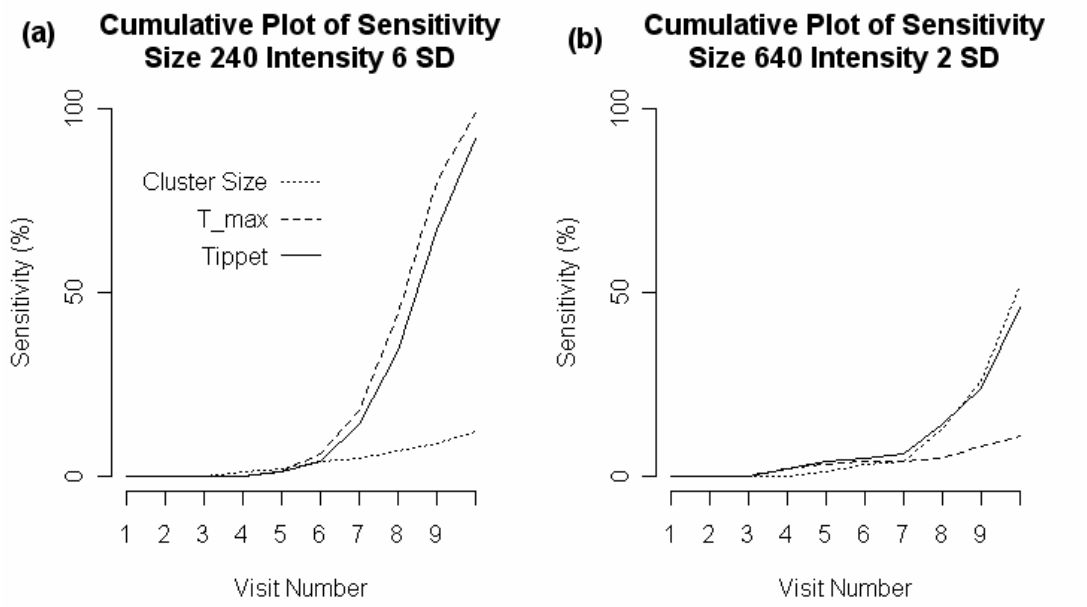


Figure 22 Computer simulation results comparing sensitivity. The sensitivities of cluster size, T-max and combining function Tippet are shown after simulating unstable patient series: (a) with high intensity and small spatial extent and (b) with low intensity and large spatial extent

4.5 Discussion

The results from these experiments suggest that when two different types of change are simulated the combining function Tippet is nearly as sensitive at detecting change as the best performing cluster size or T-max statistic. This is not surprising as the combining function inputs both these distributions and each element in the combining function distribution is simply the most extreme probability value of either cluster size or T-max which arises at each permutation. What is reassuring from the results is that these benefits are achieved without a reduction in specificity. This work is novel in that the TCA does not have a specific mechanism of incorporating depth of glaucomatous damage; the TCA flags significant change if a cluster of 20 superpixels are confirmed as significant over three consecutive visits. With the Tippet combining function SIM seems to be able to accommodate the

detection of change whether it is large by spatial extent or large by intensity. However, the results presented in this chapter are limited by the specific computer experiments which were performed.

In chapter 5 the Tippet combining function is incorporated into the SIM technique to optimize the detection of change in real patient data.

5. A comparison of SIM and global parameters

Currently an event-based analysis (comparison of most recent image with a baseline) using normalized stereometric parameters is used in the HRT software (HRT Eye Explorer v1.4.1.0) to help detect glaucomatous progression in series of images (see Figure 23).

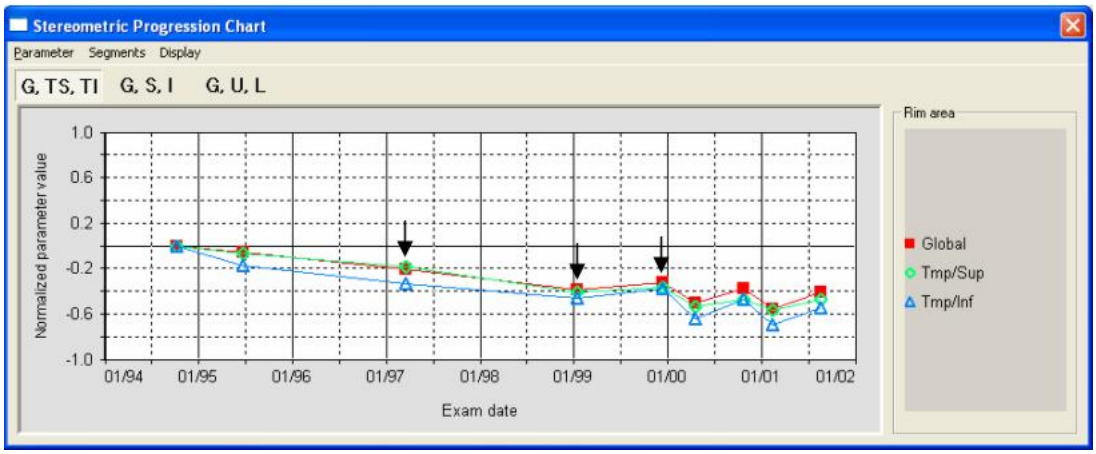


Figure 23 The parameter analysis available on the HRT software. The parameters are normalized to quantify the difference between normal controls and patients with advance glaucoma (see section 5.1 for details). Progression is confirmed if there is a difference of -0.05 or more on three consecutive occasions. In this example progression would be confirmed using global rim area (red line) at the visit corresponding to the position of the third arrow (Courtesy of Heidelberg Engineering)

In this chapter we compare SIM to this 'stereometric analysis' for structural progression in patients with glaucoma and ocular hypertension (OHT). Additionally, we demonstrate how SIM can be used as a clinically useful tool for visualizing and highlighting suspected localized areas of structural progression not detected by monitoring stereometric parameters during the follow-up period.

Stereometric parameters effectively condense all the information contained within a topography image into a single number. This is a highly data reductive process but by definition will not encounter the spatial correlation and multiple comparison problems discussed in section 1.3. However, as a result of this global parameters may suffer from low sensitivity in detecting localised areas of change, in the same

way that global indices in automated permetry fail to detect subtle glaucomatous progression. The objective of this chapter is to compare SIM against the stereometric parameters. A further objective is to evaluate if the combining function Tippet increase the sensitivity of SIM. The work in this chapter serves to test these hypotheses.

Some of the work in this chapter has formed a paper submitted to the British Journal of Ophthalmology (Patterson, Garway-Heath et al, 2005). The results in this chapter have also been presented in part at the American Academy of Ophthalmology meetings Chicago, USA, on Oct 15-18, 2005 and as a paper read before the UK and Eire Glaucoma Society Meeting in Nottingham on Dec 2, 2005.

5.1 Methods

This section first describes how the visual output of SIM was changed to allow interpretation of intensity of ONH progression. Then the 'stereometric analysis' available on the HRT is described, before describing how both techniques were compared.

Visualization of Topographic Change

Given a series of HRT images, SIM provides a map of areas of activity or progressive change that can be superimposed on the images. Thus far, change was flagged as significant or not with no information on the intensity of change which has occurred at a single pixel. In this chapter a 'change map' is presented which shows the statistically significant depressed change; also a scale-bar is generated to link colour to the total magnitude of change which occurred over the course of the follow-up, in microns, µm. Figure 24 provides an example of a patient's image series. The figure was generated by developing purpose written software using the C++ programming language to run in a windows environment (see Appendix A and B). The depressed change is colour coded from yellow, representing shallow excavation, to red, representing deeper excavation.

The change map is produced as follows: at each pixel the topographic change is quantified as the product of the rate of change times the time elapsed between baseline and the follow-up examination. As the spatial resolution of the method is so high, and the individual slope values at each pixel will be subject to error, we smooth the topographic change which has been extrapolated (from the rate and duration). This smoothing is done solely for visualisation; it does not affect the quantitative results returned by SIM. The topographic change values are smoothed using spatial convolution with a Gaussian kernel with a full width half maximum of 1 and size 3×3. The 'change map' is produced by showing depressed active change using colour lookup tables, with yellow through to red representing 'depressed' change. These maps are the first of their kind since they attempt to delineate both spatial extent and intensity or rate of change – both critical in evaluating structural progression.

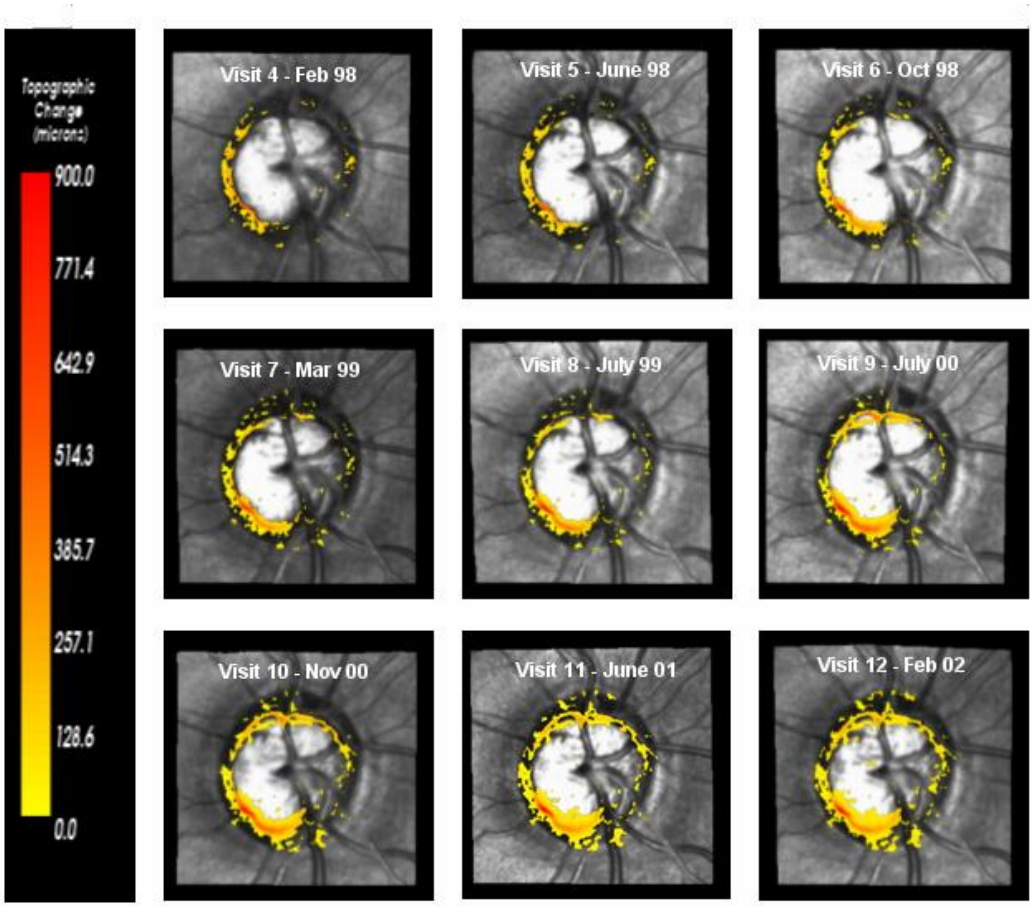


Figure 24 SIM 'change map' images overlaid on a patient's HRT image series from visit 4 through to visit 12. This OHT patient progressed to a diagnosis of glaucoma by visual field criteria (AGIS) during follow-up (note: a minimum of four visits is required to evaluate a 'change map'). The colour represents the depth of change which occurs; yellow through to red representing shallow through to deep change respectively

Stereometric Parameters

Previous studies have quantified the utility of stereometric parameters for monitoring progression (Mikelberg, Wijsman et al, 1993; Rohrschneider, Burk et al, 1994; Kamal, Viswanathan et al, 1999; Kamal, Garway-Heath et al, 2000; Tan and Hitchings, 2003). Section 1.3 provides a review of work in this area. In this study we consider five stereometric parameters: RA, RV, CSM, HVC and RNFL. Section 1.2 explains the features of the ONH that these parameters represent.

HRT software stereometric parameter analysis

HRT Eye Explorer software (v1.4.1.0) incorporates an event-based analysis comparing the most recent value for a stereometric parameter in an image series against the baseline value. Burk and colleagues (2000) classified subjects as having normal or early, moderate or advanced glaucoma by VFs (unpublished work). The averages for the various stereometric parameters were calculated for each group and are defined as P_{normal} and $P_{advanced.}$ Then the following equation is used to detect change

$$\Delta P = \frac{P_{follow-up} - P_{baseline}}{P_{normal} - P_{advanced}} \qquad (2)$$

$P_{follow\text{-}up}$ and $P_{baseline}$ are the values of the measured stereometric parameters for the patient. ΔP is essentially a coefficient of variation; simply, if the eye is stable at the follow-up visit ΔP is 0, and if a patient changes from normal to advanced glaucoma over follow-up ΔP is -1. The HRT literature accompanying the native software defines progression if ΔP is equal to or greater than -0.05 confirmed in 3 consecutive visits (www.heidelbergengineering.com). These values seem rather arbitrary but are worthy of investigation as they are recommended in the user manual of the HRT software.

Clinical Data

The techniques were applied to a group of test-retest HRT data and longitudinal HRT data. The data adhered to the Declaration of Helsinki, had local ethical committee approval and informed consent was obtained. All patients were attendees of clinics at Moorfields Eye Hospital, London.

Test-Retest Data

Seventy-four patients (43 OHT, 31 POAG) were recruited. OHT patients had an intraocular pressure (IOP) of ≥ 22 mmHg on two or more occasions, two initial reliable visual fields with AGIS score of 0, absence of other significant ocular disease that would affect visual field performance and age > 35 years. A reliable

visual field was defined as <25% fixation errors, <30% false positive errors and <30% false negative errors. POAG was defined as above, but inclusive of visual field defects quantified as AGIS scores of ≥ 1 on three consecutive reliable visual fields (AGIS, 1994; AGIS, 2000). The data were originally collected to evaluate the test-retest variability of the HRT and HRT II (Strouthidis, White et al, 2005; Strouthidis, White et al, 2005). Patients were not excluded by ONH appearance but were excluded by myopia greater than 12 dioptres of spherical power or any history of intra-ocular surgery. One eye was selected at random. In total five mean topographies were obtained by two experienced operators on two separate visits within a six-week period.

Longitudinal Data

Two hundred and seventeen OHT patients have been scanned regularly (median follow-up period 6 years, range 2.3 to 7.2 years). This study group is described in detail elsewhere (Kamal, Viswanathan et al, 1999; Kamal, Garway-Heath et al, 2000; Kamal, Garway-Heath et al, 2003). Fifty-two of the 217 OHT patients were categorized as progressing to POAG during follow-up based on a visual field analysis using AGIS, a global analysis for visual field progression used in several other studies (Kamal, Garway-Heath et al, 2003; Strouthidis, White et al, 2005; Strouthidis, White et al, 2005) or by pointwise linear regression (PLR) using PROGRESSOR software (Fitzke, Hitchings et al, 1996; Viswanathan, Fitzke et al, 1997; Viswanathan, Crabb et al, 2003). For the latter we defined progression with a highly specific PLR criteria (called 3 omitting); this is described in detail elsewhere (Gardiner and Crabb, 2002), and has been used in other studies (Nouri-Mahdavi, Hoffman et al, 2004).

Comparison

The SIM combining function Tippet and the SIM cluster-size statistic were both applied to the real data. The SIM combining function Tippet is compared with the stereometric parameters.

SIM and the stereometric parameter analyses were applied to the test-retest dataset to yield a false positive rate for each method. The image sequence was randomly reordered to compensate for any ordering effects in the study design. The techniques were first applied at 'visit' four and then 'visit' five in the series of images. The stereometric analyses were performed twice, first following guidelines from the HRT literature with ΔP set to 0.05 (see "How to interpret progression", www.heidelbergengineering.com) and second with ΔP varied so that the false-positive rate matched that of SIM.

SIM and the stereometric analyses were then applied to those patients in the longitudinal data categorized as having progressed to POAG. In order for the sensitivity (true positive rate) of all the techniques to be compared, it is important that the false-positive rate of each technique is identical. It is not meaningful to compare the sensitivity of different tests unless their respective specificity is matched; this is done to avoid one technique flagging a greater percentage of patients as progressing to glaucoma by chance, and follows an approach adopted by (Ford, Artes et al, 2003) when examining the diagnostic precision of the HRT. ΔP (see Equation 2) was altered for each individual stereometric parameter to anchor the false-positive rate to that yielded by SIM in the test-retest data (2.7%; 2 out of 74 test-retest patients progressing). Progression was recorded in the longitudinal data if the limit for ΔP was exceeded in three consecutive visits. For each technique the time to progression in the longitudinal data was recorded and plotted using Kaplan-Meier curves. The log-rank test was used to compare the time to progression between techniques; this is a non-parametric method for testing the null hypothesis that the detection rates of each technique are samples from the same population. This overcomes the problem of comparing the methods at a single time-point. The log-rank test is used to assess the significance of any differences in times to detection of progression.

5.2 Results

Test-retest data

Both SIM Tippet and SIM cluster-size falsely identified 2 (2.7%) patients as having progressed. Using the stereometric parameter analysis recommended in the HRT literature, RA (falsely) flagged 6 (8.1%), RV 19 (25.7%), CSM 19 (25.7%), HVC 22 (29.7%) and RNFL 16 (21.6%) patients as progressing.

Longitudinal data

Figure 25 shows Kaplan-Meier curves comparing the SIM Tippet and the SIM cluster-size statistic. Tippet performs better, the median time for Tippet to detect progression was 3.9 years, whereas cluster size was 6.8 years. By the end of the follow-up Tippet had flagged 78.8% as progressing, whereas cluster-size flagged 51.9%. The log-rank test comparing both showed that Tippet detected change significantly earlier ($P<0.001$).

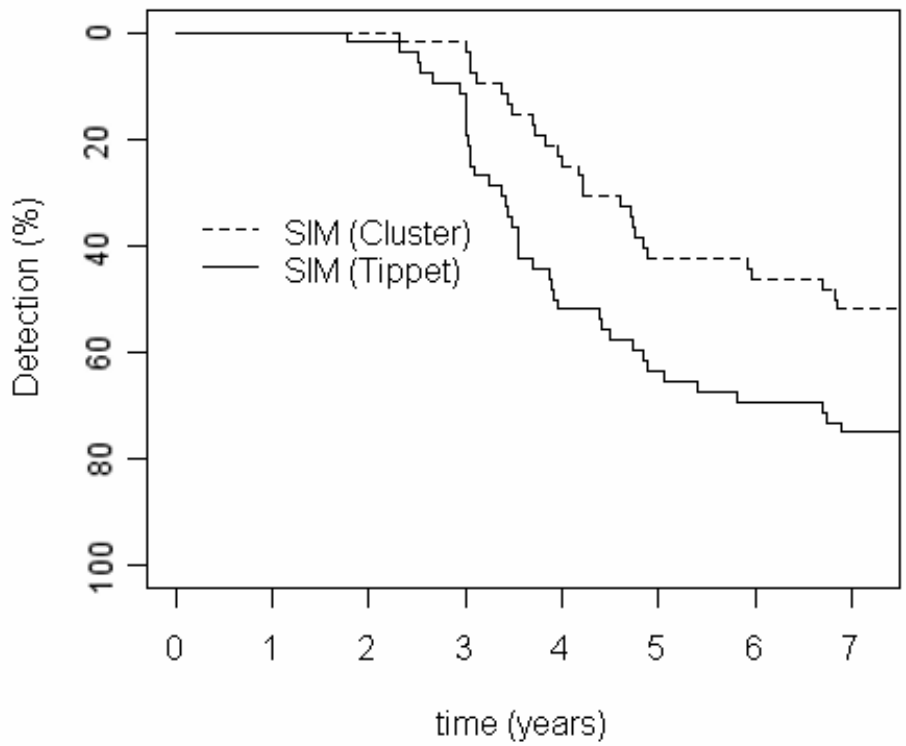


Figure 25 Kaplan-Meier plots comparing the performance of the SIM Tippet and the SIM cluster size statistic in 52 patients that have been defined as progressing based on visual field criteria. The results show that SIM Tippet flags change earlier than the SIM cluster-size statistic

The false-positive rate of stereometric analyses was anchored by varying ΔP (see Equation 2) so each parameter flagged 2 of the 74 patients (2.7%). Figure 26 shows Kaplan-Meier curves which compare SIM Tippet and the stereometric parameters. With the false-positive rates anchored, the stereometric parameters failed to detect 50% of the patients as progressing over the course of the follow-up. By the end of the study RA detected only 25.0%, RV 0%, CSM 40.4%, HVC 1.9% and RNFL 5.8% as progressing. The log-rank test comparing SIM Tippet to each of the parameters showed that in each case SIM Tippet detected change significantly earlier ($P<0.001$).

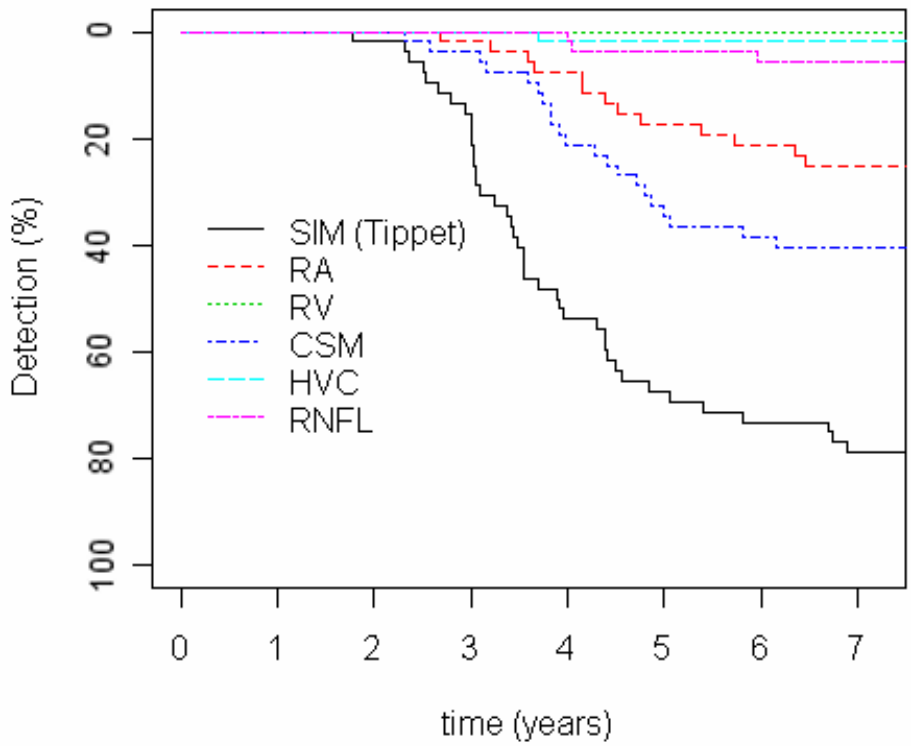


Figure 26 Kaplan-Meier plots comparing the performance of stereometric parameter analysis against SIM Tippet in 52 patients that have been defined as progressing based on visual field criteria. The comparison is made with the false positive rates anchored as described in the methods. This provides overwhelming evidence that SIM detects more true progression events and significantly earlier than the stereometric parameter analysis

Cases 1 to 3 in this section (see Figure 27 through to Figure 29) show illustrative examples of SIM 'change maps' for 3 patients from the longitudinal dataset. The normalized stereometric parameters (following Equation 2) were plotted for the patients over the follow-up period as they look in the output of the HRT software and the visual field changes are also shown in the figure. All three patients had OHT and progressed to a definition of glaucoma by two visual field criteria, AGIS and PLR, over the course of the follow-up. Figure 27 shows a patient where SIM detected a focal loss in the temporal sector. SIM detected significant change after 3.0 years and the only stereometric parameter that detected change (CSM) did so only later, after 4.0 years. In Figure 28 SIM detected diffuse loss occurring with the highest intensity at the inferior and superior poles. SIM detected change first after 2.5 years and none of the stereometric parameters detected change. In Figure 29 SIM detected diffuse change with a large intensity between the inferior and temporal regions. SIM detected change first after 4.3 years but again the stereometric parameter analyses failed to detect change.

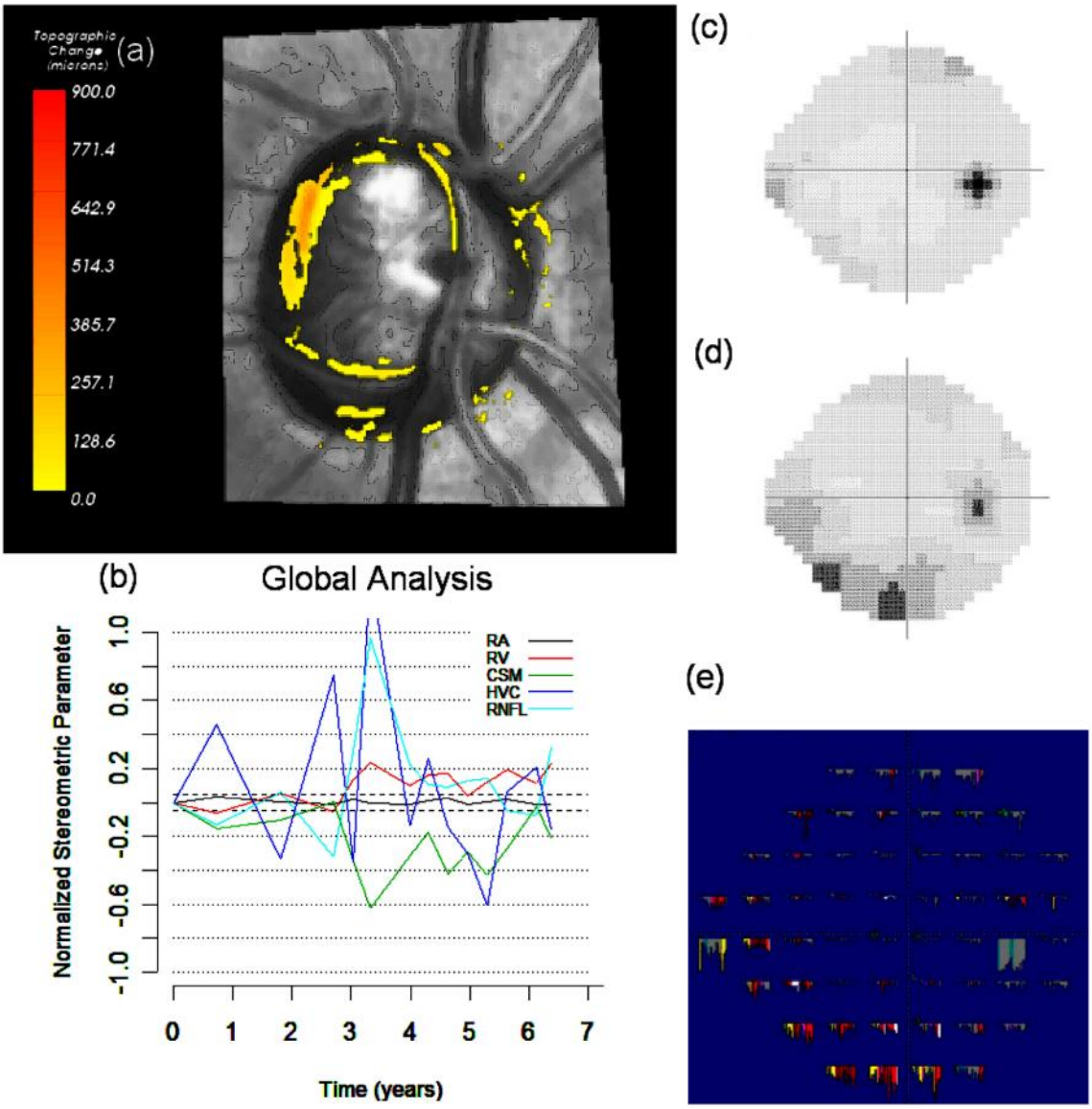


Figure 27 Case 1: An OHT patient who converted to glaucoma based on visual field testing (AGIS criteria) and PLR during the follow-up period. (a) A 'change map' with the scale bar showing topographic change (yellow to red representing optic disc deepening). The area of statistically significant change detected by SIM is overlaid onto HRT reflectance images. Change occurred mostly in the temporal superior position of up to ~450 microns (a rate of loss of ~70 microns per annum). Stereometric analysis (b): the corresponding normalized stereometric parameters are plotted for each patient. The ± 5% deviation line is represented by the dashed lines. CSM detected change after 4.0 years whereas the other measures did not detect change. (c) A greyscale of the baseline visual field, (d) a visual field obtained at the end of the follow-up period. (e) An image from PROGRESSOR showing the cumulative output from pointwise linear regression at each test point in the visual field. Each test location is shown as a bar graph in which each bar represents one test in the series. The length of the bars represents the depth of the defect. The colour of the bars relates to the p-value summarizing the significance of the regression slope with colours from yellow to red to white representing p-values of low to high statistical significance. Whereas stable points with low sensitivity are displayed as long bars and grey represent flat non-significant slopes. The patient's visual field shows progression occurring mostly in the lower nasal area

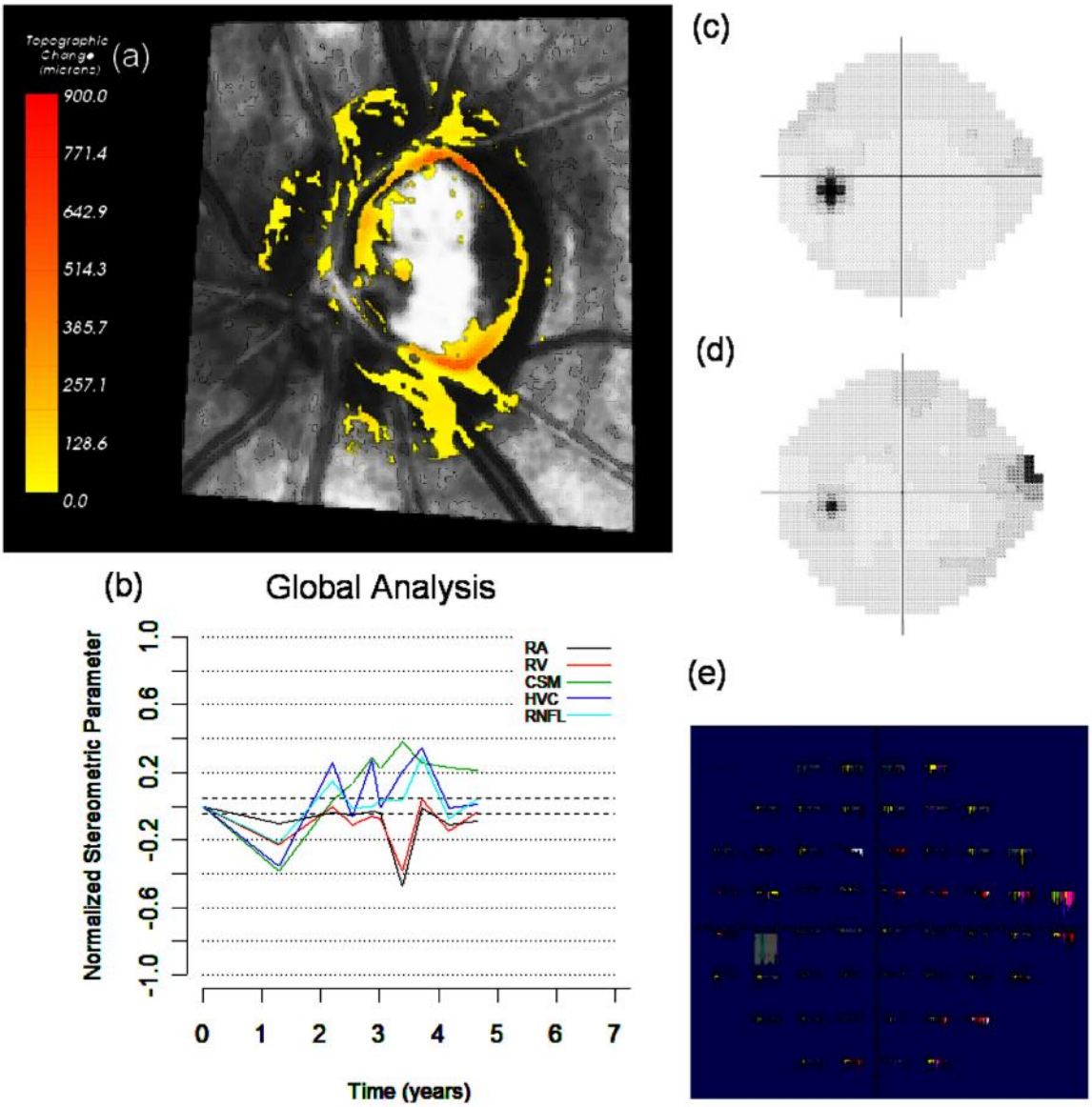


Figure 28 Case 2: An OHT patient who converted to glaucoma based on visual field testing (AGIS criteria) and PLR during the follow-up period. (a) A 'change map': change occurred mostly in the inferior and superior poles of up to ~850 microns (a rate of loss of ~180 microns per annum). SIM detected change after 2.5 years. (b) Stereometric analysis: none of the parameters detected change. (c) The baseline visual field, (d) a visual field obtained at the end of the follow-up period. (e) Output from PROGRESSOR. The visual field grey scales look remarkably similar but PROGRESSOR shows modest, but highly significant, superior paracentral arcuate progression

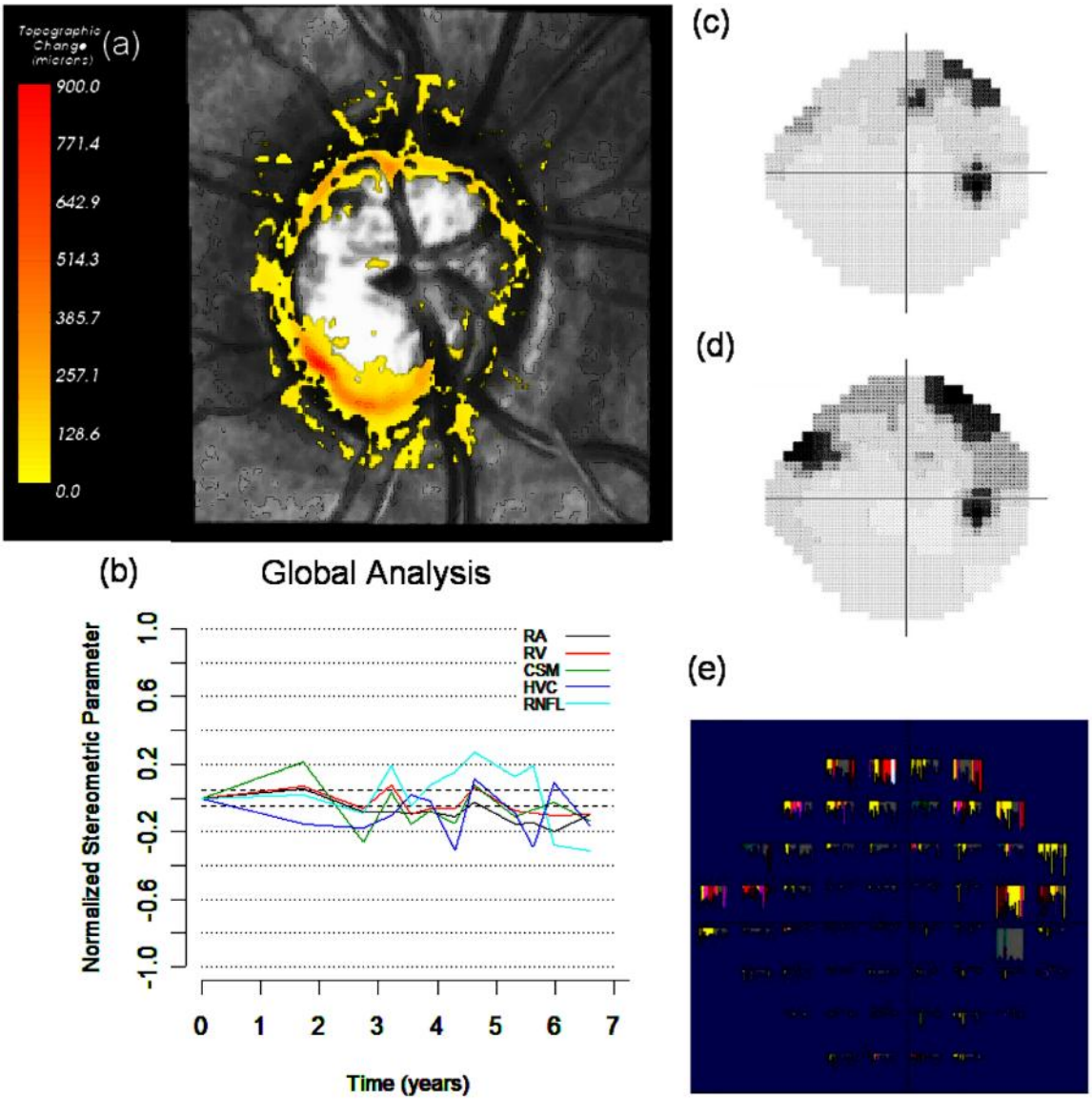


Figure 29 Case 3: An OHT patient who converted to glaucoma based on visual field testing (AGIS criteria and PLR) during the follow-up period. (a) A 'change map': change occurred mostly in the inferior temporal sector of up to ~850 microns (a rate of loss of 130 microns per annum). SIM detected change after 4.3 years (b) Stereometric analysis: none of the parameters detected change. (c) The baseline visual field, (d) a visual field obtained at the end of the follow-up period. (e) Output from PROGRESSOR. This patient has extensive visual field progression in the upper nasal to upper temporal areas

5.3 Discussion

The main finding of this chapter is that maps of structural change, in this case derived from SIM, are better at detecting progression than current statistical approaches for monitoring stereometric parameters. Moreover, the analysis of stereometric parameters suggested in the current HRT proprietary software exhibits very poor specificity, which suggests it has limited clinical utility. SIM had the highest true-positive rate, detecting significantly more patients as progressing when compared to the stereometric parameters, detecting 50% of the patients after 3.9 years; the parameter analysis failed to detect 50% of the patients over the entire course of the follow-up.

This work showed that SIM Tippet increases the sensitivity of the technique whilst flagging the same number of false-positive events when applied on real clinical data. Moreover, incorporating information on the intensity of change improves sensitivity.

SIM provided a useful alternative for detecting and visualizing progressive damage in the ONH as compared to the data reductive process of simply monitoring stereometric parameters over time. As discussed in chapter 1.3 the visualization of change is critical in the management of glaucoma where experienced clinical observation of the ONH, remains paramount to the diagnosis of disease progression. This is illustrated in cases 1-3 where a range of structural damage, varying in extent and intensity, is delineated by SIM but is not detected (or at best detected later in the follow-up) by the monitoring of the stereometric parameters for change. In Figure 27 the extent and intensity of structural damage is moderate, while a notable visual field defect is detected by pointwise linear regression. In Figure 28 a modest visual field defect is detected but the structural change is extensive. In Figure 29 both the visual field and structural defects are extensive. The main advantage of SIM may be as a method which provides a ‘change map’ flagging areas of optic nerve head damage resulting from glaucoma. By using the visual output, it is possible to

quantify the rate of loss (microns per annum). It is hoped that this may be a valuable tool in assessing a patient's response to treatment.

Current evidence has been interpreted as suggesting that visual field loss and structural progression can occur independently, or at least may not be simultaneously detectable over the course of a follow-up period (Artes and Chauhan, 2005) This hampers the experimental design of any study that uses visual field changes as the gold standard for glaucomatous progression, and this limitation applies also to this study. Further work will apply SIM to larger datasets with the hope of providing rates of morphological loss for normal subjects followed over time, OHT with stable and unstable visual fields, and for the glaucomatous population at different stages of the disease. Only after such studies will it be possible for SIM to become a clinical standard by which structural change in longitudinal series of optic disc images can be assessed. In the meantime SIM offers a new way of looking at structural change beyond the use of summary measures of the ONH, such as the stereometric parameters.

6. Deconvolution: Improving the repeatability of ONH images

This chapter examines a technique to improve the repeatability of confocal scanning laser tomography. This chapter diverges slightly from the main theme of the thesis thus far in developing techniques to detect glaucomatous progression. However, section 1.1 makes the point that the ability of a technology to measure small changes to the ONH is a function of the reproducibility of the technology. Therefore, any improvement in the repeatability of ONH images will improve longitudinal analyses.

A recent large population study using scanning laser tomography (HRT II) reported that satisfactory images (defined as average repeatability of topographic height > 68 microns) could not be obtained in 10% of a normal elderly population (Vernon, Hawker et al, 2005). A test-retest study of the HRT in an ocular hypertensive (OHT) and glaucomatous population indicated similar results, 11% > 68 microns (Strouthidis, White et al, 2005). This simply means that with HRT image acquisition as it stands a considerable amount of data is simply lost or disregarded. Computer simulations reported in chapter 3.3 demonstrate the intuitive point that improving the repeatability of image series will increase the sensitivity of techniques in detecting glaucomatous damage (both SIM and TCA).

To recap and complement the introduction to confocal scanning laser tomography in chapter 1.2, the technology is in essence a special application of confocal microscopy which regards the ocular fundus as the object. The resolution of the raw three-dimensional images obtained by the technology is roughly 'pencil' shaped, with a lateral resolution of ~10 microns in diameter and a depth resolution of ~300 microns. A topography image is generated by determining the position of peak reflectance in the confocal stack at each pixel. Reflecting layers within the retina (pigment epithelium, nerve fiber layer, inner limiting membrane, etc.) vary in thickness from between 200 to 500 microns. Due to the limited depth resolution (300

microns) of the technology, the topography image can be thought of as the position of the mean depth of the interfaces between these layers (Gaida, 1990).

In an ideal imaging system, the image acquired would be an identical scaled version of the true object being imaged. Image restoration algorithms are a collection of image processing techniques used to estimate the true object of interest given the image obtained and information of the nature of the blurring and noise in the imaging system (Gonzalez and Woods, 2002). Image restoration algorithms can be classified as linear and non-linear. Non-linear techniques, based on computationally intensive methods have increasingly gained acceptance, with (Jansson, 1997) finding they deliver superior results to linear methods in a broad spectrum of applications. Maximum-likelihood (ML) deconvolution is an example of these techniques: this estimates both the true image and the blur from the degraded image using partial information about the imaging system (Dempster, Laird et al, 1977; Holmes, 1992). Holmes and colleagues (Holmes, Bhattacharyya et al, 1995) were the first to apply the techniques to images obtained in confocal microscopy. In this study ML Blind Deconvolution is applied to series of images obtained from a test-retest study of the HRT. Our motivation is to investigate if this technique can improve the repeatability of topographic height measurements and hence increase the utility of the HRT as a technique to assess the ONH.

The work in this chapter has formed a paper which has been accepted and is currently in press at Investigative Ophthalmology and Visual Science (Patterson, Garway-Heath et al, 2006). The results within this chapter have been presented at Image Morphometry and Glaucoma in Europe Meeting, Milan, Italy on April 4-5, 2005.

6.1 Methods

To recap MPHSD is a gauge of the variability of each pixel height measurement across the three topographies used to make up the mean topography (section 1.2). In this chapter we use MPHSD to report intra-scan (within-scan) repeatability. Previous studies have used this metric to evaluate the repeatability of the technology in normal subjects and glaucoma patients (Chauhan, LeBlanc et al, 1994; Rohrschneider, Burk et al, 1994). It has been show that MPHSD is influenced by lens opacity, age and degree of astigmatism (Strouthidis, White et al, 2005). In this study MPHSD is also used to report the repeatability of mean topographies, referred throughout this chapter as inter-scan repeatability.

Confocal scanning laser tomography has known limitations; for example, although the optical setup is designed to reject most light from outside the focal plane, it by no means rejects all of it and an out-of-focus haze remains (Pawley, 1990; Vieira, Manivannan et al, 1999). The resolution in the confocal images is higher in the *x* and *y* directions compared to poorer resolution along the optical axis (*z*-axis). The resolution obtained is also limited by the optics of the eye, aberrations are generated by the cornea and the lens (Artal, Guirao et al, 2001). These limitations result in axial smearing: for example, if a spherical point object is being imaged with constant reflectivity properties, the resulting image obtained will appear elongated in the z-axis, therefore the data obtained will have an 'American football' type appearance. Another limitation of the technology is that the detector in the optical setup, is prone to Poisson noise primarily due to quantum variations in the number of photons recorded (Nourrit, Vohnsen et al, 2005). This noise obscures real data and randomly creates impossible features such as high intensity data only one pixel in size (Holmes, Bhattacharyya et al, 1995). A discussion on the principals and limitation of imaging systems is given by (Goodman, 1996).

Deconvolution is an example of an image restoration algorithm that models the imaging system with:

$g(x) = h(x)*f(x) + n(x)$

where $g(x)$ is the image obtained, $f(x)$ is the true image, $h(x)$ is the point spread function (PSF), $n(x)$ is noise, and where x is 3D Euclidean space. The PSF describes how much a single point source of light is spread over the focal planes. The image formed by a system g(x) is a combination of the PSF covering all the area where the geometrical image should be (Figure 30 provides a real world example of this phenomena). This process, is called convolution ("*" indicates convolution). The wider the PSF the more blur the image will contain. In CSLT the PSF is assumed to have a three dimensional hour-glass shape, orientated along the optical axis and of highest intensity in the central 'narrow' area (Holmes, Bhattacharyya et al, 1995). The objective of image restoration algorithms is to obtain an estimate of the true image $f(x)$ given the image obtained $g(x)$. In classical linear deconvolution, the PSF $h(x)$ is assumed to be known explicitly prior to the procedure. A long list of these techniques is available, such as the inverse filter and Wiener filter (Andrews and Hunt, 1977; Gonzalez and Woods, 2002). Unfortunately, in our situation the blur $h(x)$ is unknown, along with much information about the true image $f(x)$. Blind deconvolution refers to the task of separating two convolved signals $f(x)$ and $h(x)$, when both signals are either unknown or partially known.

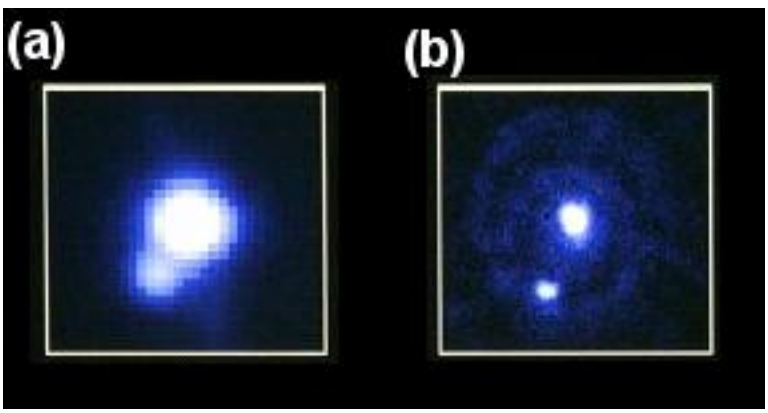


Figure 30 Images taken of Pluto (www.nasa.org). (a) An image of Pluto taken from an earth based observatory in Hawaii, in this image it is difficult to distinguish Pluto's moon 'Charon'. (b) An image of Pluto obtained from the Hubble Space Telescope, in this image it is possible to differentiate Pluto from its moon. These two images illustrate the blur induced by the atmosphere

ML Deconvolution

The ML estimation is a mathematical optimization strategy designed to produce the best guess of the true data which has been corrupted by random noise (Van-Trees, 1968). Here the ML paradigm is an adaptation of the 'Richardson-Lucy' optimization strategy (Richardson, 1972; Lucy, 1974). The ML deconvolution approach is however known to fail unless strong constraints can be applied to the properties of the PSF. Holmes and colleagues (Holmes, Bhattacharyya et al, 1995) published suitable constraints for data obtained by confocal microscopy, using assumptions about the shape of the 'hour-glass' PSF and optimized for the mathematical nature of the noise present. Previous studies evaluated ML deconvolution in confocal microscopy on simulated and real data (Kempen, Vliet et al, 1997; Rooms, Philips et al, 2003). In this study ML deconvolution is applied using software developed by Holmes and colleagues, with the commercially available suite AutoDeblur (version 9.3.6) [Autoquant, NY, USA].

The deconvolution software requires input of the optical setup and image medium to allow it to approximate and iteratively constrain the solution of the PSF. The numerical aperture of the lens (0.08 mm), wavelength of the laser beam (680 nm), refractive index of the medium being imaged (assumed to be close to water, 1.333) and spacing of the image obtained is input. For a 10º scan, the *x* and *y* spacing are 11.4 microns. With the HRT, the depth of scan can be varied from 2 mm to 4 mm whilst always acquiring 32 scans. The *z* spacing must therefore be set appropriately (as the scan depth ÷ 32).

Clinical Data

The techniques were applied to the test-retest dataset described in section 5.1. To recap briefly (Strouthidis, White et al, 2005) originally collected the data to evaluate the test-retest variability of the HRT and HRT II. The images were obtained over a six-week period in two visits to the clinic. On the first visit, three mean topographies

were obtained and in the follow-up visit a further two were obtained. All subjects had previous experience in scanning laser tomography having been imaged at least three times previously.

Comparison

This study investigated whether ML deconvolution would result in an improvement in intra-scan (within-scan) and inter-scan (between-scan) topographic height measurement repeatability. For this objective, 40 patients' image series (HRT 'Classic') were randomly selected from the test-retest dataset.

The association between lens opacity and the effect deconvolution had on repeatability was investigated. The lens opacity was measured subjectively using the Lens Opacity Classification System (LOCS) III grading system. Nuclear opalescence (NO), nuclear colour (NC), posterior subcapsular (PS) and cortical (C) scores were recorded.

Measuring intra-scan repeatability

The mean topography on the first visit was obtained and the MPHSD was recorded from the HRT software (version 2.01b). Three confocal stacks used to generate the mean topography images were deconvolved. The images were then re-input into the HRT software and three single topographies were calculated. It was observed that deconvolution had induced an artefact at the border of the single topographies. This is a known phenomenon, Gonzalez and colleagues (Gonzalez, Woods et al, 2004) suggest using a edge taper function to blur the edge of this image to minimize this effect. As a local spatial filter may affect the result, the artefact was removed using an image processing erosion algorithm (Gonzalez and Woods, 2002). The mean topography was then calculated and the MPHSD was recorded for comparison with the pre-processed data.

Measuring inter-scan repeatability

Three mean topographies were randomly selected for each patient. To measure the repeatability across scans the mean topography images were input into the HRT database as single topographies. The proprietary image registration algorithms spatially aligned the mean topographies and the computed MPHSD was recorded. The nine confocal images associated with the three mean topography images were deconvolved. Nine single topographies were then generated and the edge artefact was removed as before. Three mean topographies were generated and input into the HRT software as single topographies and the MPHSD quantifying the inter-scan repeatability was recorded.

A paired non-parametric (Wilcoxon) test was performed to determine whether the average improvement in repeatability was statistically significant because the distributions of differences in MPHSD had a positive skew. Spearman's rank correlation test was used to detect whether there was an association between the average MPHSD and the improvement in repeatability before and after deconvolution. The association between improvement and lens opacity was also quantified in this way.

6.2 Results

An example of a pre-processed and post-processed confocal image from different viewing perspectives is shown in Figure 31. The raw confocal image is in the left-hand column, while the deconvolved image is the right-hand column. Figure 31(c & d) compares the maximum reflectance images from the *xz*-plane, this viewing perspective can be thought of as a side elevation of the 3D confocal stack. The reduction in axial smearing is apparent after deconvolution. An individual slice in Figure 31(e & f) suggests a reduction in high frequency noise. Two *z*-profiles, before and after deconvolution, are plotted in Figure 31(g & h). The two *z*-profiles represent a pixel located in the neuro-retinal rim (marked by an arrow in Figure 31a), an area of the image that typically has low light reflectance. In this patient with lens opacity of NO = 2.7 and NC = 2.2 (as marked subjectively on a scale between 0 and 6), a high amount of noise is seen in the raw z-profile, while the deconvolved *z*-profile shows low noise. Deconvolution appears to improve the resolution, most apparent along the optical (z) axis, and reduces high frequency noise.

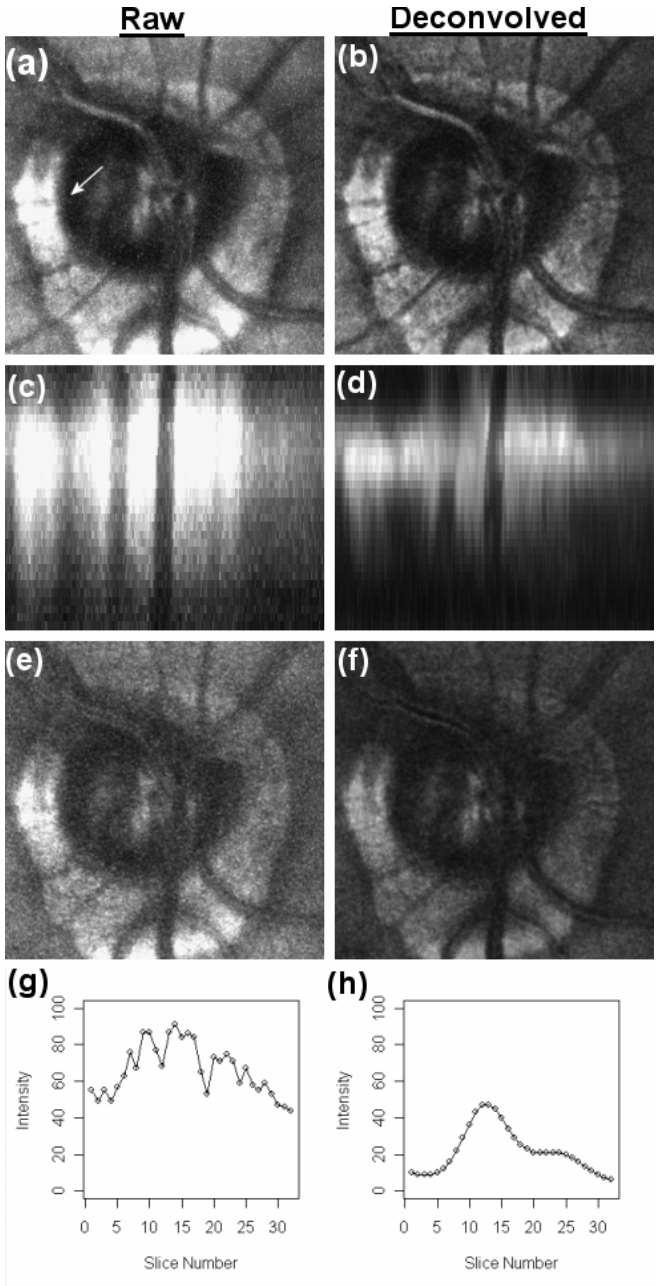


Figure 31 The raw confocal stack of optic nerve head acquired by HRT is on the left-hand column and the confocal stack after 30 iterations of ML deconvolution is on the right-hand column. The maximum projections in *xy*-plane of the raw data, otherwise known as reflectance images for the original image (a) and deconvolved image (b). The maximum projection in the *xz*-plane: original image (c) and deconvolved image (d) show axial smearing associated with confocal scanning laser tomography in the original image. There is better discrimination between slices in the deconvolved image. Slice number 15 in the original (e) and deconvolved (f) shows a reduction in high frequency noise. Two z-profiles, pre (g) and post (h) deconvolution, are shown at a position in the rim area (marked by the arrow in (a))

The effect deconvolution had on intra-scan repeatability is summarized in Figure 32. The figure shows the average MPHSD plotted against the difference between MPHSD before and after deconvolution. An improvement in repeatability after deconvolution results in a point being above the zero line parallel to the x-axis. An improvement in MPHSD occurred in 38 of the 40 images. The median improvement of 2.5 microns (inter-quartile range 2.19) was statistically significant ($P<0.001$). The figure also demonstrates that the improvement is greater for subjects with higher MPHSD. There is a statistically significant association between average intra-scan MPHSD and the improvement in repeatability Spearman's $r = 0.45$ ($P=0.004$). The association between LOCS III scores and improvement in repeatability was NO ($r = 0.37$, $P = 0.019$); NC ($r = 0.33$, $P = 0.040$); C ($r = 0.22$, $P = 0.162$) and PS ($r = 0.10$, $P = 0.526$).

In Figure 33 the same figure format shows the effect deconvolution had on inter-scan repeatability. An improvement in inter-scan repeatability of topographic height measurements occurred in 33 of the 40 images, with a mean improvement 1.80 microns ($P<0.001$). There was an association between average inter-scan MPHSD and the difference in MPHSD before and after deconvolution Spearman's $r = 0.49$ ($P=0.002$). The association between LOCS III scores and improvement in repeatability was NO ($r = 0.25$, $P = 0.125$); NC ($r = 0.18$, $P = 0.254$); C ($r = 0.15$, $P = 0.358$) and PS ($r = 0.12$, 0.459).

It takes approximately 3 minutes of computer processing to deconvolve a single HRT image ($256 \times 256 \times 32$ pixels) using a standard desktop computer (Pentium IV 1.6 GHz).

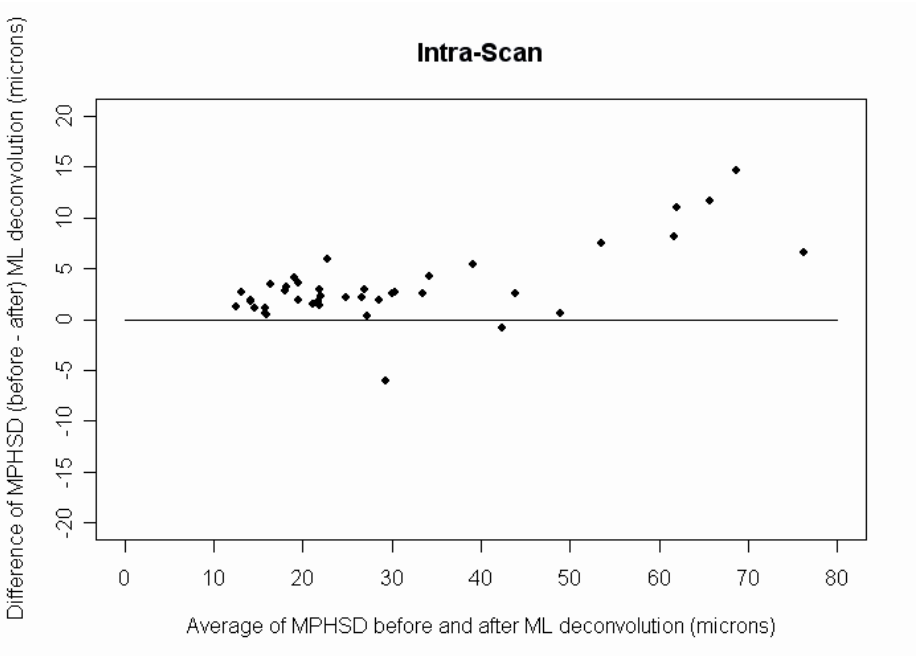


Figure 32 Effect of deconvolution on intra-scan repeatability of topographic height measures. The plot shows the difference in average MPHSD against the difference in MPHSD before and after deconvolution. An improvement in repeatability is represented by a point above the 'zero line'. An improvement in repeatability occurred in 38 of the 40 images (P<0.001)

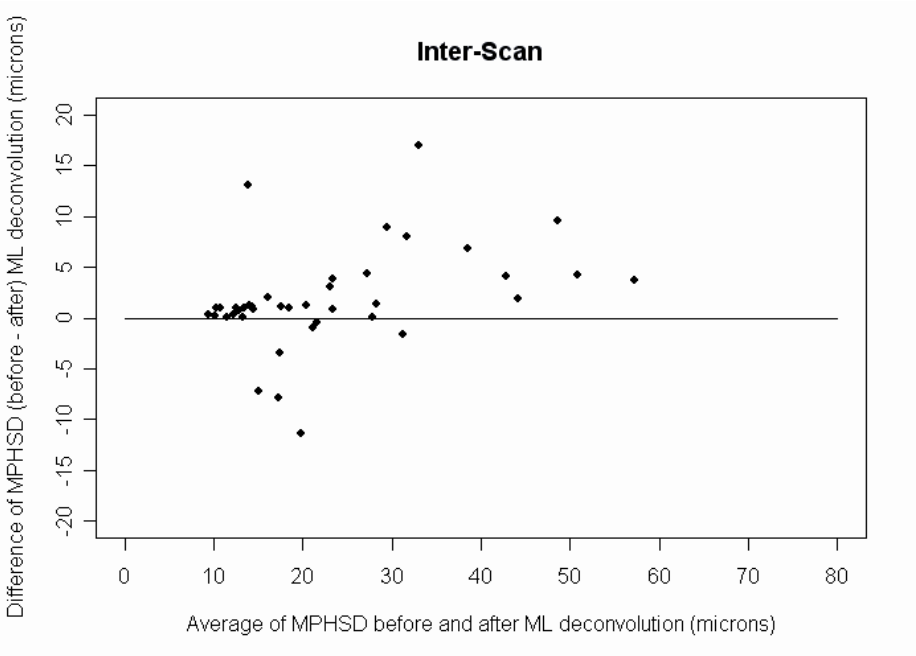


Figure 33 Effect of deconvolution on the inter-scan repeatability of topographic height measures. An improvement in repeatability occurred in 33 of the 40 images (P<0.001)

6.3 Discussion

Previously ML deconvolution has been used in confocal microscopy, wide-field epiflourescence microscopy and transmitted light brightfield microscopy (Holmes, Bhattacharyya et al, 1995). Nourrit and colleagues (Nourrit, Vohnsen et al, 2005) recently reported an improvement in high resolution images obtained using 2D ML deconvolution; this study applied deconvolution to raw data obtained from confocal scanning laser ophthalmoscopy and compared the improvement to that obtained with adaptive optics. Other examples of ML deconvolution in astronomy are described by (Hanisch, White et al, 1997). ML deconvolution has also solved one-dimensional problems in digital signal processing. In one application it was used to reconstruct a speech source recorded in a noisy, reverberant environment (Attias, Platt et al, 2001). To our knowledge, this study is the first to demonstrate the utility of image restoration algorithms in scanning laser tomography ONH images. We found ML deconvolution improved the resolution and the repeatability of HRT images.

In this study image restoration algorithms were applied to 3D confocal images before estimating the position of the peak reflectance to construct a topography image. A recent application of the HRT has been in imaging macular edema in diabetic retinopathy; here a more complex model for the z-profile, the 'edema index', is calculated by Gaussian curve fitting (Guan, Hudson et al, 2004). The improvement obtained by Deconvolution may have more affect in this area where the model fitted is more involved.

Previously, Burk and Rendon (Burk and Rendon, 2001) evaluated the impact of a new subpixel image alignment algorithm on series of HRT images and reported an improvement in intra-scan MPHSD. An average improvement from 22.9 ±8.2(SD) microns with the standard alignment procedure to 15.46 ±6.8(SD) microns with the new algorithm was reported on 132 eyes. The study did not report on the improvement in inter-scan repeatability. Note that although the average gain was

smaller by comparison, it is an increment improvement to that already shown by improved image registration algorithms.

All forms of cataract result in light scattering, reducing the repeatability of measurements made from scanning laser tomography images. Both cataract and glaucoma occurs within the same age range and the two conditions frequently coexist. It is recognized that glaucoma surgery is often followed by a speeding up of a pre-existing cataract (Hitchings, 2000). The image restoration techniques presented in this study may be particularly applicable to glaucoma patients and especially in series of data where cataract might develop during follow-up. Strouthidis and colleagues (Strouthidis, White et al, 2005) previously reported an association between lens opacity and MPHSD using NO and NC. Siik and colleagues (Siik, Chylack et al, 1999) reported a statistically significant association between NO, NC and white light scatter. We report a low association between grading of lens opacity and improvement in MPHSD which, however, was stronger for the intra-scan, than the inter-scan, improvement. This is perhaps due to the improvement afforded by deconvolution getting 'swamped' by noise induced by misalignment present after applying HRT image registration algorithms. Z-profiles at the rim area are shown after data exploration of the image in Figure 31. This important area of the image for monitoring glaucoma and detecting damage during follow-up has low light reflectance properties. In cases with high noise, there will be inherit difficulties in accurately determining the position of maximum reflectance, and hence calculating the topographic height (see Figure 31g). However, the deconvolved z-profile (Figure 31h) clearly shows the position of maximum reflection.

The problems of screening a normal elderly population (aged 65 to 89 years) with scanning laser tomography has been highlighted in a recent study which reported that 10% had a MPHSD above 68 microns, and a mean MPHSD of 26.8 ±13.3(SD) microns with the worst 10% removed (Vernon, Hawker et al, 2005). Strouthidis and colleagues (Strouthidis, White et al, 2005) reported in an OHT and POAG population which was deliberately enriched with cataract cases by scanning the eye

with the highest LOCS score; over a mean of 5 visits, MPHSD had mean 33.5 ±23.6(SD) microns, range [12 to 130 microns]. Table 2 summarises the instrument guidelines on how to interpret MPHSD.

MPHSD	
<10	Excellent
10-20	Very good
20-30	Good
30-40	Acceptable
40-50	Look for way to improve
>50	Low quality (do not use as baseline image)

Table 2 Instrument guidelines categorizing MPHSD (courtesy of Heidelberg Engineering, Heidelberg, Germany)

The results highlighted in Figure 32 and Figure 33 suggest that the improvement obtained by deconvolution was greater in the least repeatable images. For example the mean improvement in the five patients with an original MPHSD greater than 60 microns was 10.4 microns. Therefore, these techniques might have a real clinical impact in the use of images that are of a poor quality.

7. Conclusions and future work

Chapter 2 describes a computer simulation which mimics series of patient images. The simulation attempts to replicate the repeatability of image series observed in real clinical data. It is also designed to replicate the changes that occur in unstable patient series. This is used in later chapters to assess quantitative techniques for detecting structural changes in longitudinal series of ONH images. These simulations of longitudinal HRT data are the first of their kind.

Chapter 3 describes a statistical technique (SIM) which has previously been used by the neuro-imaging community for detecting localised areas of activity in images of the brain. This chapter describes how the technique has been adapted and applied to ONH images. The method is compared with the TCA technique using simulation and clinical data. The computer experiment suggests the TCA false-positive rate is high. The experiments also suggest the sensitivity to detect progression is better with SIM as compared to TCA as the series of data gets longer. The results from simulation, clinical data and individual case studies suggest SIM has a better specificity and sensitivity than the TCA. However, the 'truth' can only be established on unselected longitudinal data.

Chapter 4 describes how the global result defining progression from the SIM analysis can be adjusted with a combining function to allow detection of morphological change with different characteristics. Computer experiments are designed which simulate change of high intensity with small spatial extent and low intensity with large spatial extent. The experiments suggest that the use of combining functions in SIM can help detect different types of change, while still flagging the same number of stable patients.

Chapter 5 tests SIM against a statistical technique available on the current HRT software which uses 'stereometric' parameters of the ONH to detect change. The techniques are tested using real clinical data. The false-positive rate is recorded by

applying the techniques to a test-retest dataset and the true-positive rate is evaluated using a survival analysis on a subset of longitudinal OHT patients which developed visual field defects during follow-up. SIM is significantly more sensitive at detecting change, and the combining function improved the performance of the technique. The chapter also shows that the HRT guidelines for the 'stereometric' parameters analyses result in high false-positive rates.

Chapter 6 applies an image restoration algorithm (ML deconvolution) to test-retest series of HRT images and examines if this technique improves the repeatability of ONH images. One motivation for this work is that computer experiments suggest SIM and TCA will detect change in image series with low MPHSD (higher repeatability) with more success. ML deconvolution significantly improved both intra-scan and inter-scan repeatability and the improvement was greater in the least repeatable images. It is hoped that this technique may be able to use images which would otherwise be unused. These images typically occur in patients with cataracts, a prevalent disease that often coexists with glaucoma.

The key contributions to work in this field are:

- A novel simulation of realistic longitudinal HRT data was developed mimicking series of stable and unstable patients
- Quantitative techniques from neuro-imaging (SIM) were adapted and applied to series of ONH images. Results from simulation and real data suggest these techniques have better specificity and sensitivity than those currently available.
- The new quantitative techniques which have been developed have been incorporated into a windows based program (see Appendix A and B)
- ML deconvolution, an image restoration algorithm, has been shown to improve the repeatability of ONH images.

Potential future work includes:

- Apply SIM to other longitudinal datasets. Establish if age-related change in normal controls can be detected and quantified. Investigate the position of the changes by region: inferior, temporal, superior and nasal to investigate if age related change follows similar patterns to glaucomatous damage. To achieve separation between normal controls and damage resulting from glaucomatous progression; it may be necessary to examine if the results obtained are larger than age-related changes.
- Apply SIM to data obtained from other retinal imaging modalities such as optical coherence tomography and scanning laser polarimetry
- Assess if ML deconvolution can improve the HRT algorithms used for diagnosing glaucoma and detecting change

Appendix A - SIM software tutorial

This following appendix is a self-contained tutorial to enable a user to install SIM, and describes the steps necessary to process a patient. This section assumes the reader is familiar with the Heidelberg Eye Explorer software version 1.4.1 or later:

Installing SIM software

Double click on *setup.exe*. The installation program automatically copies the required files to the hard-drive (Figure 34). The installation by default loads three examples patient series with the SIM analysis already performed.

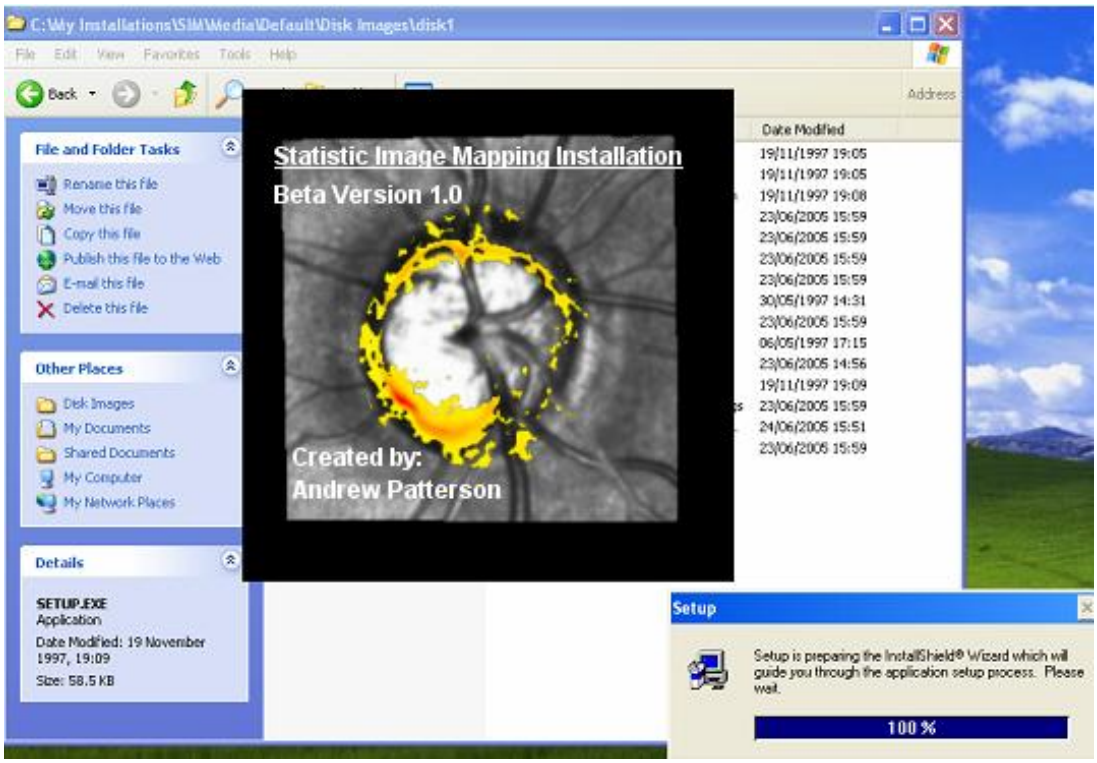


Figure 34 The setup installing SIM software

At the time of writing SIM uses the Matrox Imaging Library (MIL). To run SIM a copy of MIL version 7.1 or later will need to be installed. MIL requires a user license to enable its functionality.

Exporting images from Heidelberg Eye Explorer software

This section assumes a copy of Heidelberg Eye Explorer software version 1.4.1 or later is installed. Exporting images using the function described below aligns the patient series to compensate for differences in scan positions which occur both within visit and between visits using a proprietary HRT software alignment algorithm.

To enable exporting of HRT 1 images add the following code anywhere in *c:\heyex\plugins\HRTS.ini*. To enable exporting HRT 2 images add code 1 to file *c:\heyex\plugins\HRT.ini*.

```
[progression]
ppath=c:\pmaps
tpath=c:\topos
```

Then create a directory *c:\topos.* To export a patient series drag the required mean topography images into the light box. When all the required images are in the lightbox, highlight the images, then right click and select 'include into progression'. Then highlight and right-click the progression tab which was created and select 'export raw data'. The images will be exported to directory *c:\topos*

Creating a patient record

Double click on file *C:\SIM\sim.exe,* or the SIM icon that appears in the start menu and desktop. The SIM program will now launch (see Figure 35).

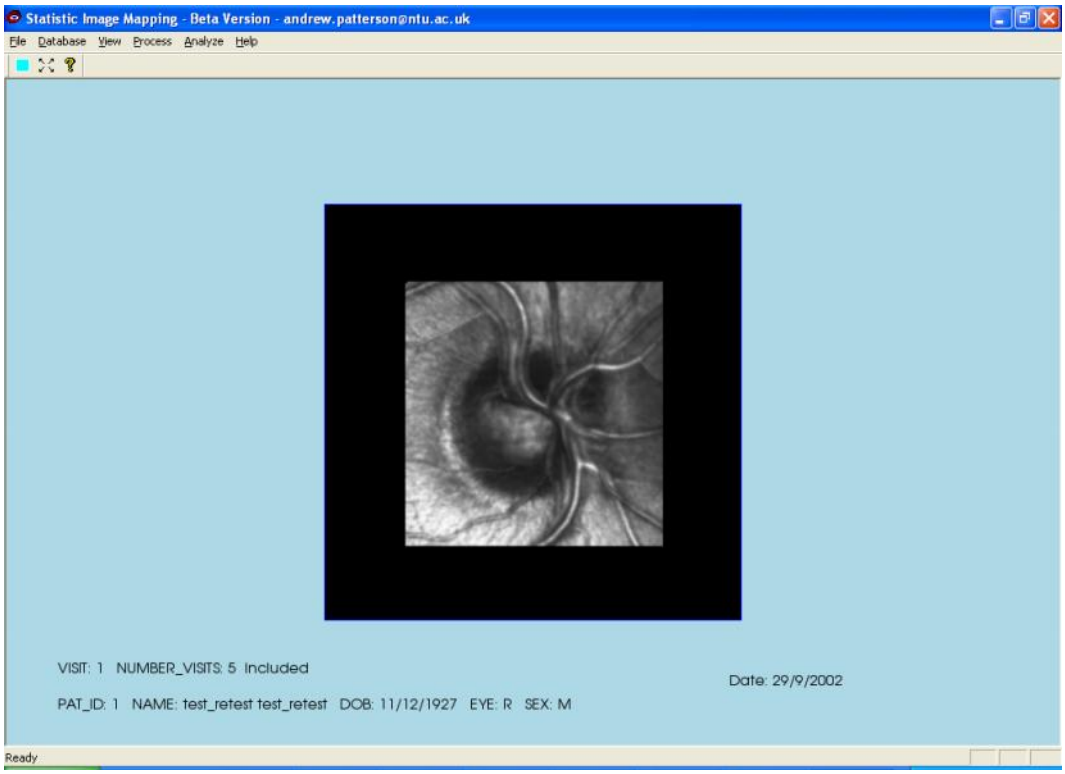


Figure 35 The SIM software user interface rendering a mean reflectance image

Select *Database* and *'Add Patient'* and fill out the required fields in the dialog (Figure 36). The date of the baseline examination must be entered at this point as this information is not exported from the Heidelberg Eye Explorer software.

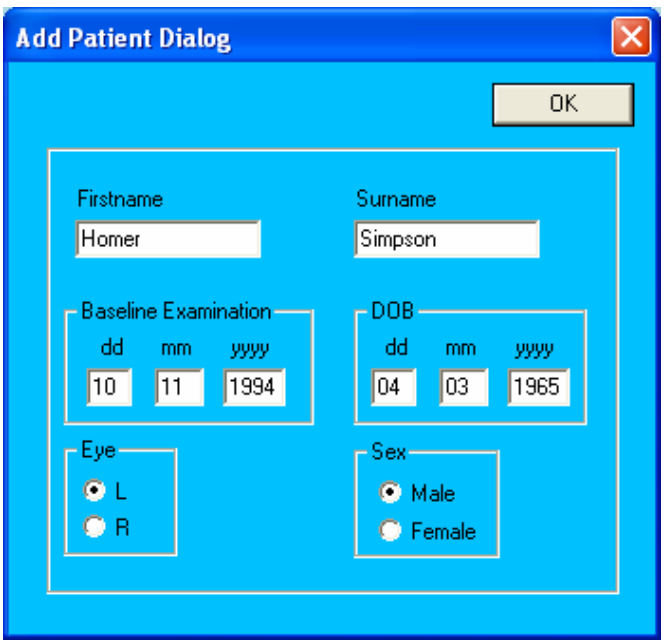


Figure 36 The 'Add Patient' dialog box

Import images into SIM

Select *Database* and *'Import Image Series'* (or press ctrl+I). The dialog box shown in Figure 37 appears. Select the patient, HRT image format and the corresponding image series in the correct sequential order. The filenames output by the HRT software contains the following information:

sssfeyyyymmdd-i.raw

where:

sss = First three letters from the patients surname

f = First letter from the patient first-name

e = eye (L or R)

yyyy = year

mm = month

dd = day

i = increments if a duplicate image was taken on a single day.

The filename therefore enables the user to select the correct topography files which correspond to the patient.

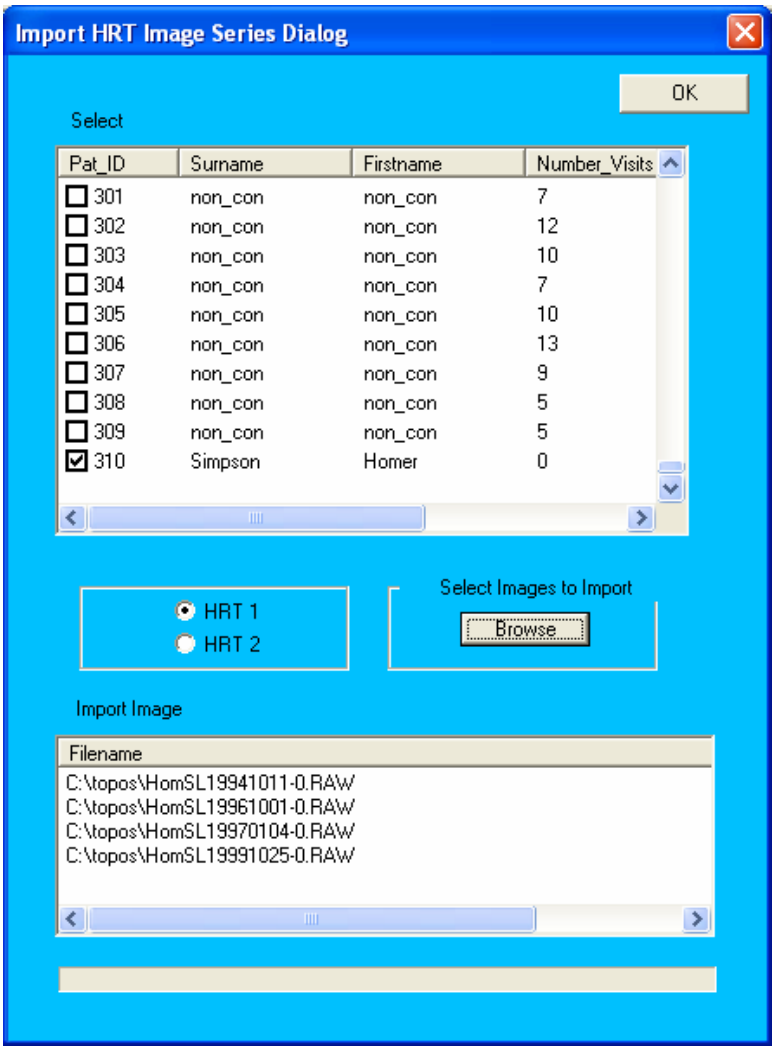


Figure 37 The 'Import HRT Image Series' dialog box. This dialog allows the patient, HRT image format (HRT 1 or HRT 2) and the topography image series to be selected

Select *OK* and the images are imported into the SIM database.

Viewing patient series

Select *View* and *'Load Patient'*. The load patient dialog appears, select the record which corresponds to the patient you wish to load. By default the mean reflectance image of the baseline examination is rendered on the screen (Figure 35).

To view the next or previous image in the series select *View* and *'Next/Previous Image'* (or alternate press 'O' or 'P'). To view a single reflectance image select *View* and *'Single Reflectance Image'*. The topography images can be viewed in 3D

to the correct scale by selecting *View* and *'Single Topography Image'* or *'Mean Topography Image'*. Press the left mouse button and position the mouse to rotate the 3D view. Press the right button and position the mouse to control zooming in and out. This visualization functions produced the images shown in Figure 38.

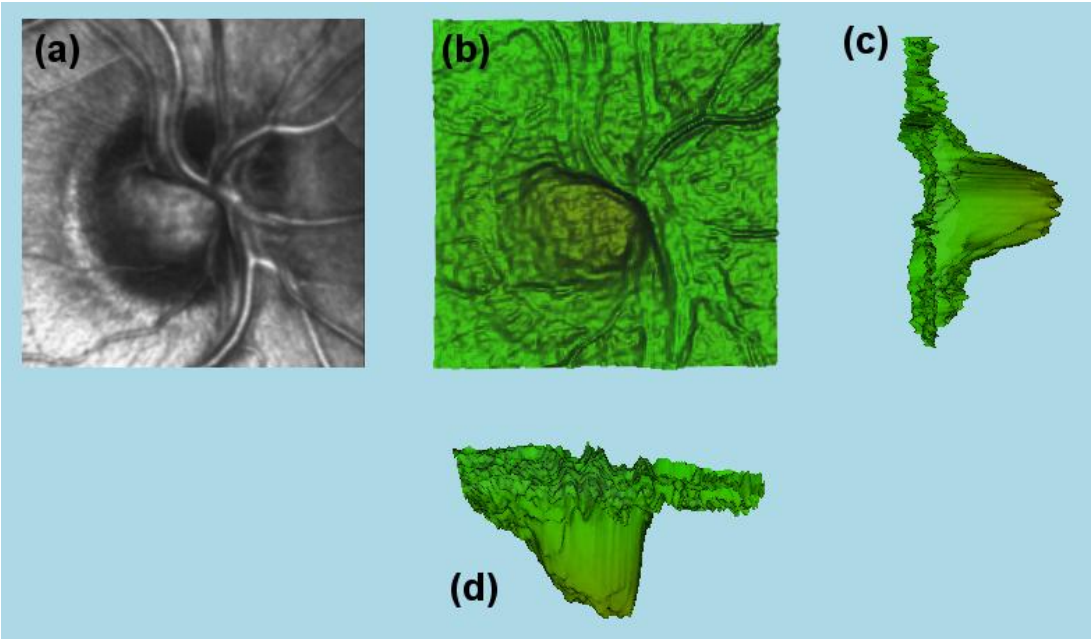


Figure 38 Visualisation of reflectance and topography images with the SIM software. (a) Mean reflectance image, (b) top elevation of the topography image, (c) side elevation and (d) front elevation

Check images for alignment and magnification error

The function 'check-board' was written to enable the user to subjectively test if images in the patient series are correctly aligned. The check-board function simultaneously renders the baseline and a follow-up image from the series. To do this select *View* and '*Check-Board*'. Then select *View* and *'Next/Previous image'* to scroll through the follow-up images in the series. Select *Process* and *'Exclude/Include examination'* to state which images in the series to process.

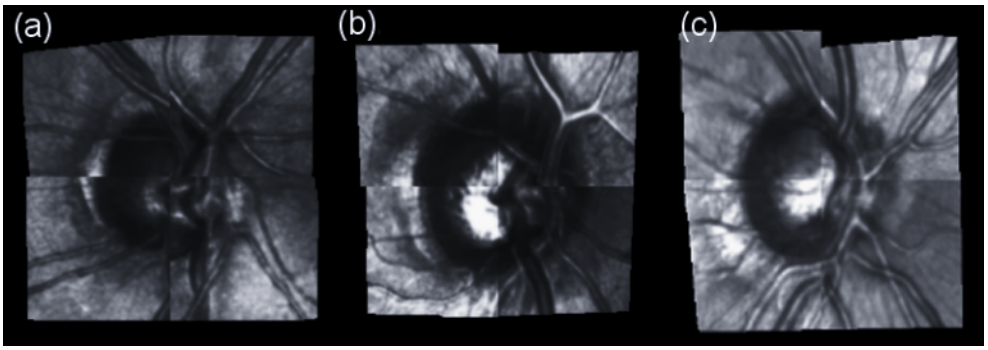


Figure 39 Image series alignment: The images contain two quadrants from the baseline image shown in the top-right and bottom-left quadrants; and two quadrants from the follow-up images shown in the top-left and bottom-right quadrants. (a) A follow-up image which has translation and rotation misalignment between the follow-up image and the baseline image, (b) a follow-up image which has magnification error, and (c) a follow-up image which is well aligned

Selecting an area-of-interest

An area-of-interest needs to be defined for each patient. To do this select *Process* and *'Draw Contour Line'*, then move the five 'handles' to position the contour line (see Figure 40). Then select *Process* and *'Accept Contour Line'* to store the position of the contour line.

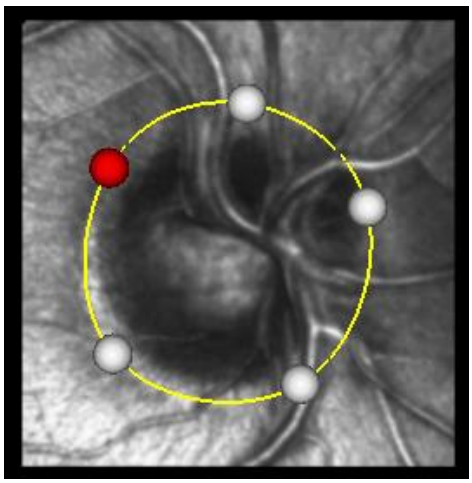


Figure 40 The position of the contour line control is determined using five 'handles'. A handle on the contour line becomes red when it has been selected or moved. The position of this contour line is used for follow-up images in the patient series. Only pixels bound within this contour line are process by the SIM paradigm (see sections 3.2 and 4.2)

Creating a batch file to process patient series

SIM software uses batch files to process multiple patient series. Select *Process* and *Batch* and the dialog shown in Figure 41 appears. Select the patients to be processed and use the options shown to control processing patient series. After 'Create Dialog' is selected, the batch file 'c:\SIM\SIM_batch.bat' is created or replaced. This batch file can be viewed in a text editor program or executed at a later date by double clicking on the batch file.

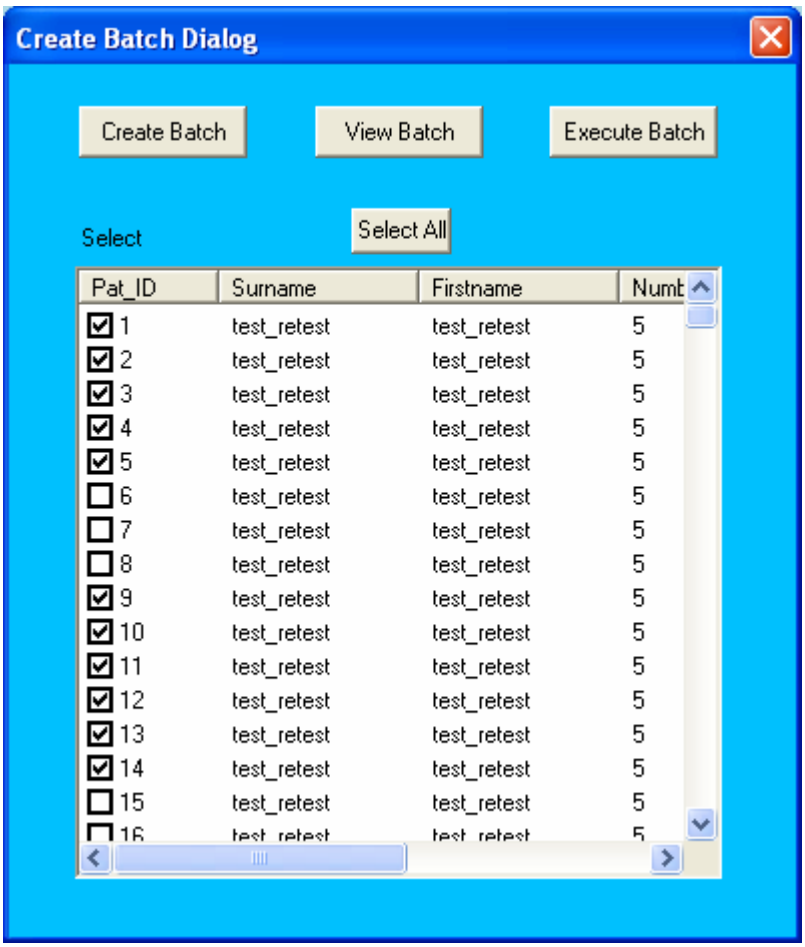


Figure 41 Creating, viewing and executing batch files is controlled by selecting (multiple) patient series using the 'Create Batch' dialog box

Viewing 'change maps'

When patient image series are processed 'change maps' are generated showing the intensity and spatial extent of change which has occurred during follow-up. The

first change map is produced at visit 4 (i.e. a minimum of 4 sets of visits are required) and for every subsequent visit after this. To view a change map select *View* and '*Change Map*'. Use *'Next/Previous image'* to scroll through a patient series.

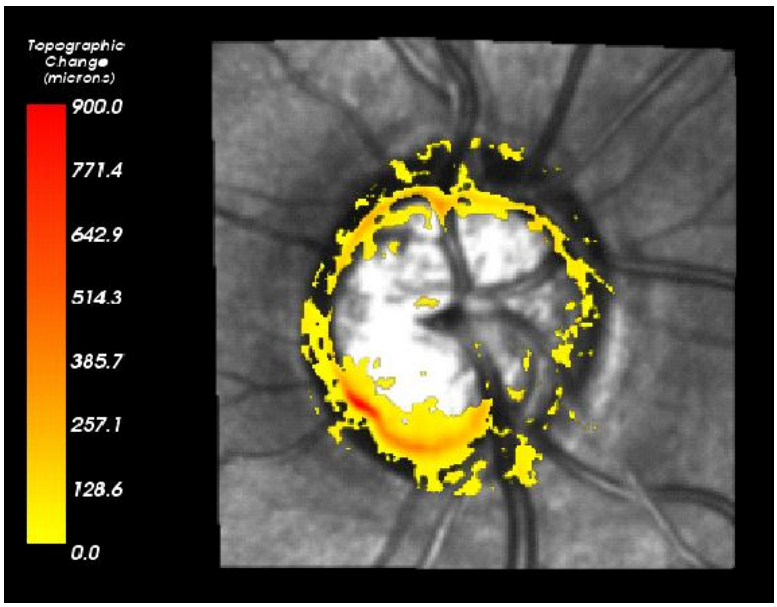


Figure 42 'Change map' showing the intensity and spatial extent of depressed morphological change which has occurred during a patients follow-up

Exporting SIM parameters

To obtain the patient details, examination dates and the corresponding parameters such as size and significance of cluster size statistic, T-max statistic or combining functions for both significant depressed and change select *Analyze* and *Filter*. Then select the patients and parameters you wish to output (Figure 43). The results are output as a tab delimited file located at *c:\SIM\results.txt.*

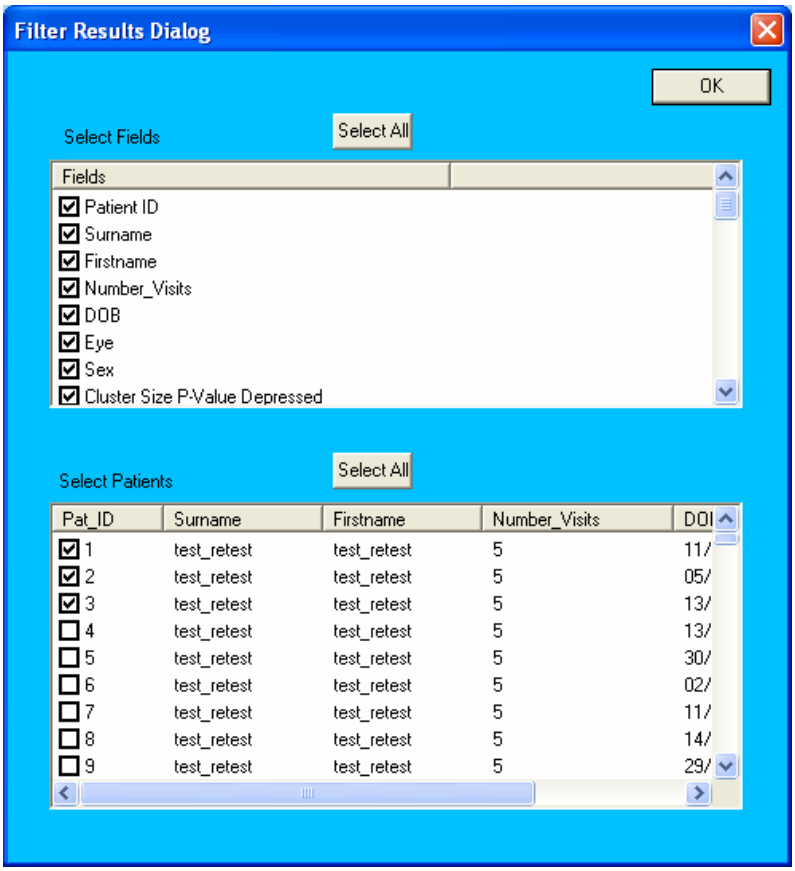


Figure 43 The 'Filter Results' dialog box outputs patient details and SIM parameters

Appendix B - SIM software design and development issues

The objective of this appendix is to provide sufficient detail to allow a research to append or modify the SIM source code. The researcher reading this appendix and intending to modify SIM would benefit from knowledge of the C++ programming language and the principals behind object orientated programming. The appendix also includes terminology associated with graphics software programming, numerical methods and GUIs. There is an abundance of literature on these topics. Woo and colleagues (Woo, Neider et al, 1997) explain the concepts behind graphics programming. Avila and colleagues (Avila, Barre et al, 2003) give examples and source code of graphics programming relevant to 2D and 3D medical images. Vetterling and Press (Press, 2002; Vetterling and Press, 2002) provide source code and examples of numerical methods. Bates and Tompkins (Bates and Tompkins, 1999) provides tutorials and examples to develop GUIs. Jones (Jones, 1999) explains GUI programming from an object orientated prospective.

The SIM software contains two executables called SIM_DOS and SIM_GUI. SIM_DOS runs in a dos shell (windows dos prompt) and implements the computational paradigms explained in section 3.2 and 4.2. SIM_GUI contains a user interface and provides the functionality documented in Appendix A.

SIM_DOS inputs a set of command line arguments. At run-time SIM_DOS inputs topography images, patient information, the time elapsed between patient visits, a Boolean vector to indicate whether to include or exclude visits from the analysis and a binary image of the area of interest to process (the area bound within the contour line). The executable checks if the files are located at the pathname specified by the command line arguments, if any files are missing the program exits and prints an error message. SIM_DOS has no GUI as the program was designed to be 'light-weight' to increase computation time.

SIM_DOS uses two C++ libraries, the numerical recipes C++ library (www.nr.com) and the matrox imaging library (www.matrox.com). SIM_DOS outputs binary files of the 'change map' (see Figure 42) and the cluster size, T-max and the combining function Tippet statistics. Table 3 contains a list and description of the files needed to compile and link SIM_DOS.

Name	**Description**
Vector.h/cpp	Template container class for one-dimensional array
Matrix.h	Template container class for two-dimensional array
Matrix3D.h	Template container class for three-dimensional array
Ptr.h	Utility class used by Vector, Matrix and Matrix3D
Binary.h	Class contains functions to input 1-byte and 2-byte binary images
MathFunctions.h	Static class of maths and statistics functions
Input.h	Inputs one/two/three-dimensional arrays from ASCII files
Output.h	Outputs one/two/three-dimensional arrays to ASCII files
nr.h, nrutil.h, nrutil_nr.h, nrtypes.h, nrtypes_nr.h	Numerical recipes header files
MIL.h	Matrox imaging library header file
Progression.h	The progression class containing the SIM paradigm (see sections 3.2 and 4.2)
DIP.h	Perform two dimensional spatial convolution and generates Gaussian kernels
SIM_Stat.h	Container class stores SIM parameters
shell.cpp	Numerical recipe which sorts arrays into ascending/descending order
main.cpp	Contains function main written to input command line arguments, main function also inputs image data and calls

progression class to compute SIM paradigm

Table 3 Header and C++ files required to compile and link SIM_DOS

Table 4 contains a list DLLs needed to execute SIM_DOS at run-time; these files need to be placed in the same folder as SIM_DOS at run-time or in a systems directory.

Name
mil.dll
Milblob.dll
milcal.dll
milcode.dll
miledge.dll
milim.dll
milmeas.dll
Milmod.dll
milocr.dll
milpat.dll
milvga.dll
Milvhook.dll

Table 4 DLLs required at run-time to execute SIM_DOS

The files listed in Table 5 are read by SIM_DOS at run-time. SIM_DOS initially checks the state of each file listed in the table and prints an error message to the screen and exits if it cannot locate these files.

Name	**Description**
top_*.raw where * is 1 → 12	A 2 byte binary files (binary format is signed short). Each file represents a single topography, a minimum of 12 single topographies (4 visits × 3 scans per visit) are

	required for the SIM permutation framework.
_window.raw	A 1 byte binary file (binary format is unsigned char). Each pixel in the file is either 0 which represents a pixel outside the contour line or 255 representing pixels inside the contour line.
_x.txt	Text file contains a vector of the time which has elapsed between visits in days
_exam.txt	Text file contains a vector of Boolean values which state whether to include or exclude an examination from the analysis
c_*_3.txt where * is 12,15,18, …,75	Text file contains lookup tables where each row represents a unique permutation.

Table 5 Files required to execute SIM_DOS

Table 6 contains a list of the files created by SIM_DOS during execution. These files are input by SIM_GUI and used to render 'change maps'.

Name	**Description**
_sim_*.txt	Text file contains the cluster size, T_max and combing function Tippet statistics
stat_*.raw	1 byte binary file (format unsigned char) containing the univariate statistic image. Pixel locations without statistically significant change are represented by 0, elevated statistically significant change is represented by a value of 1 and depressed change is represented by a value of 2 and.
tippet_dep_*.raw, tippet_ele_*.raw	2 byte binary file (format signed short) containing a 'change map' (as described in chapter 5). Pixel locations

without statistically significant change are represented by 0 and positive values represent the intensity of change which has occurred between baseline and follow-up calculated by the rate of change times the time which has elapsed.

Table 6 Files created by SIM_DOS during execution (* represents the visit number)

SIM_GUI is the second executable which runs in a windows environment (tested on WindowXP and Windows 2000). The executable is a GUI which provides functions to allow a user to input data from Heidelberg Eye explorer software (tested on v1.4.1 and later versions), check image alignment quality, determine an area of interest (by defining a contour line), and generate batch files which the SIM_DOS executable processes. The SIM_GUI can visualise imaging data in a 2D and 3D mode, view 'change maps' and output global parameters. The functions are shown in a step by step manner in the tutorial in Appendix A.

The SIM_GUI is a merger of two programming technologies: the visualisation toolkit (VTK, (www.vtk.org)) and microsoft foundation classes (MFC, (www.msdn.com)). The VTK contains high level functions to rendering points, lines, scale-bars and has specific functions suitable for visualising 2D and 3D medical images. The VTK default behaviour gives an application controls for panning, zooming or rotating the scene being rendered. The VTK library contains examples, source code and online documentation. SIM_GUI is based on a VTK example called *VTK_SDI* which is a merger of the VTK and MFC technologies. The MFCs is best described as the 'glue' which allows a user to interact with a program allowing the developer to write code which will be executed after a user presses a specific key or selects a menu item from the menu bar.

Table 7 contains a list of header and C++ files needed to compile and link SIM_GUI.

Name	Description
VTK_Header_Files.h	Includes all require VTK header files
vtkAJPViewSingleReflectance.h/cpp	Reads and renders a single reflectance image
vtkAJPContourLine.h/cpp	Renders a contour line in a default position and allows a user to manipulate the contour line position.
vtkAJPGrid.h/cpp	Renders a grid
vtkAJPHeatMapStatisticImage.h/cpp	Renders a ‘change map’ overlaid on top of a mean reflectance image, this class colour codes intensity of change from yellow to red representing shallow through to deep excavation.
vtkAJPScaleBar.h/cpp	Renders a scale bar used to link the colour in the ‘change map’ with the intensity of change
vtkAJPSetCamera.h/cpp	Sets the camera position to a default two dimensional view (as seen at start-up)
vtkAJPUnivariateStatisiticImage.h/cpp	Renders a statistic image overlaid on top of a mean reflectance image, this class colour codes depressed change as a monotone red colour and elevated change as a monotone green colour.
vtkAJPViewCheckBoard.h/cpp	Renders a baseline and a follow-up image simultaneously, the baseline image is rendered in the bottom left and top right quadrants with the follow-up image rendered in the other two quadrants.
vtkAJPViewMeanReflectance.h/cpp	Reads and renders a mean reflectance image
vtkAJPViewMeanTopography.h/cpp	Reads and renders a mean topography

	image, the image is rendered as a surface in three dimensions.
vtkAJPViewPatientInfo.h/cpp	Displays patient identification, name, date of birth, sex, and number of visits.
vtkAJPViewSIMInfo.h/cpp	Displays p-values for cluster size, T-max and the combining function Tippet
vtkAJPViewSingleTopography.h/cpp	Reads and views a single topography image, the image is rendered as a surface in three dimensions.
Prelude.h, sfl.h, version.h	Header class used by SFL (standard function library, (www.imatix.com)
sfldate.h/c	Provides functions to perform numerical operations to dates
Sflprint.h/c, sflstr.h/c, sflsymb.h/c, sflenv.h/c, sflfind.h/c, sflmem.h/c	Required by SFL
DataExtraction.h/cpp	Inputs HRT files and saves the image information in a directory structure
geometry.h/cpp	Contains functions to determine if a point is inside a closed polygon (require to calculate area-of-interest)
Utils.h/cpp	Required by geometry.h/cpp
Patient.h/cpp	Container class for patient information
SIM_Result.h/cpp	Utility class for reading SIM results from text files
StereometricParameters.h/cpp	Container class for SIM results
Vector.h/cpp	Template container class for one-dimensional array
Ptr.h	Utility class used by Vector.h/cpp
DialogAddPatient.h/cpp, DialogCreateBatch.h/cpp, DialogDeletePatient.h/cpp,	Control dialog boxes and data validation, each class inherits the MFC CDialog class

DialogFilterSIMResults.h/cpp, DialogImportImageSeries.h/cpp, DialogImportParameters.h/cpp, DialogLoadPatient.h/cpp, DialogSIMSettings.h/cpp, DialogStatisticImageSettings.h/cpp	
MainFrm.cpp	Controls the rendering and state of the *frame*, a windows frame typically has functions to minimize, maximize and restore the application window size as well as displaying the applications name in the titlebar.
StdAfx.h/cpp, vtkSDI.h/cpp	Utility files generated by MFC
vtkSDI.rc	Includes information on the appearance of the GUI, such as the size of the dialog boxes and position of edit boxes within each dialog box.
vtkSDIDoc.cpp	The file contains the document class (inherits *CDocument)*
vtkSDIView.cpp	This file contains the view class which inherits *CView*. This class contains a function called *pipeline()* which is called every time a user forces the GUI to update.

Table 7 Header and C++ files required to compile and link SIM_GUI

The following section provides two examples of code taken from the SIM_GUI source code. The objective is to illustrate how MFCs has been used to control the execution of compiled binary code at run-time. It is hoped that this will allow other researchers to append functionality to the SIM software in the same style.

The view class (class *vtkSDIView*) controls the behaviour of SIM_GUI. Each menu function is typically linked to a Boolean variable which is initialised to a default state at start-up. The pipeline function contains a series of *if()* statements. When a user selects for example “View Single Topography” from the menu bar, the Boolean member variable *b_single_topography* is changed state from *false* to *true*. This causes the following code within the pipeline function to render a single topography, for the patient and visit currently selected.

```
Pipeline()
{
…
      if(b_single_topography == true)
      {

      vtkAJPViewSingleTopography::ViewSingleTopography(patient,
      visit);
      //comment: call view single topography function
      //from class vtkAJPViewSingleTopography
      }
}
```

A user can also select an item from the menu bar such as “Add Patient” which prompts a dialog box to be rendered to the screen (see Figure 36). The default behaviour of a dialog box is to appear in front of the application in focus (which makes the rest of the application inactive until the dialog box is closed). In this application the dialog boxes are called from within the view class (class *vtkSDIView*) by first initializing a dialog object and then calling the *domodal()* member function of the dialog object. The dialog box prompts a user to input the required data. One advantage of the MFCs is that it incorporates data validation. For example, the date is rejected if it is outside defined limits, in this case the month in the date of birth field must be between 1 and 12 inclusive. If the user inputs the required data which is valid and then presses the ‘OK’ button the domodal() function returns a *true* state.

The data is then typically processed within the application as shown in the example below.

```
OnMenuAddPatient()
{
      //Initial an AddPatient dialog object
      CDialogAddPatient AddPatient;

      if(AddPatient.domodal()==false)
      {
            //comment: if the user closes the dialog or
            //enters invalid data the return statement
            //forces the function to close
            return;
      }
      else
      {
            //comment: otherwise the data is added to the
            //database
            …
      }
}
```

The files listed in Table 8 are the DLL which need to be located in the same directory as the SIM_GUI executable or in a systems directory at runtime.

Name

vtkCommon.dll

vtkexpat.dll

vtkFiltering.dll

vtkfreetype.dll

vtkftgl.dll

vtkGraphics.dll
vtkHybrid.dll
vtkImaging.dll
vtkIO.dll
vtkjpeg.dll
vtkParallel.dll
vtkPatented.dll
vtkpng.dll
vtkRendering.dll
vtktiff.dll
vtkzlib.dll

Table 8 DLLs required at run-time to execute SIM_GUI

SIM_GUI is linked to a database and directory structure which is documented in Table 10. The directory structure enables SIM_GUI to locate and render the correct patient series and image at run-time.

Name	**Description**
database.txt †	SIM_GUI needs a database file to link a patient name with the directory number. By default the file *c:/sim/data/database.txt* is the database file. SIM_GUI also accepts a command line argument, for example you can create a desktop shortcut and set the shortcut to reference the SIM_GUI executable, then enter -c:/sim/data/moorfields_normals.txt as a command line argument. This allows the SIM software can have multiple databases.
c:/sim/patients//*	SIM automatically creates a new directory each time a patient record us successfully entered using the 'add patient' dialog box. The software creates new directories in an ascending

numerical order. When HRT data is imported, the image data is placed in the directory associated with the patient record.

Table 9 Files and directory structure required to execute SIM_GUI

Table 10 contains the list of files which can be generated by SIM_GUI.

Name	**Description**
_x.txt †	Text file containing vector of the time in days between each follow-up visit
_exam.txt	Text file containing vector of visits to include or exclude from the analysis (0 = exclude) and (1 = include)
m_ref_*.raw †, * is the visit number	Mean reflectance image stored as 1 byte binary file (unsigned char) containing 384 × 384 pixels
m_top_*.raw †, * is the visit number	Mean topography image stored as 2 byte binary file (signed short) containing 384 × 384 pixels
ref_*.raw †, * is the sequence in which the images were obtained	Single reflectance image, stored as 1 byte binary file (unsigned char) containing 384 × 384 pixels
top_*.raw †, * is the sequence in which the images were obtained	Single topography image stored as 2 byte binary file (signed short) containing 384 × 384 pixels
c:/SIM/results.txt	Text file of results is created by the 'Filter Results' function (see Figure 43)
C:/SIM/sim_batch.bat	Batch file is created by the 'Batch' function (see Figure 41)

Table 10 List of files created by SIM_GUI († files created when HRT data is imported, see Figure 37)

Limitations and suggestions for future improvements

The SIM_DOS executable currently uses image processing functions from the Matrox Imaging Library. During run-time SIM_DOS initialises functions from the Matrox Imaging Library which requires a commercial license (in the form of a USB or serial dongle). To make SIM freely available SIM_DOS needs to be developed using a freely available image processing library.

References

Adler, W., Hothorn, T. and Lausen, B. (2004). "Simulation based analysis of automated, classification of medical images." Methods Inf Med 43(2): 150-5.

AGIS (1994). "Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability." Ophthalmology 101(8): 1445-55.

AGIS (2000). "The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration.The AGIS Investigators." Am J Ophthalmol 130(4): 429-40.

Airaksinen, P. J. and Drance, S. M. (1985). "Neuroretinal rim area and retinal nerve fiber layer in glaucoma." Arch Ophthalmol 103(2): 203-4.

Andrews, H. and Hunt, B. (1977). Digital image restoration. Englewood Cliffs, Prentice-Hall.

Arndt, S., Cizadlo, T., Andreasen, N. C., Heckel, D., Gold, S. and O'Leary, D. S. (1996). "Tests for comparing images based on randomization and permutation methods." J Cereb Blood Flow Metab 16(6): 1271-9.

Artal, P., Guirao, A., Berrio, E. and Williams, D. (2001). "Compensation of corneal aberrations by the internal optics in the human eye." Journal of Vision 1: 1-8.

Artes, P. and Chauhan, B. (2005). "Longitudinal changes in the visual field and optic disc in glaucoma." Progress in Retinal and Eye Research 24: 333-354.

Asawaphureekorn, S., Zangwill, L. and Weinreb, R. (1996). "Ranked-segment distribution curve for interpretation of optic nerve topography." J Glaucoma 5: 79-90.

Attias, H., Platt, J., Acero, A. and Deng, L. (2001). "Speech denoising and dereverberation using probabilistic models." Advances in Neural Information Processing Systems 13: 758-764.

Avila, L., Barre, S., Geveci, B., Henderson, A., Hoffman, W., King, B., Law, C., Martin, K. and WJ, S. (2003). The VTK User's Guide (VTK 4.2). USA, Kitware Inc.

Balazsi, A. G., Drance, S. M., Schulzer, M. and Douglas, G. R. (1984). "Neuroretinal rim area in suspected glaucoma and early chronic open-angle glaucoma. Correlation with parameters of visual function." Arch Ophthalmol 102(7): 1011-4.

Bates, J. and Tompkins, T. (1999). Practical Visual C++ 6. USA, QUE.

Bathija, R. (2000). "Optic nerve blood flow in glaucoma." Clin Exp Optom 83(3): 180-184.

Bathija, R., Zangwill, L., Berry, C. C., Sample, P. A. and Weinreb, R. N. (1998). "Detection of early glaucomatous structural damage with confocal scanning laser tomography." J Glaucoma 7(2): 121-7.

Bellezza, A. J., Rintalan, C. J., Thompson, H. W., Downs, J. C., Hart, R. T. and Burgoyne, C. F. (2003). "Deformation of the lamina cribrosa and anterior scleral canal wall in early experimental glaucoma." Invest Ophthalmol Vis Sci 44(2): 623-37.

Bengtsson, B., Olsson, J., Heijl, A. and Rootzen, H. (1997). "A new generation of algorithms for computerized threshold perimetry, SITA." Acta Ophthalmol Scand 75(4): 368-75.

Bennett, A. G. and Rabbetts, R. B. (1998). Bennett and Rabbetts' clinical visual optics. Oxford ; Boston, Butterworth-Heinemann.

Brett, M., Penny, W. and Kiebel, S. (2003). An introduction to random field theory. Human Brain Function II. London, UK, Academic Press.

Brigatti, L., Weitzman, M. and Caprioli, J. (1995). "Regional test-retest variability of confocal scanning laser tomography." Am J Ophthalmol 120(4): 433-40.

Bueno, J. (2004). "The influence of depolarization and corneal birefringence on ocular polarization." J Opt. A: Pure Appl. Opt. 6: 91-99.

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E. and Brammer, M. J. (1999). "Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain." IEEE Trans Med Imaging 18(1): 32-42.

Burk, R., Rohrschneider, K., Völcker, H. and Zinser, G. (1990). Analysis of three-dimensional optic disc topography by laser scanning tomography. Laser

Scanning Ophthalmoscopy and Tomography. B. R. Nasemann JE, München, Quintessenz**:** 161–76.

Burk, R. O. and Rendon, R. (2001). "Clinical detection of optic nerve damage: measuring changes in cup steepness with use of a new image alignment algorithm." Surv Ophthalmol 45 Suppl 3: S297-303; discussion S332-4.

Burk, R. O., Vihanninjoki, K., Bartke, T., Tuulonen, A., Airaksinen, P. J., Volcker, H. E. and Konig, J. M. (2000). "Development of the standard reference plane for the Heidelberg retina tomograph." Graefes Arch Clin Exp Ophthalmol 238(5): 375-84.

Chauhan, B. C. (1996). "Interpreting technology: confocal scanning laser tomography." Can J Ophthalmol 31: 152-156.

Chauhan, B. C., Blanchard, J. W., Hamilton, D. C. and LeBlanc, R. P. (2000). "Technique for detecting serial topographic changes in the optic disc and peripapillary retina using scanning laser tomography." Invest Ophthalmol Vis Sci 41(3): 775-82.

Chauhan, B. C., Drance, S. M. and Douglas, G. R. (1990). "The use of visual field indices in detecting changes in the visual field in glaucoma." Invest Ophthalmol Vis Sci 31(3): 512-20.

Chauhan, B. C., LeBlanc, R. P., McCormick, T. A. and Rogers, J. B. (1994). "Test-retest variability of topographic measurements with confocal scanning laser tomography in patients with glaucoma and control subjects." Am J Ophthalmol 118(1): 9-15.

Chauhan, B. C. and MacDonald, C. A. (1995). "Influence of time separation on variability estimates of topographic measurements with confocal scanning laser tomography." J Glaucoma 4: 189-193.

Chauhan, B. C. and McCormick, T. A. (1995). "Effect of the cardiac cycle on topographic measurements using confocal scanning laser tomography." Graefes Arch Clin Exp Ophthalmol 233(9): 568-72.

Chauhan, B. C., McCormick, T. A., Nicolela, M. T. and LeBlanc, R. P. (2001). "Optic disc and visual field changes in a prospective longitudinal study of patients with glaucoma: comparison of scanning laser tomography with

conventional perimetry and optic disc photography." Arch Ophthalmol 119(10): 1492-9.

Ciulla, T., Regillo, C. and Harris, A. (2003). Retina and Optic Nerve Imaging. Philadelphia, Lippincott Williams and Wilkins.

Coffey, M., Reidy, A., Wormald, R., Xian, W. X., Wright, L. and Courtney, P. (1993). "Prevalence of glaucoma in the west of Ireland." Br J Ophthalmol 77(1): 17-21.

Coleman, A. L., Sommer, A., Enger, C., Knopf, H. L., Stamper, R. L. and Minckler, D. S. (1996). "Interobserver and intraobserver variability in the detection of glaucomatous progression of the optic disc." J Glaucoma 5(6): 384-9.

Colton, T. and Ederer, F. (1980). "The distribution of intraocular pressures in the general population." Surv Ophthalmol 25(3): 123-9.

Crabb, D., Owen, V., Strouthidis, N. and Garway-Heath, D. (2006). Towards estimating rates of structural progression in glaucoma. Image Morphometry and Glaucoma in Europe, Heidelberg, Germany.

Crabb, D. P., Fitzke, F. W. and Hitchings, R. A. (1999). "Detecting gradual and sudden sensitivity loss in series of visual fields." In Permetry Update: 131-138.

Crawford, M. L., Harwerth, R. S., Smith, E. L., 3rd, Shen, F. and Carter-Dawson, L. (2000). "Glaucoma in primates: cytochrome oxidase reactivity in parvo- and magnocellular pathways." Invest Ophthalmol Vis Sci 41(7): 1791-802.

Da Pozzo, S., Iacono, P., Marchesan, R., Vattovani, O. and Ravalico, G. (2005). "Comparing measurements of retinal nerve fiber layer thickness obtained on scanning laser polarimetry with fixed and variable corneal compensator." Eur J Ophthalmol 15(2): 239-45.

Dempster, A., Laird, N. and Ruben, D. (1977). "Maximum Likelihood from Incomplete data via the EM Algorithm." J R Stat Soc B 39: 1-37.

Dielemans, I., Vingerling, J. R., Wolfs, R. C., Hofman, A., Grobbee, D. E. and de Jong, P. T. (1994). "The prevalence of primary open-angle glaucoma in a population-based study in The Netherlands. The Rotterdam Study." Ophthalmology 101(11): 1851-5.

Diggory, P. and Franks, W. A. (1997). "Glaucoma therapy may take your breath away." Age Ageing 26(2): 63-7.

Drance, S. M. (1975). "Optic disc in glaucoma." Trans Ophthalmol Soc N Z 27: 18-9.

Dreher, A., Tso, P. and Weinreb, R. (1991). "Reproducibility of topographic measurements of the normal and glaucomatous optic nerve head with the laser tomographic scanner." Am J Ophthalmol 111: 221-229.

Everitt, B. and Dunn, G. (1998). Statistical analysis of medical data : new developments. London, Arnold.

Fitzke, F. W., Hitchings, R. A., Poinoosawmy, D., McNaught, A. I. and Crabb, D. P. (1996). "Analysis of visual field progression in glaucoma." Br J Ophthalmol 80(1): 40-8.

Ford, B. A., Artes, P. H., McCormick, T. A., Nicolela, M. T., LeBlanc, R. P. and Chauhan, B. C. (2003). "Comparison of data analysis tools for detection of glaucoma with the Heidelberg Retina Tomograph." Ophthalmology 110(6): 1145-50.

Foster, P. and Johnson, G. (2000). Primary angle closure: epidemiology and mechanism. Glaucoma. R. Hitchings. London, BMJ Publishing Group.

Frackowiak Richard, S. J. (1997). Human brain function. San Diego ; London, Academic Press.

Friston, K. J., Holmes, A. P., Worsley, K. J. and Poline, J. B. (1995). "Statistical parametric maps in functional imaging: a general linear approach." Human Brain Mapping 2: 189-210.

Gaida, G. (1990). Prespectives and limits of three-dimensional microscopy. Handbook of biological confocal microscopy. B. Pawley James. NY, Plenum**:** 253-57.

Gardiner, S. K. (2003). Statistical methods for the analysis of visual field data in glaucoma, The Nottingham Trent University, UK.

Gardiner, S. K. and Crabb, D. P. (2002). "Examination of different pointwise linear regression methods for determining visual field progression." Invest Ophthalmol Vis Sci 43(5): 1400-7.

Garway-Heath, D. (2000). The identification of progression in cupping of the optic disc. Glaucoma. R. Hitchings. London, BMJ Publishing.

Gonzalez, R. and Woods, R. (2002). Digital image processing. Upper Saddle River, NJ, Prentice Hall.

Gonzalez, R. C., Woods, R. E. and Eddins, S. L. (2004). Digital Image processing using MATLAB. Upper Saddle River, N. J., Pearson Prentice Hall.

Good, P. I. (2000). Permutation tests: A practical guide to resampling methods for testing hypotheses, Springer-Verlag, Heidelberg.

Goodman, J. (1996). Introduction to Fourier Optics. Singapore, McGraw-Hill.

Gordon, M. O., Beiser, J. A., Brandt, J. D., Heuer, D. K., Higginbotham, E. J., Johnson, C. A., Keltner, J. L., Miller, J. P., Parrish, R. K., 2nd, Wilson, M. R. and Kass, M. A. (2002). "The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma." Arch Ophthalmol 120(6): 714-20; discussion 829-30.

Guan, K., Hudson, C. and Flanagan, J. G. (2004). "Comparison of Heidelberg Retina Tomograph II and Retinal Thickness Analyzer in the assessment of diabetic macular edema." Invest Ophthalmol Vis Sci 45(2): 610-6.

Gundersen, K. and Asman, P. (2000). "Comparision of ranked segment analysis (RSA) and cup to disc ratio in computer-assisted optic disc evaluation." Acta Ophthalmol Scand 78: 137-141.

Haley, M. (1987). The Field Analyzer Primer. Allergan Humprey, San Leandro, CA.

Hanisch, R., White, R. and Gilliland, R. (1997). Deconvolution of Hubble Space Telescope Images and Spectra. Deconvolution of Images and Spectra. J. PA. NY, Academic Press**:** 310-360.

Harrington, D. O. (1976). The Visual Fields. A Textbook and Atlas of Clinical Perimetry. St Louis.

Hatch, W. V., Flanagan, J. G., Etchells, E. E., Williams-Lyn, D. E. and Trope, G. E. (1997). "Laser scanning tomography of the optic nerve head in ocular hypertension and glaucoma." Br J Ophthalmol 81(10): 871-6.

Hatch, W. V., Flanagan, J. G., Williams-Lyn, D. E., Buys, Y. M., Farra, T. and Trope, G. E. (1999). "Interobserver agreement of Heidelberg retina tomograph parameters." J Glaucoma 8(4): 232-7.

Hayasaka, S. and Nichols, T. E. (2003). "Validating cluster size inference: random field and permutation methods." Neuroimage 20(4): 2343-56.

Hayasaka, S. and Nichols, T. E. (2004). "Combining voxel intensity and cluster extent with permutation test framework." Neuroimage 23(1): 54-63.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J. and Nichols, T. E. (2004). "Nonstationary cluster-size inference with random field and permutation methods." Neuroimage 22(2): 676-87.

Heijl, A., Leske, M. C., Bengtsson, B., Hyman, L. and Hussein, M. (2002). "Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial." Arch Ophthalmol 120(10): 1268-79.

Heijl, A., Lindergren, G. and Lindergren, A. (1991). Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. Perimetry Update 1990/1991. M. RP and A. Heijl. Amsterdam, Kugler & Ghedini**:** 303-315.

Henson, D. (2000). Visual fields. Oxford, Butterworth-Heinemann.

Hitchings, R. (2000). Glaucoma. London, BMJ Publishing Group.

Holmes, A. P., Blair, R. C., Watson, J. D. and Ford, I. (1996). "Nonparametric analysis of statistic images from functional mapping experiments." J Cereb Blood Flow Metab 16(1): 7-22.

Holmes, T., Bhattacharyya, S., Cooper, J., Hanzel, D., Krishnamurthi, V., Lin, W., Roysam, B., Szarowski, D. and Turner, J. (1995). Light Microscopy Images Reconstructed by Maximum Likelihood Deconvolution. Handbook of Biological and Confocal Microscopy. J. Pawley. NY, Plenum Press**:** 389-402.

Holmes, T. J. (1992). "Blind deconvolution of quantum-limited incoherent imagery." Journal of Optical Society America 9(1052-1061).

Iester, M. and Mermoud, A. (2005). "Retinal nerve fiber layer measured by Heidelberg retina tomograph and nerve fiber analyzer." Eur J Ophthalmol 15(2): 246-54.

Iester, M., Mikelberg, F. S. and Drance, S. M. (1997). "The effect of optic disc size on diagnostic precision with the Heidelberg retina tomograph." Ophthalmology 104(3): 545-8.

Jansson, P. (1997). Deconvolution of Images and Spectra. NY, Academic Press.

Jay, J. L. and Allan, D. (1989). "The benefit of early trabeculectomy versus conventional management in primary open angle glaucoma relative to severity of disease." Eye 3 (Pt 5): 528-35.

Jonas, J. and Garway-Heath, D. (2000). Primary glaucomas: optic disc features. Glaucoma. R. Hitchings. London, BMJ Publishing.

Jonas, J. B., Budde, W. M. and Panda-Jonas, S. (1999). "Ophthalmoscopic evaluation of the optic nerve head." Surv Ophthalmol 43(4): 293-320.

Jones, R. (1999). Introduction to MFC programming with Visual C++. New Jersey, USA, Prentice-Hall Inc.

Kamal, D., Garway-Heath, D., Ruben, S., O'Sullivan, F., Bunce, C., Viswanathan, A., Franks, W. and Hitchings, R. (2003). "Results of the betaxolol versus placebo treatment trial in ocular hypertension." Graefes Arch Clin Exp Ophthalmol 241(3): 196-203.

Kamal, D. S., Garway-Heath, D. F., Hitchings, R. A. and Fitzke, F. W. (2000). "Use of sequential Heidelberg retina tomograph images to identify changes at the optic disc in ocular hypertensive patients at risk of developing glaucoma." Br J Ophthalmol 84(9): 993-8.

Kamal, D. S., Viswanathan, A. C., Garway-Heath, D. F., Hitchings, R. A., Poinoosawmy, D. and Bunce, C. (1999). "Detection of optic disc change with the Heidelberg retina tomograph before confirmed visual field change in ocular hypertensives converting to early glaucoma." Br J Ophthalmol 83(3): 290-4.

Kempen, G. v., Vliet, L. v. and Verveer, P. (1997). Application of image restoration methods for confocal fluorescence microscopy. 3-D Microscopy: Image Acquisition and Processing IV, SPIE.

Kerrigan-Baumrind, L. A., Quigley, H. A., Pease, M. E., Kerrigan, D. F. and Mitchell, R. S. (2000). "Number of ganglion cells in glaucoma eyes

compared with threshold visual field tests in the same persons." Invest Ophthalmol Vis Sci 41(3): 741-8.

Klein, B. E., Klein, R., Sponsel, W. E., Franke, T., Cantor, L. B., Martone, J. and Menage, M. J. (1992). "Prevalence of glaucoma. The Beaver Dam Eye Study." Ophthalmology 99(10): 1499-504.

Krieglstein, G. (2000). Medical treatment of glaucoma. Glaucoma. R. Hitchings. London, BMJ Publications.

Lehmann, T. M., Gonner, C. and Spitzer, K. (1999). "Survey: interpolation methods in medical image processing." IEEE Trans Med Imaging 18(11): 1049-75.

Leske, M. C. (1983). "The epidemiology of open-angle glaucoma: a review." Am J Epidemiol 118(2): 166-91.

Leske, M. C., Connell, A. M., Schachat, A. P. and Hyman, L. (1994). "The Barbados Eye Study. Prevalence of open angle glaucoma." Arch Ophthalmol 112(6): 821-9.

Libby, R. T., Gould, D. B., Anderson, M. G. and John, S. W. (2005). "Complex Genetics of Glaucoma Susceptibility." Annu Rev Genomics Hum Genet.

Linner, E. and Stromberg, U. (1967). "Ocular hypertension: A five year study of the total population in a Swedish town, and subsequent discussion." Glaucoma: Tutzing Symposium. Basel: Karger 187.

Lucy, L. (1974). "An iterative technique for the rectification of observed distributions." Astronomical Journal 79(June): 745-754.

Manly, B. (1991). Randomization and Monte Carlo methods in biology. London, Chapman and Hall.

Medeiros, F. A., Zangwill, L. M., Bowd, C., Sample, P. A. and Weinreb, R. N. (2005). "Use of progressive glaucomatous optic disk change as the reference standard for evaluation of diagnostic tests in glaucoma." Am J Ophthalmol 139(6): 1010-8.

Migdal, C., Gregory, W. and Hitchings, R. (1994). "Long-term functional outcome after early surgery compared with laser and medicine in open-angle glaucoma." Ophthalmology 101(10): 1651-6; discussion 1657.

Mikelberg, F., Pafitt, C., Swindale, N., Graham, S., Drance, S. and Gosine, R. (1995). "Ability of the Heidelberg Retinal Tomograph to detect early glaucomatous visual field loss." J Glaucoma 4: 242-247.

Mikelberg, F., Wijsman, K. and Schulzer, M. (1993). "Reproducibility of topographic parameters obtained with the Heidelberg Retina Tomograph." J Glaucoma 2(101-103).

Mitchell, P., Smith, W., Attebo, K. and Healey, P. R. (1996). "Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study." Ophthalmology 103(10): 1661-9.

Neter, J., Wasserman, W. and Kutner, M. (1985). Applied Statistical Models. Homewood, IL.

Newell, F. W. and Krill, A. E. (1964). "Diurnal Tonography in Normal and Glaucomatous Eyes." Trans Am Ophthalmol Soc 62: 349-74.

Nichols, T. E. and Holmes, A. P. (2002). "Nonparametric permutation tests for functional neuroimaging: a primer with examples." Hum Brain Mapp 15(1): 1-25.

Nouri-Mahdavi, K., Hoffman, D., Gaasterland, D. and Caprioli, J. (2004). "Prediction of visual field progression in glaucoma." Invest Ophthalmol Vis Sci 45(12): 4346-51.

Nourrit, V., Vohnsen, B. and Artal, P. (2005). "Blind deconvolution for high-resolution confocal scanning laser ophthalmoscopy." J. Opt. A: Pure Appl. Opt. 7: 585–592.

Odberg, T. and Riise, D. (1985). "Early diagnosis of glaucoma. The value of successive stereophotography of the optic disc." Acta Ophthalmol (Copenh) 63(3): 257-63.

Orgul, S., Croffi, G. A. and Van Buskirk, E. M. (1997). "Variability of contour line alignment on sequential images with the Heidelberg Retina Tomograph." Graefes Arch Clin Exp Ophthalmol 235(2): 82-6.

Oster, S. F., Deiner, M., Birgbauer, E. and Sretavan, D. W. (2004). "Ganglion cell axon pathfinding in the retina and optic nerve." Semin Cell Dev Biol 15(1): 125-36.

Patterson, A., Garway-Heath, D. and Crabb, D. (2006). "Improving the repeatability of topographic height measurements in confocal scanning laser imaging using maximum-likelihood deconvolution." Invest Ophthalmol Vis Sci (In press).

Patterson, A. J., Garway-Heath, D. F., Strouthidis, N. G. and Crabb, D. P. (2005). "Beyond the parameters: Mapping areas of structural change in longitudinal series of optic disc images." British Journal of Ophthalmology (under review).

Patterson, A. J., Garway-Heath, D. F., Strouthidis, N. G. and Crabb, D. P. (2005). "A new statistical approach for quantifying change in series of retinal and optic nerve head topography images." Invest Ophthalmol Vis Sci 46: 1659-1667.

Pawley, J. (1990). Handbook of biological confocal microscopy, New York, N.Y., Plenum.

Pederson, J. and Anderson, D. (1980). "The mode of progressive disc cupping in ocular hypertension and glaucoma." Ophthalmology 98: 490-5.

Pohjanpelto, P. E. and Palva, J. (1974). "Ocular hypertension and glaucomatous optic nerve damage." Acta Ophthalmol (Copenh) 52(2): 194-200.

Poline, J. B., Worsley, K. J., Evans, A. C. and Friston, K. J. (1997). "Combining spatial extent and peak intensity to test for activations in functional imaging." Neuroimage 5(2): 83-96.

Press, W. H. (2002). Numerical recipes in C and C[plus plus] : the art of scientific computing. Cambridge, Cambridge University Press.

Quigley, H. A., Dunkelberger, G. R. and Green, W. R. (1989). "Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma." Am J Ophthalmol 107(5): 453-64.

Reus, N. J. and Lemij, H. G. (2004). "Diagnostic accuracy of the GDx VCC for glaucoma." Ophthalmology 111(10): 1860-5.

Richardson, W. (1972). "Bayesian Based Iterative Method of Image Restoration." Journal of the Optical Society of America 62(Janurary): 55-59.

Rohrschneider, K., Burk, R. O., Kruse, F. E. and Volcker, H. E. (1994). "Reproducibility of the optic nerve head topography with a new laser tomographic scanning device." Ophthalmology 101(6): 1044-9.

Rooms, F., Philips, W. and Oostveldt, P. V. (2003). Estimation of Anisotropic Blur for the Restoration of confocal images. SPIE Conference - Wavelet Applications in Industrial Processing, Providence, Rhode Island, USA, SPIE.

Schwartz, A. L., Love, D. C. and Schwartz, M. A. (1985). "Long-term follow-up of argon laser trabeculoplasty for uncontrolled open-angle glaucoma." Arch Ophthalmol 103(10): 1482-4.

Schwartz, B. (1976). "The optic disc in glaucoma: introduction." Trans Am Acad Ophthalmol Otolaryngol 81(2): 191.

Shiose, Y., Kitazawa, Y., Tsukahara, S., Akamatsu, T., Mizokami, K., Futa, R., Katsushima, H. and Kosaki, H. (1991). "Epidemiology of glaucoma in Japan--a nationwide glaucoma survey." Jpn J Ophthalmol 35(2): 133-55.

Sihota, R., Gulati, V., Agarwal, H. C., Saxena, R., Sharma, A. and Pandey, R. M. (2002). "Variables affecting test-retest variability of Heidelberg Retina Tomograph II stereometric parameters." J Glaucoma 11(4): 321-8.

Siik, S., Chylack, L. T., Jr., Friend, J., Wolfe, J., Teikari, J., Nieminen, H. and Airaksinen, P. J. (1999). "Lens autofluorescence and light scatter in relation to the lens opacities classification system, LOCS III." Acta Ophthalmol Scand 77(5): 509-14.

Smith, S. D., Katz, J. and Quigley, H. A. (1996). "Analysis of progressive change in automated visual fields in glaucoma." Invest Ophthalmol Vis Sci 37(7): 1419-28.

Sommer, A., Katz, J., Quigley, H. A., Miller, N. R., Robin, A. L., Richter, R. C. and Witt, K. A. (1991). "Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss." Arch Ophthalmol 109(1): 77-83.

Sommer, A., Pollack, I. and Maumenee, A. (1979). "Optic disc parameters ona onset of glaucomatous field loss. I. Methods and progressive changes in disc morphology." Arch Ophthalmol 97: 1444-8.

Sommer, A., Tielsch, J. M., Katz, J., Quigley, H. A., Gottsch, J. D., Javitt, J. and Singh, K. (1991). "Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey." Arch Ophthalmol 109(8): 1090-5.

Spaeth, G. L., Hitchings, R. A. and Sivalingam, E. (1976). "The optic disc in glaucoma: pathogenetic correlation of five patterns of cupping in chronic open-angle glaucoma." Trans Am Acad Ophthalmol Otolaryngol 81(2): 217-23.

Strouthidis, N., White, E., Ho, T. and Garway-Heath, D. (2005). "Improving the repeatability of Heidelberg Retina Tomograph and Heidelberg Retina Tomograph II measurements." Br J Ophthalmol In Press.

Strouthidis, N., White, E., Owen, V. M., Ho, T., Hammond, C. and Garway-Heath, D. (2005). "Factors affecting the test-retest variability of Heidelberg Retina Tomograph and Heidelberg Retina Tomograph II measurements." Br J Ophthalmol In Press.

Swindale, N. V., Stjepanovic, G., Chin, A. and Mikelberg, F. S. (2000). "Automated analysis of normal and glaucomatous optic nerve head topography images." Invest Ophthalmol Vis Sci 41(7): 1730-42.

Tan, J. C. and Hitchings, R. A. (2003). "Approach for identifying glaucomatous optic nerve progression by scanning laser tomography." Invest Ophthalmol Vis Sci 44(6): 2621-6.

Tielsch, J. M., Katz, J., Sommer, A., Quigley, H. A. and Javitt, J. C. (1995). "Hypertension, perfusion pressure, and primary open-angle glaucoma. A population-based assessment." Arch Ophthalmol 113(2): 216-21.

Tielsch, J. M., Sommer, A., Katz, J., Royall, R. M., Quigley, H. A. and Javitt, J. (1991). "Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey." Jama 266(3): 369-74.

Tonnu, P. A., Ho, T., Newson, T., El Sheikh, A., Sharma, K., White, E., Bunce, C. and Garway-Heath, D. (2005). "The influence of central corneal thickness and age on intraocular pressure measured by pneumotonometry, non-contact

tonometry, the Tono-Pen XL, and Goldmann applanation tonometry." Br J Ophthalmol 89(7): 851-4.

Tonnu, P. A., Ho, T., Sharma, K., White, E., Bunce, C. and Garway-Heath, D. (2005). "A comparison of four methods of tonometry: method agreement and interobserver variability." Br J Ophthalmol 89(7): 847-50.

Uchida, H., Brigatti, L. and Caprioli, J. (1996). "Detection of structural damage from glaucoma with confocal laser image analysis." Invest Ophthalmol Vis Sci 37: 2396-2401.

Van-Trees, H. (1968). Detection, Estimation, and Modulation Theory. NY, Wiley.

Vernon, S. A., Hawker, M. J., Ainsworth, G., Hillman, J. G., Macnab, H. K. and Dua, H. S. (2005). "Laser scanning tomography of the optic nerve head in a normal elderly population: the Bridlington eye assessment project." Invest Ophthalmol Vis Sci 46(8): 2823-8.

Vetterling, W. and Press, W. (2002). Numerical recipes example book (C++). Cambridge, Cambridge Press.

Vieira, P., Manivannan, A., Lim, C. S., Sharp, P. and Forrester, J. V. (1999). "Tomographic reconstruction of the retina using a confocal scanning laser ophthalmoscope." Physiological Measurements 20: 1-19.

Viswanathan, A. C., Crabb, D. P., McNaught, A. I., Westcott, M. C., Kamal, D., Garway-Heath, D. F., Fitzke, F. W. and Hitchings, R. A. (2003). "Interobserver agreement on visual field progression in glaucoma: a comparison of methods." Br J Ophthalmol 87(6): 726-30.

Viswanathan, A. C., Fitzke, F. W. and Hitchings, R. A. (1997). "Early detection of visual field progression in glaucoma: a comparison of PROGRESSOR and STATPAC 2." Br J Ophthalmol 81(12): 1037-42.

Webb, R. and Hughes, G. (1981). "Scanning laser ophthalmoscope." IEEE Trans Biomed Eng 28: 488-492.

Weinreb, R. N. and Khaw, P. T. (2004). "Primary open-angle glaucoma." Lancet 363(9422): 1711-20.

Weinreb, R. N., Lusky, M., Bartsch, D. U. and Morsman, D. (1993). "Effect of repetitive imaging on topographic measurements of the optic nerve head." Arch Ophthalmol 111(5): 636-8.

Werner, E. (1991). Manual of visual fields. Churchill Livingstone, NY.

Werner, E. B., Bishop, K. I., Koelle, J., Douglas, G. R., LeBlanc, R. P., Mills, R. P., Schwartz, B., Whalen, W. R. and Wilensky, J. T. (1988). "A comparison of experienced clinical observers and statistical tests in detection of progressive visual field loss in glaucoma using automated perimetry." Arch Ophthalmol 106(5): 619-23.

Wollstein, G., Garway-Heath, D. F. and Hitchings, R. A. (1998). "Identification of early glaucoma cases with the scanning laser ophthalmoscope." Ophthalmology 105(8): 1557-63.

Woo, M., Neider, J., Davis, T. and Shreiner, D. (1997). OpenGL programming guide: the official guide to learning OpenGL, Version 1.2. Indianapolis, USA, Addison Wesley.

Worsley, K. J., Evans, A. C., Marrett, S. and Neelin, P. (1992). "A three-dimensional statistical analysis for CBF activation studies in human brain." J Cereb Blood Flow Metab 12(6): 900-18.

www.heidelbergengineering.com How to interpret progression, Heidelberg Engineering. Available at: http://www.heidelbergengineering.com.

www.imatix.com Standard Function Library, Available at: http:/www.imatix.com.

www.matrox.com Matrox Imaging Library, Matrox Electronic Systems Ltd, Berkshire, UK. Available: http://www.matrox.com.

www.msdn.com "Microsoft Foundation Classes (Visual C++), Available at: http://www.msdn.com."

www.nr.com "Numerical Recipes, Cambridge, MA, USA. Available at: http://www.nr.com."

www.vtk.org Visualization Toolkit, Available at http://www.vtk.org.

Yagci, R., Eksioglu, U., Midillioglu, I., Yalvac, I., Altiparmak, E. and Duman, S. (2005). "Central corneal thickness in primary open angle glaucoma,

pseudoexfoliative glaucoma, ocular hypertension, and normal population." Eur J Ophthalmol 15(3): 324-8.

Zangwill, L., Irak, I., Berry, C. C., Garden, V., de Souza Lima, M. and Weinreb, R. N. (1997). "Effect of cataract and pupil size on image quality with confocal scanning laser ophthalmoscopy." Arch Ophthalmol 115(8): 983-90.

Zinser, G., Wijnaendts-van-Resandt, R., Dreher, A., Weinreb, R., Harbarth, U. and Burk, R. (1989). "Confocal scanning laser tomography of the eye." Proc SPIE 1161: 337-344.

List of Publications

Manuscripts

- Patterson, A. J., D. F. Garway-Heath, N. G. Strouthidis and D. P. Crabb (2005). "A new statistical approach for quantifying change in series of retinal and optic nerve head topography images." Invest Ophthalmol Vis Sci 46: 1659-1667.
- Patterson, A. J., D. F. Garway-Heath, N. G. Strouthidis and D. P. Crabb (2005). "Beyond the parameters: Mapping areas of structural change in longitudinal series of optic disc images." British Journal of Ophthalmology (under review)
- Patterson, A., D. Garway-Heath and D. Crabb (2006). "Improving the repeatability of topographic height measurements in confocal scanning laser imaging using maximum-likelihood deconvolution." Invest Ophthalmol Vis Sci (in press).

Oral Presentations:

- "Testing A New Approach To Detecting Change In Series Of Retinal Images Acquired From Scanning Laser Tomography", International Perimetric Society conference, Barcelona, Spain on June 29 - July 2, 2004.
- "Adaptive Blind Deconvolution: Improving the Repeatability of Topographic Height Measurements in Scanning Laser Tomography", Image Morphometry and Glaucoma in Europe Meeting, Milan, Italy on April 4-5, 2005
- "A New Approach for Quantifying Change in Series of Retinal and Optic Nerve Head Topography Images", UK and Eire Glaucoma Society Meeting, Nottingham, UK on Dec 2, 2005
- "Beyond the parameters: Detecting areas of structural progression in longitudinal series of optic disc images", Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, USA on April 30-May 4, 2006. (Read by Mr DF Garway-Heath).

Poster Presentations:

- "A Novel Method of Generating Mean Topography Images of the Optic Nerve Head", Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, USA on April 30-May 4, 2004.
- "Image Deconvolution Improves the Repeatability of Topographic Height Measurements in Scanning Laser Tomography", Association for Research in Vision and Ophthalmology, Fort Lauderdale, Florida, USA on May 1-May 5, 2005.