

Some pages of this thesis may have been removed for copyright restrictions.

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown policy](#) and contact the service immediately (openaccess@aston.ac.uk)

TECHNOLOGICAL ENHANCEMENTS TO OPTOMETRIC
CLINICAL TESTS

ALEC KINGSNORTH

Doctor of Philosophy

ASTON UNIVERSITY

March 2015

©Alec Kingsnorth, 2015

Alec Kingsnorth asserts his moral right to be identified as the
author of this thesis

This copy of the thesis has been supplied on the condition that
anyone who consults it is understood to recognise that its
copyright rests with its author and that no quotation from the
thesis and no information derived from it may be published
without appropriate permission or acknowledgement.

Thesis Abstract

A sizeable amount of the testing in eye care, requires either the identification of targets such as letters to assess functional vision, or the subjective evaluation of imagery by an examiner. Computers can render a variety of different targets on their monitors and can be used to store and analyse ophthalmic images. However, existing computing hardware tends to be large, screen resolutions are often too low, and objective assessments of ophthalmic images unreliable.

Recent advances in mobile computing hardware and computer-vision systems can be used to enhance clinical testing in optometry. High resolution touch screens embedded in mobile devices, can render targets at a wide variety of distances and can be used to record and respond to patient responses, automating testing methods. This has opened up new opportunities in computerised near vision testing. Equally, new image processing techniques can be used to increase the validity and reliability of objective computer vision systems.

Three novel apps for assessing reading speed, contrast sensitivity and amplitude of accommodation were created by the author to demonstrate the potential of mobile computing to enhance clinical measurement. The reading speed app could present sentences effectively, control illumination and automate the testing procedure for reading speed assessment. Meanwhile the contrast sensitivity app made use of a bit stealing technique and swept frequency target, to rapidly assess a patient's full contrast sensitivity function at both near and far distances. Finally, customised electronic hardware was created and interfaced to an app on a smartphone device to allow free space amplitude of accommodation measurement.

A new geometrical model of the tear film and a ray tracing simulation of a Placido disc topographer were produced to provide insights on the effect of tear film breakdown on ophthalmic images. Furthermore, a new computer vision system, that used a novel eye-lash segmentation technique, was created to demonstrate the potential of computer vision systems for the clinical assessment of tear stability.

Studies undertaken by the author to assess the validity and repeatability of the novel apps, found that their repeatability was comparable to, or better, than existing clinical methods for reading speed and contrast sensitivity assessment. Furthermore, the apps offered reduced examination times in comparison to their paper based equivalents. The reading speed and amplitude of accommodation apps correlated highly with existing methods of assessment supporting their validity. Their still remains questions over the validity of using a swept frequency sine-wave target to assess patient's contrast sensitivity functions as no clinical test provides the range of spatial frequencies and contrasts, nor equivalent assessment at distance and near. A validation study of the new computer vision system found that the authors tear metric correlated better with existing subjective measures of tear film stability than those of a competing computer-vision system. However, repeatability was poor in comparison to the subjective measures due to eye lash interference.

The new mobile apps, computer vision system, and studies outlined in this thesis provide further insight into the potential of applying mobile and image processing technology to enhance clinical testing by eye care professionals.

Acknowledgments

This thesis could not of been undertaken without the help, support and friendship of the following people.

Firstly I would like to thank my supervisor Prof James Wolffsohn and my associate supervisor Dr Amy Sheppard for their excellent supervision, answering questions both great and small, and guiding me through the PhD process.

A special thanks to Dr Thomas Drew whose technical management and insights have been an inspiration; always there to offer guidance, support and enthusiasm to all of my endeavors. I would also like to thank the rest of the of my friends and colleagues in MB171. Notably, Ben, Joe, Jon and Ria whose warmth, friendship and humor have made this PhD enjoyable.

I must also thank my family. My mother, Jenny, whose constant trips to the library as a child inspired my love of science and learning. My father, Terry, whose love of tinkering with motorcycles inspired me as an engineer. And to my brothers and sisters, Brett, Sally, Kimberley and Jessica, who have all offered much needed love and support.

Finally, to my long suffering fiancée Jessica Cheung. You have been phenomenal.

Contents

Abstract	2
Acknowledgments	3
Nomenclature	8
1 Introduction	13
1.1 Mobile computing	13
1.1.1 Mobile device adoption	13
1.1.2 Hardware	15
1.1.3 Software Development	18
1.2 Mobile computing in optometry	19
1.2.1 Imaging	19
1.2.2 Refraction	22
1.2.3 Test charts, functional vision and psychophysics	22
1.2.4 Discussion and Opportunities	25
1.3 Image processing	28
1.3.1 Image processing theory	28
1.3.2 Dynamic imaging	31
1.4 Aim	34
2 The Development and Validation of an Objective Reading Speed Test.	36
2.1 Introduction	36
2.1.1 Visual impairment and reading speed	37
2.1.2 Reading presentation methods	40
2.1.3 Reading Texts	41
2.1.4 Reading speed curve and metrics	43
2.1.5 Current charts, texts and devices	44
2.1.6 Outcomes	46

2.2	Aims and objectives	49
2.2.1	Aim	49
2.2.2	Objectives	49
2.3	Development-Aston Read	50
2.3.1	Mobile device and Software development kit	50
2.3.2	Software	54
2.4	Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test	59
2.4.1	Aim	59
2.4.2	Subjects	59
2.4.3	Protocol	59
2.4.4	Statistical testing	62
2.4.5	Study results	64
2.4.6	Discussion and Conclusion	73
2.5	Further development of Aston Read	75
2.5.1	Additions and improvements	75
3	Development of a Mobile, Near and Distance Contrast Sensitivity Test.	82
3.1	Introduction	82
3.1.1	Contrast sensitivity function	83
3.1.2	Threshold Measurement	88
3.1.3	Charts and Instrumentation	93
3.1.4	Outcomes-What makes a good clinical test for contrast sensitivity?	96
3.2	Aims and Objectives	98
3.2.1	Aim	98
3.2.2	Objectives	98
3.3	Development	99
3.3.1	Concept	99
3.3.2	Profiling and Calibration	101
3.3.3	Stimuli design	106
3.3.4	Software-Near Contrast Sensitivity	109
3.3.5	Software and Hardware-Distance Contrast Sensitivity	112

3.4	Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods	116
3.4.1	Aim	116
3.4.2	Subjects	116
3.4.3	Protocol	116
3.4.4	Statistical Testing	117
3.4.5	Results	119
3.4.6	Discussion and conclusion	128
4	The Development Of A Computerised, Mobile Test For Subjective Amplitude of Accommodation	132
4.1	Introduction	132
4.1.1	Objective AOA	133
4.1.2	Subjective AOA	134
4.1.3	Background summary	140
4.2	Aims and objectives	141
4.2.1	Aim	141
4.2.2	Objectives	141
4.3	Development-Aston Accommodometer	142
4.3.1	Distance measurement	142
4.3.2	Concept	145
4.3.3	Hardware	146
4.3.4	Software	150
4.4	Study - Comparison between the Aston Accommodometer and the RAF rule	155
4.4.1	Aim	155
4.4.2	Subjects	155
4.4.3	Protocol	155
4.4.4	Statistical testing	156
4.4.5	Results	157
4.4.6	Discussion and Conclusions	165
5	A New Model For Simulating Localized Tear Film Breakdown in Video-Keratometry images	167
5.1	Introduction	167
5.1.1	Tear Film Functions	168
5.1.2	Tear Film Structure.	170

5.1.3	Tear Film Maintenance, Thinning, Rupture and Dry eye. . . .	172
5.1.4	Placido based devices and modelling	179
5.1.5	Current Models	180
5.1.6	Outcome	183
5.2	Aims and objectives	185
5.2.1	Aim	185
5.2.2	Objectives	185
5.3	New Combined Cornea and Tear Break-up Models	186
5.3.1	Parametric method	186
5.3.2	Numerical approximation	190
5.4	Ray Tracing	195
5.4.1	Ray tracer	195
5.4.2	Diffuse and specular reflection	195
5.4.3	Simulation Results and Discussion	198
5.5	Conclusion	202
6	Automated Assessment of Tear Film Stability	203
6.1	Introduction	203
6.1.1	Fluorescein tear film break up time	204
6.1.2	Non-invasive break up time	206
6.1.3	Interferometry	207
6.1.4	Video Keratography tear stability measures	209
6.1.5	Outcome	213
6.2	Aims and Objectives	215
6.2.1	Aim	215
6.2.2	Objectives	215
6.3	Hardware Development	216
6.3.1	Design and manufacture	216
6.4	Development	223
6.4.1	Blink detection	224
6.4.2	Centre detection	225
6.4.3	ROI - Stage 1, Eye lid removal	227
6.4.4	ROI - Stage 2, Eye lash removal	230
6.4.5	Tear metric	232
6.5	Study	234
6.5.1	Aim	234

6.5.2	Subjects	234
6.5.3	Protocol	234
6.5.4	Statistics	235
6.5.5	Results	236
6.5.6	Discussion and conclusions	243
7	Conclusion	246
7.1	Introduction	246
7.2	Discussions and conclusions	246
	Bibliography	252

List of Figures

1.1	Smart-phone usage and biomedical app statistics	14
1.2	User interface builder [taken from Apple, 2014]	18
1.3	Retinal imaging attachment taken from [Maamari et al., 2014].	21
1.4	Diagram of the most commonly used image enhancement techniques	29
1.5	Enhancement of raw Placido topography images	29
1.6	Segmentation in pupillometry images	30
1.7	Segmentation in pupillometry images	30
1.8	Quantification of retinal images, disc and cup ratios	31
1.9	Thesis structure	35
2.1	Reading rate versus print sizes taken from a clinical reading adapted from [Legge et al., 1985a]	44
2.2	iPad screen calibration	54
2.3	Aston Read - User Interface	57
2.4	Aston Read - Programming flow diagram	58
2.5	Sentence presentation	60
2.6	Raw reading speed data and objective CPS measurement	62
2.7	Mean reading speed results	64
2.8	Difference plot RR versus Chart 1	65
2.9	Difference plot RR versus Chart 2	65
2.10	Difference plot Chart 1 versus Chart 2	66
2.11	Difference plot RR versus Chart 1	67
2.12	Difference plot RR versus Chart 2	68
2.13	Difference plot Chart 1 versus Chart 2	68
2.14	Difference plot RR versus Chart 1	69
2.15	Difference plot RR versus Chart 2	70
2.16	Difference plot Chart 1 versus Chart 2	70
2.17	Difference plot RR versus Chart 1	71
2.18	Difference plot RR versus Chart 2	72

2.19	Difference plot Chart 1 versus Chart 2	72
2.20	Audio sample of a single Radner sentence	76
2.21	Estimated outcomes of distance measurement	79
2.22	Tested outcomes of distance measurement	80
2.23	Eye Detection and pixel separation.	81
3.1	Human Contrast Sensitivity Function	83
3.2	Multiple Spatial Channels and adaptation	84
3.3	Electrophysiology of macaque striate cells	87
3.4	Detection theory	91
3.5	Signal Detection Theory	92
3.6	Paper Based Contrast Sensitivity Charts	95
3.7	iPad concept adapted from charts first described by Campbell and Robson [1968]	100
3.8	Luminance profiles of the R,G and B channels for an iPad Screen . .	102
3.9	Luminance profiles of the R,G and B channels for the HDTV	103
3.10	Luminance steps available around 50% of the peak luminance	105
3.11	Luminance Matrix	106
3.12	Swept frequency characteristics for the near contrast sensitivity chart	108
3.13	Near Contrast Sensitivity Chart	108
3.14	Near Contrast Sensitivity App User Interface	110
3.15	Near Contrast Sensitivity Programming Flow Diagram	111
3.16	Distance Contrast Sensitivity Hardware	112
3.17	Distance Contrast Sensitivity App User Interface	114
3.18	Distance contrast sensitivity programming flow diagram	115
3.19	Mean results for twenty patients for all charts and occlusion conditions	120
3.20	Pelli-Robson Repeatability Bland-Altman Plots, n=20	124
3.21	Near App Repeatability Bland-Altman Plots, n=20	125
3.22	Distance App Repeatability Bland-Altman Plots , n=20	126
3.23	CSV-1000 Repetability Bland-Altman Plots, n=20	127
4.1	Near and far points of accommodation	133
4.2	Pseudo accommodation caused by pupillary constriction	135
4.3	Mean accommodative amplitudes for subjective and objective meth- ods taken from [Ostrin and Glasser, 2004]	136
4.4	R.A.F near point rule	136
4.5	Comparison of objective and subjective measures of AOA and age. .	138

4.6	Compact accommodometer taken from [Ide et al., 2012]	139
4.7	Calibration test rig and model head	144
4.8	Calibration results	145
4.9	Operating principles of the Aston Accommodometer	146
4.10	Component and pin out diagram of the Aston Accommodometer	147
4.11	Aston Accommodometer Prototype	149
4.12	Oscilloscope sampling of the iPhones microphone port for binary data modulated to FSK	150
4.13	Microcontroller flow diagram	151
4.14	User interface for the Aston Accommodometer	153
4.15	iPhone programming flow diagram	154
4.16	Bland Altman analyses for the push-up test	157
4.17	Bland Altman analyses for the pull-down test	160
4.18	Difference plot between the push-up and the pull-down technique for all testing methods	161
4.19	Difference plot between the 1st and 3rd repeated measure on each instrument for the push-up technique	163
4.20	Difference plot between the 1st and 3rd repeated measure on each instrument for the pull-down technique	164
5.1	Tear film location	168
5.2	The effect of an irregular tear film on the optical path of light	168
5.3	The classical trilaminar and the aqueous-mucin gel model of the tear film.	170
5.4	Lacrimal System	173
5.5	Lid movement and the build up phase.	176
5.6	Tear film break-up cross sections and 3D shape, taken from [Licznarski, Tomasz J and Kasprzak, Henryk T and Kowalik, 1999]	177
5.7	3D Tear film dimple	178
5.8	Placido based corneal topographer diagram	179
5.9	Illustration of Placido rings under different corneal conditions	180
5.10	Perturbation diagram	180
5.11	Geometrical model of a Placido disc topographer	181
5.12	Placido Cones and Spheres modeled as an equivalent flat Placido disc	181
5.13	Skew ray error	182
5.14	Spherical break up model	184

5.15	Conventions and relationships between parameters	188
5.16	Parametric tear film model.	189
5.17	Initial Mesh	191
5.18	Surface mesh	192
5.19	Diagram illustrating the rotation method	193
5.20	Creating and applying the tear dimple	194
5.21	Ray tracer	196
5.22	Diffuse and specular reflection	197
5.23	Specular only simulation results	199
5.24	Diffuse and Specular simulation results	200
5.25	Results of Alonso-Caneiro et al. [2011], who used a half-width of $p_{hw} = 1.8mm$ and depth of $a = 1\mu m$ taken from [Alonso-Caneiro et al., 2011]	201
5.26	Bulbous like tear distortions in real Placido topographer images . . .	201
6.1	Fluorescein staining pre and post break up	205
6.2	Objective fluorescein imaging taken from [Yedidya et al., 2007a] . .	206
6.3	Tearscope mires pre and post break-up	207
6.4	Lipid layer interference fringes as observed on two instruments . . .	208
6.5	Lateral shearing interferometry images taken from [Dubra et al., 2004]	209
6.6	Tear stability maps using topographical data taken from [Goto et al., 2003]	210
6.7	Image processing steps undertaken by [Iskander et al., 2005]	211
6.8	Background data subtraction to increase efficiency of edge detection algorithms	212
6.9	Placido Disc Concept	216
6.10	Illumination System	218
6.11	Illumination System	218
6.12	Beam detector setup	219
6.13	Imaging system designs taken from [Mejia-Barbosa and Malacara-Hernandez, 2001]	220
6.14	Ring spacing determination	221
6.15	Placido cone manufacture	222
6.16	Software concept for custom tear tear analysis	223
6.17	Pixel standard deviation plotted against frame number for two subjects.	224
6.18	Image processing steps in center detection	225
6.19	Image processing steps to exclude eyelids	227

6.20 ROI intensity images	229
6.21 Eyelash removal image processing steps	230
6.22 Tear metric image processing steps	233
6.23 OSDI, TSFQ_M and TSFQ_A Bland-Altman plots for repeatability	241
6.24 TSBUT, VBUT and IBT Bland-Altman plots for repeatability . . .	242
6.25 Segmentation faults in the TSFQ_A metric	244

List of Tables

1.1	Specifications of the Apples iPhone 4S [Apple Inc, California, US]	16
2.1	Visual impairment and reading speed studies	39
2.2	Criterion and iPad specification	53
2.3	ORS agreement and correlation values	67
2.4	RA agreement and correlation values	69
2.5	Objective CPS agreement and correlation values	71
2.6	Correlation of metrics between charts.	74
2.7	Repeatability of the Aston Read and the Radner Reading test.	74
3.1	Luminance Coefficients and Gamma Values	102
3.2	Bit stealing steps	105
3.3	Variables used in the contrast charts	107
3.4	Chart Characteristics	116
3.5	Paired t-tests at 1.5 cpd	121
3.6	Paired t-tests at 1.5 cpd	122
3.7	Coefficients of repeatability (COR) and Pearson's correlations using the 0.8 Bangerter foil	123
4.1	Component specification and rationale	148
4.2	Push-Up Correlation table	158
4.3	Pull-Down Correlation table	159
4.4	Push-up and Pull-Down Correlations	162
4.5	Pooled Standard Deviations for each instrument	163
4.6	Bland-Altman repeatability data	164
5.1	Summary of tear film functions	174
5.2	Simulation Parameters	198
6.1	Dry eye testing methods	204

6.2	Device summary	217
6.3	Image processing values used in the center detection	226
6.4	Image processing values used in the eye lid removal	228
6.5	Image processing parameters used in the removal of eye lashes	231
6.6	Image processing parameters used in determining the tear metric	232
6.7	Symptomatic dry eye as evaluated by the OSDI Questionnaire, n=29	236
6.8	Means and standard deviations for the metrics across Session 1, n=29	236
6.9	Spearman's rank correlation coefficients ρ between each metric in the first session, n=29	237
6.10	Symptomatic dry eye as evaluated by the OSDI Questionnaire, n=20	238
6.11	Means and standard deviations for the tear film metrics across session 2, n=20	238
6.12	Spearman's rank correlation coefficients ρ between each tear film measure in the second session n=20	239
6.13	Wilcoxon signed rank tests, Spearman's correlation coefficients and Bland Altman repeatability data	240

Nomenclature

(x_0, y_0)	Tear Dimple Centre Position
$\Delta\sigma(\theta, \varphi)$	Central Angle
θ	Longitude
φ	Latitude
a	Tear Dimple Depth
A_{hw}	Angular Half-Width
$c(i)$	Contrast at Matrix Row i
$c(x, y)$	Radial distance at position x, y
$C_{triangles}$	Vertex Connectivity Matrix
d	Distance
$D_{reading}$	Reading Distance
f	Focal Length
f_{end}	End Frequency
f_o	Starting Frequency
i	Matrix Column Position
I^2C	Inter-Integrated Circuit Serial Computer Bus
j	Matrix Row Position
k	Rate of Exponential Increase

$\text{LogMAR}_{\text{lastLineRead}}$	Reading Acuity of the Last Line Read
m	Matrix No. of Rows
$M_{\text{syllables}}$	Syllable Mistakes
n	Matrix No. of Columns
p_{hw}	Tear Dimple Half-Width
P_{render}	Pixels to Render
P_{hw}	Half Distance Along Arc Perimeter
R	Spherical Radius
r_0	Central Angle
R_{acuity}	Reading Acuity
R_c	Radius of the Cornea
R_{screen}	Screen Resolution
$r_{tf}(\theta, \varphi)$	Change In Radius Due To The Tear Film
S_{reading}	Reading Speed
S_{target}	Target Size
$TSFQ_A$	Authors Tear Film Surface Quality Indicator
$TSFQ_M$	Medmont's Tear Film Surface Quality Indicator
V_{coords}	Matrix of Vertex Coordinates
w	Screen Width
$x(j)$	Visual Angle at Matrix Row Postion j
X_{height}	X-height (Letter Size)
x_{max}	Maximum Visual Angle
$x_s(\theta, \varphi), y_s(\theta, \varphi), z_s(\theta, \varphi)$	Cartesian Coordinates For The Parametric Tear Model

$y_{i,j}$	Luminance Relative to Peak at Matrix position i, j
2-AFC	Two Alternative Forced Choice
3D	Three Dimensional
8-AFC	8 Alternative Forced Choice
ANOVA	Analysis of variance
AOA	Amplitude Of Accomodation
API	Application Programmable Interfaces
ARMD	Age Related Macular Degeneration
BUT	Break Up Time
CNC	Computer Numerical Control
COR	Coefficent of Repeatability
CPS	Critical Print Size
CRT	Cathode Ray Tube
CSF	Contrast Sensitivity Function
DAC	Digital to Analog Converter
DOF	Depth Of Field
ECG	Electrocardiogram
ETDRS	Early Treatment Diabetic Retinopathy Study
FACT	Functional Acuity Contrast Test
FBUT	Flourescein Break Up Time
FSK	Frequency Shift Keying
HDTV	High Definition Television
IBT	Inter-Blink Time

IR	Infra-Red
IREST	International Reading Speed Texts
LASIK	Laser-Assisted in situ Keratomileusis
LCD	Liquid Crystal Display
LED	Light Emitting Diode
MNREAD	Minnesota Near Vision Reading Test
MRS	Mean Reading Speed
NBC	Normal Blink Conditions
NETRA	Near Eye Tool For Refractive Assessment
NIBUT	Non Invasive Break Up Time
ORS	Optimal Reading Speed
OSDI	Ocular Surface Disease Index
PC	Personal Computer
PD	Pupillary Distance
qCSF	Quick Contrast Sensitivity Function
RA	Reading Acuity
RAF	Royal Air Force
RBU	Ring Break Up
ROC	Receiver Operating Curve
ROI	Region Of Interest
RR	Radner Reading
RSVP	Rapid Serial Visualisation Presentation
SAI	Surface Asymmetry Index

SBC	Suppressed Blink Conditions
SDK	Software Development Kit
SDT	Signal Detection Theory
SRI	Surface Regularity Index
TFBUT	Tear Film Break Up Time
TSBUT	Tearscope Break Up Time
V1	Primary Visual Cortex
VA	Visual Acuity
VBUT	Video-Keratoscope Break Up Time
VCTS	Vision Contrast Test System

1 Introduction

More than any other computing device to date, the rapid rise of smart phone and tablet computer ownership has revolutionised not only digital communication across the world, but our culture, community and sense of identity. The impact of this technological shift towards mobile computing can-not be understated; whole multi-billion pound industries have sprung up to serve content, computer programs, and accessories in response to the burgeoning demand of smart-phone owning consumers. As a by product of such demand, the increasing competition between smart-phone manufacturers has led to the development of ever increasingly sophisticated devices. One can expect accelerometers, gyroscopes, cameras, light sensors and high resolution screens from the most inexpensive of smart-phones. Both the ubiquitous nature of such devices and their plethora of sensors has in turn spurred significant software advances in the related areas of image processing, voice recognition and pattern detection.

It is clear that recent advances in technology, particularly those of mobile computing and image processing, have the potential to transform clinical practice in optometry. Thus there is great scope to leverage the power of these technologies, to create new and novel mobile applications that not only seek to digitize existing tests and practices in optometry, but which can enhance the measurement, time efficiency and practicality of optometric tests.

1.1 Mobile computing

1.1.1 Mobile device adoption

The rate of adoption of both smart-phones and tablets within the western world has been tremendous. Figure 1.1a displays a chart using survey data taken between 2010-2014 reporting the ownership of smart phones within the UK population [OF-COM, 2014]. Notice the rise in both smart phone and tablet adoption across the

years of the survey. While smart-phones in particular have reached over 60% of the population, the same survey found an uneven adoption across age groups, with the elderly less likely to own and use mobile computing devices than their younger contemporaries (65 years: 14% adoption, 16-24 years: 88% adoption). This may provide a barrier to medical app usage in healthcare today, particularly with older health care professionals that may benefit from using apps (or even the patients themselves) being unwilling to adopt them. However, the youth of today will be the elderly people of tomorrow, hence patients will become more familiar with their use, improving the prospect of using them for routine healthcare testing and monitoring.



(a) Adoption rates of mobile computing devices, data [taken from OFCOM 2014]

Figure 1.1: Smart-phone usage and biomedical app statistics

As of 2014 there were approximately 21,900 medical apps available on one of the most popular mobile platforms (Apples iOS) representing 2.11% of all apps currently available according to market research [PGBIZ, 2014]. More important is the use of mobile apps and devices within healthcare itself. A survey of junior medical practitioners found 72% have used medical apps to varying degrees within clinical practice [Payne et al., 2012], the majority of which use them for medication/drug reference [Grasso et al., 2006] or to view clinical guidelines [Carter et al., 2014]. Amongst eye care professionals such as ophthalmologists the use of smart-phones in practice has been equally high with a recent survey suggesting 83% of cataract and

refractive surgeons use their smart-phones to accomplish their professional responsibilities [Davis et al., 2010]. With both patients and clinicians willing to adopt and use mobile devices as part of their normal routine, the time is right to broaden their application within optometry.

1.1.2 Hardware

Much of the work presented in this thesis would not be possible without the technical advances in hardware associated with smart-phones. The specifications of the most popular smart-phone as of 2012, the Apple iPhone 4S, are listed in Table 1.1. This shows the breadth of communication devices, sensors, and user input available within such a compact pocket sized form factor.

Unique to the optometry profession, a sizeable amount of the testing requires the identification of targets such as letter charts or the use of specific targets for patients to fixate on during testing. Current computerized test charts, such as the popular Test Chart 2000 [Thompson Software Solutions, St Albans, UK] use standard LCD monitors which lack the screen resolution to be able to test patients at near distances, and so are limited to testing at distance. For example with a PC standard monitor resolution of 96 ppi the lowest acuity that can be tested at 40 cm is 0.36 logMAR. Paper based near vision tests, including those of reading [Luebker and Lamay, 1989], contrast sensitivity [Weatherhead, 1980], and acuity [Bailey and Lovie, 1980], have been created but they lack the flexibility of computerized test charts, making post test analysis of the results a laborious process. The screen resolution offered by most commercial tablets and smart-phones surpasses the resolution of existing custom made computerized charts. The iPhone 4s, for example, can render optotypes down to -0.17 logMAR at 40 cm. Moreover, due to the compact size and long battery life, mobile devices can be readily used at the examination chair making them ideal candidates to create new and novel near vision testing methods.

One of the interesting aspects of testing for near vision on mobile devices is that when testing within reaching distance of patients there is the opportunity to make use of the embedded touch screens to collect user feedback during clinical testing. This helps automate existing methods, releasing the examiner from the burden of not only applying the test but collecting and analysing the data manually. Furthermore, creating testing methods that can be implemented by the patient themselves, opens up the possibility of electronically monitoring patients near functional vision. For example, an app that tests for reading speed could be used to monitor the functional

Table 1.1: Specifications of the Apples iPhone 4S [Apple Inc, California, US]

Body	Dimensions:	115.2 x 58.6 X 9.3 mm
	Weight:	137g
Processor	Type	A5-ARM processor with on-board GPU
	Speed:	800 Mhz
	Ram:	512mb
Display	Type:	Capacitive Touch Screen LCD
	Resolution:	326 ppi
	Size:	960x640
Camera	Front:	0.3 megapixel
	Rear:	8.0 megapixel
Battery	Stand by time:	200 hrs
	Usage time:	6 hrs
Communication	GPRS EDGE WLAN Bluetooth	
Sensors	Accelerometer Gyro Proximity Compass Microphone	

vision of patients whose cataracts are progressing, alerting patients or health care professionals when the cataract has impaired vision enough for further treatment.

A myriad of communication devices embedded within mobile devices also have major implications for their application in healthcare. Data taken from these apps need not rely on the analysis of the immediate examiner, or even in the limited computing power of their microchips, but can be sent wirelessly for off site analysis. On a smaller scale, near by radio communication allows smart-phone users to interact with external instruments and sensors, providing custom controls and data display capabilities. This has happened in the consumer market with phones controlling light

switches, thermometers, speakers, and to a certain extent within the medical device industry with electrocardiogram (ECG) [Wong, 2013] and glucose monitoring apps [Pfützner et al., 2012] that rely on the use of external sensors. Within optometry, this trend toward using a mobile device to interact with equipment has led to smart-phone controlled acuity charts [Gounder et al., 2014].

As seen in Table 1.1 most mobile devices are equipped with a plethora of sensors. It would make sense to harness these capabilities to enhance clinical testing within optometry. For example, in creating a near vision test that requires patients to be positioned at a certain distance, the front facing camera could track the distance to the patient ensuring that the requirements of the tests are met. Making use of microphones and voice technology could allow another method of user input for patients unable to see or use touch screens. Optometry relies on visual diagnosis for many of the most common eye related conditions, so the opportunities to adapt the imaging systems of existing mobile devices to suit those for eye examinations are vast.

1.1.2.1 Hardware Limitations

While the benefits of mobile devices are numerous there are also limitations to their clinical use. The first being their size. Whilst this gives many of the positive attributes such as portability, it also reduces the size of the testing screen. = A distance acuity app would not be large enough to have many of the large testing letters all on the screen at the same time. Testing one letter at a time may actually increase the time duration of tests and mask crowding . Crowding occurs when the recognition of objects such as a letter becomes impaired by the presence of neighbouring objects. In a standard acuity chart crowding occurs due to letters placed above, below and to the side of the letter that is to be read [Bailey and Lovie, 1980].

Additionally, most hardware in mobile devices is fixed and not customized to suit the task at hand. Using mobile devices for their sensors, particularly the camera, even with additional accessories, limits the design and hence the results from such instruments. Whilst this may be a suitable compromise if no tools currently exist, say in the third world, customized hardware are more likely to produce better results.

1.1.3 Software Development

One of the key driving factors for the large growth in smart-phone ownership has been not only their capability as a phone, or even internet browser, but also the extended capabilities on offer via third party apps [Jayaraman, 2013]. It has been within the interests of the owners of the two leading platforms of mobile devices (Apples iOS and Google's Android) to further their development environments (software used to create apps) and offer a number of enhanced application programmable interface (API's) and software tools to ease the burden on developers.



Figure 1.2: User interface builder [taken from Apple, 2014]

For medical software this can mean reduced costs of development and a increased focus on the functionality of their apps. For example the user interface builder tool available through iOS, Figure 1.2, means several user interfaces can be created and optimized within a short space of time. Simplified high level communication API's also allow the iPad to control medical hardware with relative ease. Medical app creators need not be concentrating on the implementation but on the strength of their respective ideas. Additionally, digital platforms typically called "App stores", that are used to distribute software to mobile devices, have lowered the barrier

to entry in the software industry [Lee et al., 2010]. Without these tools much of this thesis, and arguably many of the recent medical apps available, would not be possible.

1.2 Mobile computing in optometry

There are several specific categories into which medical apps fall:

- Medical reference, calculators and education
- Patient records
- Remote health, consultation and monitoring
- Assessment tools

This list is generally accepted for optometry and ophthalmology [Zvornicanin et al., 2014, Bastawrous et al., 2012a], however, uniquely to these professions, much of the testing requires a variety of targets to be presented and/or is critically dependant on visual diagnosis [Chakrabarti and Perera, 2014]. High resolution screens in mobile devices provide an opportunity to render targets and the on-board cameras can be used to input visual data. Hence a broader range of assessment tools are available than would be expected of healthcare in general. For this thesis, focused as it is on improving testing in optometry, apps concerning student education, medical reference and, to a certain extent, those directly targeted at patients, will be ignored. This thesis will cover the peer reviewed apps and hardware accessories that can be used as assessment tools.

1.2.1 Imaging

1.2.1.1 Slit lamp attachments and the anterior eye

Several smart-phone slit lamp attachments [Chan et al., 2014, Ye et al., 2014, Chakrabarti, 2012] have been used in the literature to support the digitization of images in optometry. They normally consist of a plastic attachment that fixes on to a slit lamp, that enables the smart-phone camera to image through the eyepiece; thus it provides a quick, cheap and convenient form of capturing images within the clinic.

Digitization of images can be important, especially when monitoring the progression and treatment of ocular diseases. Smart-phones, with their plethora of communication technologies, can easily disseminate the photos to a patient record system or to a further specialist. Attachment designers have also stated the cost-benefits of this approach in comparison to buying new video slit-lamps [Ye et al., 2014]. However, the reliance on the slit lamp, a large and often bulky piece of equipment, does not help in creating more mobile imaging devices. To increase portability, Myung et al. [2014b] proposed combining a macro lens, custom light source and a smart-phone to create a hand held anterior eye imaging device.

While authors have been quick to point out the advantages of using these attachments, no quantitative studies comparing the capability of using these specific devices to those of the traditional methods can be found.

1.2.1.2 Fundus Imaging

Fundus imaging is an essential part of an optometrist's tool kit to monitor and diagnose the progression of diseases such as ARMD or diabetic retinopathy. Traditionally this has relied upon large, desk bound fundus cameras to capture and store retinal images. The size of these instruments restricts retinal imaging to the optometry practice or hospital. Their cost also prohibits their use in developing countries. Smart-phones have been touted as a possible solution to overcome the portability and cost issues associated with fundus imaging. Conveniently many smart-phones come with a coaxial light source as well as a suitable camera as standard.

An approach to fundus imaging using a smart phone camera was described by Haddock et al. [2013]. . The camera's on-board flash and a hand-held condensing lens were used to image the retina. In the study only qualitative descriptions of the high quality of the images and the ease of this technique were mentioned. A near identical approach was taken by Lin et al. [2014] who used an identical setup for the fundus imaging of six infants. They reported that using the smart-phone and condensing lens setup was a cost effective approach to imaging the fundus where no equipment was available. However the quality of the images were not equivalent to those taken by professional equipment. Myung et al. [2014a] created an iPhone attachment that held the lens at a suitable distance and provided a custom light source to further enhance images. A more complicated attachment was designed by Maamari et al. [2014]. Rather than attempt an off axis approach to illumination that would involve

the use of the inbuilt flash, they attempted onaxis illumination using a beam splitter and a series of polarizers to reduce specular reflection from the cornea.



Figure 1.3: Retinal imaging attachment taken from [Maamari et al., 2014].

Another approach taken by previous researchers has been to combine the smart-phone with existing specialised equipment for fundus imaging. Using custom equipment improves the quality of the fundus images whilst also making use of the extra utility of smart-phones for their computing and networking capacity. Prasanna [2013] used a smart-phone to enhance and process images produced by a hand-held ophthalmoscope. They found that their image processing techniques had a sensitivity of 85% in detecting retinal abnormalities. Thus rather than just try and imitate existing techniques they extended the capability of fundus imaging techniques using a smart-phone.

What is apparent is the lack of validation for many of the aforementioned imaging techniques. No studies could be found besides Prasanna [2013] who looked at the effectiveness of these imaging devices for tele-ophthalmology or within clinical practice. A probable cause for this is the recent invention of such instruments. It may take time for large studies to be conducted to prove the effectiveness of smart-phone imaging within optometry.

1.2.2 Refraction

An innovative use of a smart-phone as a platform to calculate refractive errors was examined by Bastawrous et al. [2012b]. In their study, they used a smart phone attachment called NETRA (Near eye tool for refractive assessment) was used, which consisted of a micro-lens array placed over a smart-phone display. The patient places their eye on the front of the device and the image formed on the retina indicated the refractive properties of the eye to be tested. Using the touch screen, patients interact with the image and try to closely match a specified result. For example, they may be asked to manipulate an image until two separate lines presented on the screen meet. Once finished, the adjusted image can be analysed to determine a patients refractive error. Spherical refractive error taken with NETRA on 40 subjects correlated well with refractive measures taken by trained optometrists but agreement was poorer with a mean difference of 0.24D and 95% CI of 1.6 D between the optomterist and the device.

1.2.3 Test charts, functional vision and physophysics

While imaging the eye has been one of the most popular uses of mobile devices, a second but equally important use has been the sudden influx of a number of charts and tests for functional vision that that have been traditionally paper-based. Not only is screen resolution sufficient to render randomised targets at a wide variety of sizes than current computerised charts but mobile devices also offer an opportunity for patients to interact with the tests they are receiving. As many of the tests or charts don't require additional hardware, they are a cost effective and simple way of evaluating patients visual function where the access to printed test charts is limited, as would be the case in home visits, schools and emergency rooms [Zvornicanin et al., 2014].

1.2.3.1 Distance Acuity

Black et al. [2013] tested an iPad acuity app called Visual acuity XL (Kybervision, Montreal, Canada), that presented a digital version of the ETDRS acuity charts, against those of a standard computerised test system and a paper based ETDRS chart. They found in the 56 patient study, that the iPad consistently measured worse acuity by an average of 0.18 logMAR than all of the other tests when no action was

taken to reduce glare. A second testing session was carried out in which the iPad was placed upright, an anti-glare cover was attached and the device was moved to reduce the reflection of light sources. No significant difference was found between charts in normal clinical testing conditions so long as the patient had a distance acuity of less than 1.0 logMAR. A similar result was found by Zhang et al. [2013] when an acuity chart app, Eye Chart Pro [Dok LLC, SanDiego, US], composed of tumbling E letters was used on patients, and Gounder et al. [2014] who used a Snellen chart based app. Both studies took care to limit glare reflection and they found no significant differences testing on the iPad than those of paper based charts. Another study by Toner et al. [2014], specifically looking at distance acuity in children, used an app that presents a series of hand gestures that the children are required to identify and found no difference in acuity found with the app and a paper based chart.

1.2.3.2 Near vision tests

Previously, computer chart systems such as the popular, Test Chart 2000 were restricted to testing only at distance due to the resolution limitations of computer monitors. Evaluating near functional vision mostly involved the use of paper based charts or cards that provide no interaction with the patient. With mobile devices having increased screen resolutions and touch screens the opportunity has arisen for new computerised near vision tests.

Yu [2014] tested an interactive near vision acuity app, SightBook (DigiSight Technologies Inc, Portola Valley, US) , displayed on a smart-phone on normal patients and those with cataract, ARMD and diabetes. The app presented a letter target at 30 cm, which patients were required to identify by tapping the corresponding button, located underneath the main target, from six alternative choices. The order and size of targets during the tests followed either the standard visual acuity testing procedure [National Academy of Sciences, 1980] or a custom staircase adaptive algorithm. In accordance with the results from distance acuity apps, the near app and adaptive testing procedure had no effect on acuity measurements when compared with the results of paper based near acuity charts. On the weight of evidence, all visual acuity testing on mobile devices is clinically acceptable and they offer advantages over paper and current computerised testing charts.

Wang et al. [2013] investigated the validity of a shape discrimination task performed on an iPhone, to monitor the performance of ARMD patients. Patients were presen-

ted with three different shapes and asked to select the shape that appeared distorted by tapping on the screen. They found that the shape discrimination app could differentiate between patients with early and advance stage ARMD at least as effectively as distance visual acuity. Furthermore, the shape discrimination task was less sensitive to viewing distance and contrast changes than those of near and distance acuity [Wilkinson et al., 1998]. So it would be ideal to use for home monitoring. A further study by Kaiser et al. [2013], that looked at the practicality of this app for home monitoring in 160 patients, found 86% of patients complied with daily self testing for shape discrimination thus proving that patients are willing and capable of home monitoring.

1.2.3.3 Psychophysical tests

Psychophysical testing methods have traditionally been hard to implement within clinical optometry. Testing methods such as the method of limits, adjustment and two-alternative forced choice procedures require custom stimuli to be presented in response to patient feedback. Paper based charts are often limited in the number of stimuli they can present on one chart and require the examiner to manually implement and respond to the feedback from the patient [Maudgal et al., 1988]. Hence, clinical measurements of psychophysical thresholds can only be assessed in broad discrete steps. For example, contrast sensitivity testing on the VCTS paper chart (see subsection 3.1.3 on page 93) evaluated five spatial frequencies at nine values of contrast . PC based psychophysical tests have remained a research tool rather than a clinical test owing to the large size, expense, and testing time required to perform such tests [Owsley, 2003, Dorr et al., 2013]. Thus there has been a push for psychophysical testing on mobile devices to mitigate some of these problems [Dorr et al., 2013, Kollbaum and Jansen, 2014, Turpin et al., 2014].

Dorr et al. [2013] tested the reliability and accuracy of an iPad based contrast sensitivity app against a standard psychophysics laboratory setup. For each trial, patients were presented with two targets, one of which was blank and another which contained a contrast grating, and were asked to identify the grating. An adaptive two alternate forced choice procedure (2-AFC) called the qCSF [Hou et al., 2010] predicted the possible threshold of patients in response to the feedback received and rendered new contrast targets on the iPad appropriately. Dorr and colleagues found that the results from the iPad app were indistinguishable from those taken on a laboratory system, and the novel implementation of an adaptive procedure reduced

testing times to ten minutes. Meanwhile, a psychophysical platform for researchers to implement 2-AFC methods on tablet computers was created by Turpin et al. [2014]. Psychophysical stimuli, either static or moving can be uploaded from the web and presented on the iPad in the form of a 2-AFC procedure without the need for complicated coding. Kollbaum and Jansen [2014] created a Pelli-Robson contrast letter chart for the iPad where letter optotypes were presented to patients at decreasing contrast. Measures taken with the app were as repeatable as the paper based Pelli-Robson chart, but agreement between the two was poor. The app gave significantly higher thresholds. Moreover the app merely rendered fixed targets and offered no interaction with either the examiner or patient.

1.2.4 Discussion and Opportunities

Attachments for smart-phones that allow the imaging of the fundus and the anterior eye have become increasingly popular due their cost and size advantages. Nevertheless, they do not offer the imaging capability currently available from custom made equipment like fundus cameras [Lin et al., 2014]. Currently, only anecdotal evidence exists about their imaging capabilities and hence their utility in clinical optometry. That being said using the smart-phone as a communication and or image processing device to be coupled to existing equipment like slitlamp attachments and ophthalmoscopes shows great promise. Already automated detection of retinal abnormalities has taken place with this approach [Prasanna, 2013].

Innovative use of smart-phone attachments need not solely rely on imaging to enhance testing. The near refractive tool [Bastawrous et al., 2012b] was a good example of combining a low cost adapter with the use of the interactive capabilities of touch screens to gain accurate measures of refraction. With communication devices such as blue-tooth and wifi available, external sensors can even be used to extend the current capabilities of smart-phones.

While current mobile apps have proven to be both repeatable and reliable when used for measurement of distance acuity, they offer no clear advantages besides portability, to traditional computer charts when used in a clinical setting. They may even be a hindrance due to the small screen sizes restricting the amount of letters that can be displayed on the screen for the larger acuities. Testing one letter at a time may increase the time required to perform the test [Toner et al., 2014], and overestimate visual acuity by reducing crowding [Flom et al., 1963, Morad et al., 1999]. Their real advantages become apparent only in situations where traditional

paper or computerised charts are unavailable or inconvenient; say in emergency waiting rooms or domiciliary visits. That being said, future apps that remotely control distance acuity charts, presenting a way of soliciting patient feedback for automated scoring of visual acuity, may reduce testing times and thus impact clinical assessment of acuity.

Mobile devices with their high resolution screens can display near vision tests that would otherwise have been strictly paper based, hence they offer the greatest opportunity for mobile devices to enhance clinical testing. Both Yu [2014] and Wang et al. [2013] offered near functional vision apps that took full advantage of a smartphone touch screens to provide interaction with patients. Testing methods could be undertaken automatically, saving the examiner effort and time of implementing these manually, a feature that has also been adopted by the physophysical apps [Dorr et al., 2013, Turpin et al., 2014]. Computerised testing of near acuity [Yu, 2014] and contrast [Dorr et al., 2013] has made use of adaptive algorithms, whose complexity could never be attempted manually, to further reduce examination times. If mobile technology can be used to reduce testing times, functional vision tests that were typically the reserve of research can find utility within clinical testing.

1.2.4.1 Reading speed

One functional vision test that would benefit from being placed on a mobile device, for which an app is not currently available, is the assessment of reading speed. Presently, examiners wishing to assess reading speed at natural reading distances must make use of paper charts through a time consuming testing method. Examiners are required to unveil sentences one by one to patients, record time duration using a stopwatch as the patient reads the sentences aloud, and make note of reading errors. A tablet-based app could automate both the presentation of sentences and the stopwatch functionality in order to reduce both testing time and errors associated with reading speed assessment.

1.2.4.2 Contrast sensitivity

Whilst contrast testing on computer tablets is not novel, current systems have not used their plethora of sensors and touch screens to reduce testing times enough to be widely adopted in practice. Even with adaptive testing methods, Dorr et al.

[2013] mentioned that it would take around 10 minutes (120 trials) to plot a patient's contrast sensitivity curve. A tablet based app that does not require discrete presentations of stimuli, abandoning the lengthy 2-AFC procedure, may have clinical use within optometry, and should be investigated.

1.2.4.3 Amplitude of Accommodation

A common form of assessing the amplitude of accommodation is the subjective push up test. In this method a target, usually consisting of fixed letters, is placed at a far distance to the patient (typically over 50cm) on a sliding rule that is rested upon a patient's cheek. The target is then slowly brought closer to the eyes by the examiner until blur is reported by the patient. Creation of a new attachment to a smart-phone, that can measure distance accurately, will allow the subjective push up test to be performed in free space without the need for an invasive rule. Additionally, with the high resolution screens on offer, targets that dynamically resize with distance keeping the visual angle constant, may enhance the clinical testing of amplitude of accommodation.

1.3 Image processing

It has often been quoted that both optometry and ophthalmology rely heavily on visual diagnosis for detection of a range of conditions [Patton et al., 2006, Chakrabarti and Perera, 2014]. With the falling cost and size of computer hardware, and the ever increasing power of their processors, automated analysis techniques that can quantify ophthalmic images have become commonplace in instruments such as Placido disc topographers. Most image processing tools analyse static images, but the opportunity to track physical changes as they occur within the eye, such as those for accommodation, tear film, pupil response, vein dilation, saccades, blinking, fixation and alignment can help quantify existing subjective measures or improve the accuracy of instrumentation that may be affected by those changes. As hand-held mobile imaging devices become more common the application of dynamic image processing is going to get broader.

1.3.1 Image processing theory

Digital ophthalmic images are often distorted by noise from various sources that affect the quality of an acquired image. Noise is caused by a combination of the imaging system used and the interference of undesirable actions such as the micro-movement of eyes, gaze change, head and blinking [Ben Abdallah et al., 2013, Alonso-Caneiro et al., 2011, Buehren et al., 2001]. Furthermore, details in images are often too small for examiners to analyse or dynamic changes too quick to be studied manually. Image processing can be used to mitigate these phenomena and enhance clinical observations.

For dynamic imaging there are three main classes of algorithm: enhancement, segmentation, and quantification. Not all image processing requires or uses all these classes. For example, image processing can be used to provide a clearer image for examiners, where the details in raw images may be difficult to perceive, but quantification may be left to subjective judgement. That being said, computer vision applications, that is the automated analysis of images free of subjective observations, often use a combination of these three types of algorithms to arrive at an optimal solution.

1.3.1.1 Enhancement

Algorithms for enhancement aim at reducing image noise and increase the prominence of certain structures (signals) within the image. Enhancement methods (Figure 1.4) tend to accentuate and sharpen image features such as edge boundaries, increase the dynamic range of images and filter out high frequency noise. This frequently occurs as a preprocessing step before any segmentation or quantification takes place. In corneal topographers the main structures to be observed are the Purkinje rings surrounded by extraneous iris and scleral data. Adjusting the contrast and or the gamma value of images will make the boundary between the rings and the iris easier to see, Figure 1.5. In ophthalmic computer vision applications, where automated algorithms may have to deal with different lighting conditions or different eyes, enhancement techniques may reduce between subject variability of images ready for further processing.

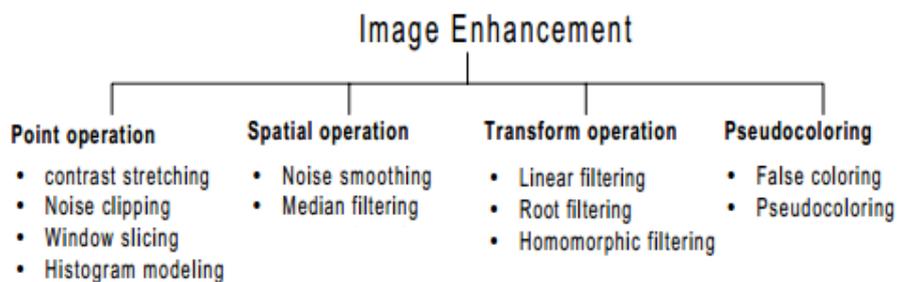


Figure 1.4: Diagram of the most commonly used image enhancement techniques

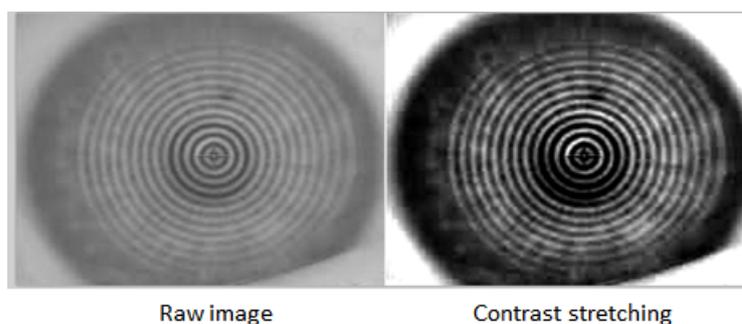


Figure 1.5: Enhancement of raw Placido topography images

1.3.1.2 Segmentation

Delineating structures of interest from background data is the main purpose of segmentation. Human observers find this an easy thing to do (to recognise a car against a road or a bird against sky is an effortless task) but in machine vision this segmentation can be challenging. Complete segmentation in one step often can't be done except in very simple images. Partial segmentation that feeds into higher level processing is typically used.

A simple example of this type of segmentation can be found in pupillometers, which use a thresholding technique (pixels larger than a certain value are differentiated from those below) to first abstract the pupil region from the rest of the eye, Figure 1.6. Post processing to remove the reflected lights, which in this case involves quantifying each separate particle (red circles in the image) and removing those that are too small, completes the segmentation and leaves the shape of the pupil intact. The most common segmentation procedures are outlined in Figure 1.7.



Figure 1.6: Segmentation in pupillometry images

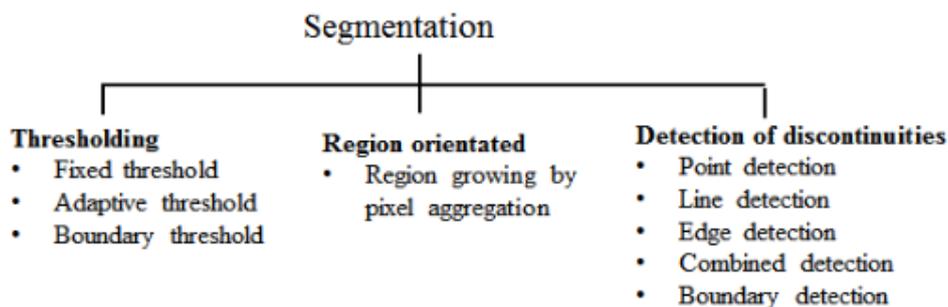


Figure 1.7: Segmentation in pupillometry images

1.3.1.3 Quantification

Quantification algorithms aim to provide a diagnostic measurement from segmented images. In machine vision applications, the last image processing step is image quantification and so it disproportionately relies on the segmentation techniques that precede it. For example, an important parameter considered in the progression of glaucoma is the cup to disc ratio. For automated assessment of this metric, fundus images must first be segmented to separate the disc and cup of the optic nerve head, Figure 1.8. Quantification tools like the mass moment can be used to find the centre of a segmented image, if the border of the segmented image is known, the distance between the centre and the border can be found. Taking the mean of the distances over all parts of the border can give an estimation of the radius of the cup and disk. Common quantification algorithms normally measure parameters such as shape, texture, orientation and motion.

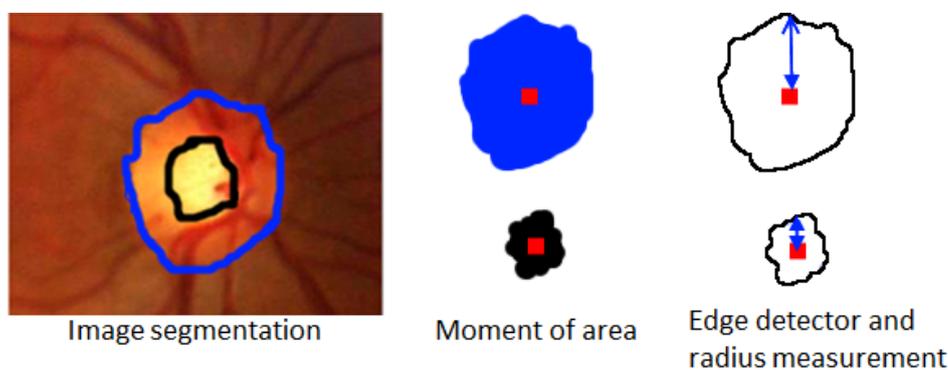


Figure 1.8: Quantification of retinal images, disc and cup ratios

1.3.2 Dynamic imaging

Taking subjective dynamic measurements using ophthalmic imaging equipment can present a problem to examiners. Often changes across time are brief, or the amplitude of the change is small and easily missed. Even recording a video and manually assessing certain attributes is a long and laborious process. Quantitative measures of dynamic changes in the eye therefore can get overlooked in a clinical setting. Computer vision systems can automate the process increasing accuracy of dynamic measurements. This section is going to look briefly at the existing uses of computer vision regarding dynamic measurement in ophthalmic imaging.

1.3.2.1 Current applications of computer vision in dynamic measurement

Spontaneous retinal vein pulsations are small changes in retinal vein thickness close to the optic disc and can often be seen when viewing the retina. Both cameras, and confocal laser ophthalmoscopes can be used to record video images of the fundus containing vein pulsations. Seifert and Vilser [2002] used a flickering light to provoke vein pulsation responses. An automated image processing algorithm segmented veins and tracked their widths from images obtained on a commercial fundus camera. Moret et al. [2011] created an image processing system that rejected erroneous data caused by eye micro-movements, segregated the dynamic and static components between the image frames and tracked the amplitude of retinal vessels in real time on a confocal laser ophthalmoscope.

Assessing pupillary light reflexes is used in clinical practice as an indicator of neurological function. Subjective measurements require estimating the size and time of pupil responses across both eyes. Pupillometers can be used to track the dynamic changes in pupil size objectively and give accurate readings for pupil lag and shape. The image processing systems used have to mitigate the interference from both ocular and eyelid movement while assessing the pupil diameter. Automated pupillometers have been shown to be more repeatable in measurements of the pupillary light reflex than examiners themselves [Meeker et al., 2005]. More over, accurate quantitative measures of the pupil centre and shape have become important in Laser-Assisted in situ Keratomileusis (LASIK) refractive surgery, where pupillometry measurements can determine the suitability of surgery for patients [Rosen et al., 2002].

Subjective assessment of oculomotor function is part of a routine eye examination but it disproportionately relies on qualitative observations [Elliott, 2013a]. Automated eye tracking can be a useful tool to give high spatial and temporal resolution measurements of oculomotor functioning to help differentiate diseases such as acquired and infantile nystagmus, which require recording the waveform characteristics of eye movement [Dell'Osso and Daroff, 1999].

1.3.2.2 Opportunities-Tear film

A growing field of automated dynamic measurements look specifically at the tear film. When the tear film is spread evenly it takes the approximate shape of the cornea, however degradation (instability leading to eventual rupture) of the tear film causes exposure of the corneal surface. Abnormal or fast degradation of the

tear film can be used as an indicator of dry eye disease [Abelson et al., 2002]. Currently, the most popular non-invasive methods of tear film stability assessment are predominantly subjective, offer poor agreement with many other tear measures and generally lack repeatability [Cedarstaff and Tomlinson, 1983, Nichols, 2004]. Automated imaging systems for tear film stability assessment have thus been developed using video keratoscopy but the validity and repeatability of many of the procedures have also been mixed [Goto et al., 2003, Pult and Riede-Pult, 2011, Best et al., 2012, Tyagi et al., 2012]. Thus, research into new image processing methods and computer vision systems to aid in the objective measurement of tear stability may be useful.

1.4 Aim

Having discussed the opportunities that mobile devices and dynamic image processing systems can bring to clinical testing, the aim of the study presented in this thesis was to demonstrate the potential of both dynamic image processing and mobile computing technology to enhance clinical measurement in optometric practice by exploring the following two questions:

1. What are the advantages of clinical testing on mobile devices in comparison to existing measures?
2. How can a new computer vision system improve the objective measurement of tear stability?

These questions will be answered through the development of several novel mobile apps, the creation of customised hardware, ray tracing simulations, and a new computer vision system to demonstrate the potential of different elements of technological advancement (Figure 1.9). Working distance monitoring with the in-built camera, voice detection and computerized scoring are demonstrated through the development of a reading speed app (chapter 2). A method to improve the dynamic range of the screen and touch technology to replace visual psychophysical testing are showcased through the development of a contrast sensitivity app test (chapter 3). Additional hardware interfacing with iPhone software allows free-space amplitude of accommodation assessment (chapter 4). Finally, dynamic image processing is conceptualized (chapter 5) and orchestrated (chapter 6) to improve the objective assessment of tear film stability.

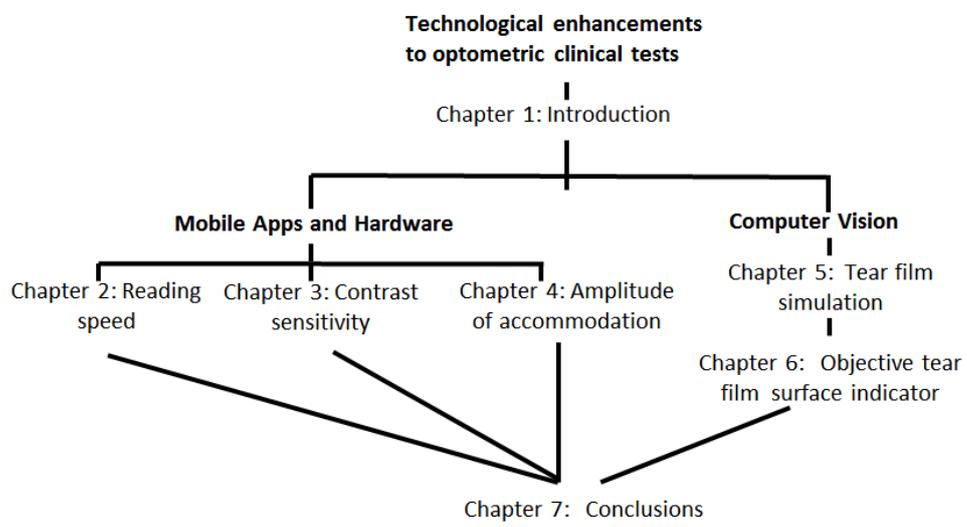


Figure 1.9: Thesis structure

2 The Development and Validation of an Objective Reading Speed Test.

2.1 Introduction

Reading is one of the most vital and common skill's for engaging, communicating and interpreting ideas. Any visual loss that affects reading ability will have a disproportionate impact on a patient's quality of life and is often cited as a major factor in patients seeking professional help [Elliott et al., 1997] for eye related problems. Reading performance (speed) is not standard practice in optometry Current paper based reading performance charts [Mansfield et al., 1993, Radner and Radner, 2008], are often cumbersome to use when manual time measurement, sentence unveiling and error recording have to be undertaken simultaneously by the examiner. Additionally reading performance metrics are determined by plotting reading performance data graphically, a laborious and time consuming process that would be better suited to automation. Furthermore, any metrics determined subjectively by the examiner can be variable; data around the critical print size (CPS) is particularly noisy and is likely to be misjudged [Cheung et al., 2008]. Computerised reading systems have been used for reading speed assessment [Mansfield et al., 1993], but display technology was not sufficiently advanced to be able to test a wide range of print sizes at typical reading distances. Moreover they were often coupled with bulky hardware unsuitable for use in the clinic.

A quick and efficient reading test based on a mobile computing platform may be a viable alternative to current PC or paper charts. They can combine the utility of a computerised test with the portability of a paper based chart. High resolution displays in products such as mobile phones and tablets can render very small text sizes perfectly at standard reading distances. Moreover, due to the compact size and long battery life, mobile devices can be readily used at the examination chair. Therefore this chapter charts the development and validation of a mobile app based

reading speed test.

2.1.1 Visual impairment and reading speed

Full functional impairment of vision cannot be assessed by visual acuity alone [Riusala et al., 2003, Martin et al., 2001]. Visual acuity (VA) scores for subjects do not represent sufficiently a patient's satisfaction with visual performance [Mangione and Phillips, 1994, Riusala et al., 2003]. Patients require not only ample visual acuity to identify letters but also sufficient visual span to recognise words and determine eye movements [Ahn and Ledge, 1995]. Hence the need for clinical tests that can assess a broader spectrum of visual impairment and thus correlate accurately with patient visual satisfaction. Reading performance has been shown to have a significant impact on self-reported satisfaction with vision [Hazel et al., 2000, Javitt et al., 1993] in patients suffering with macular disease or cataracts as well as increased performance, being the primary rehabilitation goal by the majority of elderly low-vision patients [Kleen and Levoy, 1981].

Assessing reading performance in clinical examinations may improve discrimination between eye conditions. Reading tests may also be able to improve clinical diagnosis of conditions if the effect on reading performance metrics between different diseases are known. For example, low-vision patients with central vision loss have been shown to have severely reduced reading speeds in comparison to low vision patients with cloudy ocular media, even when visual acuities are comparable [Legge et al., 1985b]. Many of the key diseases and disorders that have previously been studied with respect to reading speed are shown in Table 2.1.

Reading speed is closely related to macular function. Consequently the majority of the research on reading and visual impairment has concerned itself with differentiating macular conditions from other conditions such as cataracts. Several studies in Table 2.1 have shown optimal reading speed to be a good differentiator between ARMD and cataract patients. A cataract only study has shown that optimal reading speed was greater in patients with nuclear cortical cataracts than for those with posterior sub-capsular cataracts [Stifter et al., 2004a]. Additionally the RA/VA ratio was much smaller in those with posterior sub-capsular cataracts, an indication that VA measurements may be underestimating the impact of visual impairments in patients with certain conditions.

Of particular interest are two studies, firstly where optimal reading speed could dis-

criminate between dry (drusen maculopathy) and wet (sub-retinal fibrosis) ARMD subjects (see Table 2.1). Secondly, a study found that reading speed correlated with lesion size in those with cystoid macular edema (see Table 2.1). These two studies suggest that reading metrics may well have use as a clinical tool for studying the progression of macular diseases in patients.

Table 2.1: Visual impairment and reading speed studies

Impairment Studied and parameters measured	Details- [Abbreviations: ARMD (age related macular degeneration), CPS (critical print size), VA (visual acuity), ORS (optimal reading speed), RA (reading acuity) and MRS (mean reading speed)]	Reference
Cataract and ARMD (CPS, VA, ORS)	Optimal reading speed- remained constant for control and cataract (99.1 and 106.1 wpm respectively). ARMD significant lower (39.3 wpm). Critical print size showed differences between the three groups, control 0.43, cataract 0.88 and ARMD 1.06. Optimal reading speed unaffected by cataract but severely reduced by ARMD. As expected CPS and VA affected by both cataract and ARMD.	[Elliott et al., 2001]
Cataract, ARMD, Peripheral vision loss. (ORS)	Optimal reading speed read slower for ARMD patients (27 wpm). No significant difference between healthy (103.8 wpm), cataract (91.9 wpm) and peripheral vision loss (91.5 wpm) at large 1.20 logMAR text sizes.	[Pesudovs et al., 2002]
Pure nuclear, nuclear cortical, dense nuclear and posterior sub-capsular cataracts. (CPS, VA, ORS, RA, RA/VA ratio)	Subjects with pure nuclear or nuclear-cortical cataracts showed no significant reduction in optimal reading speed pre and post operatively for large text sizes. However for dense nuclear and posterior sub-capsular cataracts optimal reading speed increased post operatively. RA/VA ratio (reading threshold / visual acuity threshold) of more than 90% were seen for pure nuclear and nuclear cortical cataracts whilst for posterior sub-capsular ratios of 66-78% were observed. Hence clinically significant visual reading impairments may be underestimated by VA measurements alone.	[Stifter et al., 2004b]
Drusen maculopathy, subretinal fibrosis after choroidal neovascularisation. (VA, RA, ORS, MRS, CPS)	Patients VA was comparable between the two groups, wet and dry ARMD. RA was comparable for both groups when VA scores were between 0.2-0.4 logMAR but drusen maculopathy patients showed statistically lower reading acuity when VA was between 0.4-0.7 logMAR. ORS, MRS and CPS were significantly lower in patients with sub-retinal fibrosis, thus reading tests can be used to discriminate between conditions.	[Radner et al., 2004]
Uveitis associated cystoid macular edema (VA, CPS, MRS, ORS, RA/VA ratio)	RA and ORS were considerably more impaired than VA for those with cystoid macular edema (CME). RA and CPS correlated better with lesion size ($r=0.63$ and $r=0.61$ respectively) than VA with lesion size $r=0.49$. Thus reading speed assessment may be a quick indicator for monitoring progression of disease.	[Kiss et al., 2006]

2.1.2 Reading presentation methods

Evaluating reading performance in patients requires repeatable, reliable and reproducible testing methods. Currently several methods exist for presenting text to the patient in order to record reading speeds: rapid serial visualisation presentation (RSVP), drifting text, and static text.

2.1.2.1 Static Text

Static text methods present a fixed amount of text either on paper, flash card or computer screen. The patient is asked to read the text aloud in full while the examiner records the time and errors from when the first word is spoken to the last. Time measurement is typically done on a stopwatch for paper based charts, or by the computer in computer based charts. In paper based charts the text is presented at various sizes on the same page. Typically an examiner will cover the whole chart and will either uncover the words/sentences/paragraphs for a particular text size, or let the patient uncover it themselves. Static methods are quicker to administer than RSVP (rapid serial visualisation presentation) and drifting text as only the text size is varied, the speed at which the text is read is down to the patient [Luebker and Lamay, 1989]. Optimal reading speeds for sentence based static text methods are around 200wpm for normally sighted patients [Stifter et al., 2004a, 2005].

2.1.2.2 Rapid serial visual presentation (RSVP)

The RSVP method presents text at the centre of a computer screen, word by word. Using a computer, the size of the text and the time between presented words can be varied on the display. Differing both presentation speeds and text sizes requires substantial time. Recording the speed threshold at which a patient can read fluently for differing print sizes produces a standard reading speed curve.

RSVP presentation negates the oculomotor functioning involved in performing static and drifting text. Whilst this does not change the overall shape of the reading speed curve, optimal reading speeds (ORS) are significantly better [Rubin and Turano, 1992, Fine and Peli, 1995]. Öquist and Goldstein [2003] found that RSVP presentation increased reading speeds by 33% in comparison to static text methods with no corresponding drop in comprehension. However, if comprehension is untested normally sighted patients can read as fast as 1000 wpm [Chung et al., 1998].

2.1.2.3 Drifting text

The drifting text method presents a line of text that moves across the screen at a designated speed. Drifting speeds are increased until reading becomes severely impeded, and the drift speed recorded. Obtaining the reading thresholds for a range of text sizes produces the reading curve. As noted previously with RSVP, this method not only takes considerable time but also the reading speed data may only be obtained in discrete steps. Eye movement patterns between drifting and static text also appear different, although the retinal image processed in both cases are strikingly similar [Legge et al., 1997]. It has been found that low vision subjects can read 15% faster than static text whereas normal subjects read static text faster than drifting text [Luebker and Lamay, 1989].

2.1.3 Reading Texts

To ensure that differing reading speeds are attributable to visual defects in vision rather than intellectual ability, reading tests require linguistic factors to be controlled. For this purpose many different reading texts have been developed. Standardized reading texts have been created using full-length paragraphs, single sentences and randomly assigned words with the hope of isolating reading performance to visual function. The main linguistic criteria for reading speed text are as follows: passage length, word frequency, syntactic complexity and content.

2.1.3.1 Length

Passage length refers to the amount of characters words or syllables that a passage of text is said to own. As reading speed is typically defined as words per minute, many of the popular texts such as MNREAD [Mansfield et al., 1993] and the Radner Reading Test [Radner et al., 2002] have text passages of equal word length. Hahn et al. [2006] argued that the number of characters is a more appropriate measure of passage length, since words can vary significantly in length. Thus the time needed to read and repeat words can differ considerably. While this may be true of single sentence reading texts, for those with large paragraphs the effect of small changes in character length will be diminished.

2.1.3.2 Word frequency

Word frequency refers to the frequency in which a word is said to appear in a language's written texts. Experiments have suggested that the speed and accuracy in identifying words are directly related to this word frequency metric [Savin, 1963]. Unsurprisingly, more regular words can be identified quicker than those which are seldom used [Rayner and Duffy, 1986]. This word frequency effect can be controlled in reading speed texts by setting a frequency limit on included words. For example, only words with a lexical frequency of 0.0005% or higher can be included.

2.1.3.3 Syntactic complexity

Identification of words themselves are not the only linguistic factor that affects reading speed. The complexity of the sentence structure within a text (syntactic complexity) can alter the speed and accuracy at which text is read. An example of where syntactic complexity is explicitly used is in children's books and websites. Writers will typically create shorter sentences with fewer clauses to reduce the intellectual burden on younger readers. Different approaches to controlling the syntactical complexity of text between single sentence and paragraph reading texts exists. The Radner reading test sentences all contain relative clauses [Radner et al., 2002]. Relative clauses are parts of sentences that contain a subject, verb, relative pronoun and an adjective . The syntactical structure remains constant whilst the lexical units (words) change for each sentence. However for routine paragraph reading texts this is not suitable; paragraphs must generally contain sentences of varying structures to be legible [Trauzettel-Klosinski et al., 2012]. Therefore a paragraph template, with pre-determined sets of differing sentence structures, may be used to control the syntactic complexity of long text passages.

2.1.3.4 Content

The content of any text presented in reading speed tests must also be controlled. If the lexical difficulty of words vary across the reading texts, reading performance may fluctuate irrespective of their visual function. For this reason many of the texts set arbitrary rules for the content. They may create a lexicon of words that can be used, typically corresponding to a certain reading age or proficiency. Long winding and complex prose is substituted for shorter simpler sentences. Scientific terms and niche topics are avoided.

2.1.4 Reading speed curve and metrics

Although the magnitude of the reading speed data can differ across all three testing methods (RSVP, static and drift), they all share a distinctive three-phase reading speed curve. For a broad range of intermediate text sizes, reading speed remains fairly constant. A sharp decrease from this plateau in reading speed occurs for small characters until reading becomes impossible. Meanwhile, for large characters a more gradual decline is observed. Any drop off in reading speed before the cut-off suggests that fluent reading is replaced by spot or survival reading.

Clinicians have used reading performance primarily to assess real world impairment in low vision patients. As the majority of reading takes place at the small and intermediate text sizes very little interest is given to large text sizes. Many of the clinical reading speed tests reflect this by omitting very large character sizes and produce reading curves such as in Figure 2.1.

Several metrics have been used to quantify reading performance data in order to differentiate between patients and conditions. Critical print size (CPS), optimal reading speed (ORS), mean reading speed (MRS), and reading acuity (RA) have all been used in the literature.

2.1.4.1 Optimal and mean reading speed

Optimal reading speed (ORS) is the maximum reading speed that can be maintained where print size is not a limiting factor. Visual inspection of the plateau region of the graph (as seen in Figure 2.1) can determine ORS subjectively. Another alternative method is to take the median of the reading speed for the three largest text sizes presented on the chart [Rice et al., 2005]. Mean reading speed (MRS) of all print sizes has been used intermittently as a metric. Reading speed is normally expressed as wpm (words/min) or cpm (characters/min).

2.1.4.2 Critical print size

Critical print size is defined as the smallest print size at which a patient can maintain optimal reading speed. The sudden drop in reading speed on the reading speed curve can be an indicator of CPS for examiners, see Figure 2.1. However data around the CPS can be noisy and difficult to gauge subjectively so objective methods exist to determine CPS. Asymptotic curves have been fitted to reading speed data, with the



Figure 2.1: Reading rate versus print sizes taken from a clinical reading adapted from [Legge et al., 1985a]

curve approaching a maximum value akin to the ORS. Where the curve reaches a value of 80 to 90% of the magnitude of the asymptote, a print size value is recorded as the CPS [Cheung et al., 2008].

2.1.4.3 Reading acuity

Reading acuity (RA) is the smallest print size than can be read by a patient and is analogous to VA in letter identification. Simply the lower cut-off value on the graph in Figure 2.1 determines the RA of a patient. Much like in VA scoring, any mistakes that are made in reading can be added to the RA value. RA is always a smaller value than CPS, any measured difference between them is called the acuity reserve [Whittaker and Lovie-Kitchin, 1993].

2.1.5 Current charts, texts and devices

2.1.5.1 Minnesota near reading chart

The MNREAD test was originally designed to be presented on a computer screen using the drifting text method [Legge et al., 1989]. Examiners soon realized the test was difficult to administer and created a static reading chart [Mansfield et al.,

1993]. The chart consisted of a series of standardised test sentences printed on the page. Each sentence decreased in text size starting from 1.3 to -0.6 LogMAR. All of the sentences were declarative statements and consisted of sixty Times New Roman characters positioned over three lines. Individual word length and syntactic structure was allowed to differ between sentences. In order to control lexical difficulty, the vocabulary was selected from word lists that contain the 1000 most frequent words in children's literature (10 year old level).

As with all paper based charts, determining reading speed metrics must be carried out manually, a laborious and time consuming process. Working distance's are generally measured once at the start of the test using a ruler, string or similar measuring device [Ahn et al., 1995].

2.1.5.2 Radner reading Chart.

The Radner reading test has many similarities with the earlier MNREAD test. Primarily both are sentence based static reading charts, whose text decreases in size down the chart. Moreover, both charts took consideration of word frequency and intellectual ability when choosing their respective vocabularies. Nevertheless in the Radner reading charts the sentences have been standardised further to ensure that syllables, nouns, and verbs are positioned across each sentence consistently [Radner et al., 2002]. As a consequence each different sentence follows a set pattern, termed a "sentence optotype". In the case of the Radner reading chart a bound, non-restrictive, relative clause is always used. It has been claimed that the variability of reading speed across the Radner reading sentences is less than those of MNREAD [Maaijwee et al., 2008].

2.1.5.3 International reading speed texts

It had been noted that standardised single sentence reading charts (MNREAD and RR) are well suited for assessing critical print size. Paragraphs however may be preferable for accurate speed measurement as the percentage error in reading time measurement is smaller for longer texts. The international reading speed texts (IREST [Trauzettel-Klosinski et al., 2012]) presents standardised paragraphs for reading speed assessment. Paragraphs were adapted from an encyclopaedia for children; word frequency and lexical difficulty were controlled. Each paragraph was

moderated for syntactical complexity using a metric of structural complexity theorized by Gibson's syntactic prediction theory [Gibson, 1998, 2000].

2.1.5.4 Salzburg Reading Desk

In light of the many challenges with assessing reading speed on paper based charts, namely those of patient distance, illumination and metric determination, Dextl et al. [2010] created the Salzburg reading desk. It consisted of a table mounted reading desk, controlled illumination source and a laptop computer that was equipped with two infra-red cameras to measure reading distance. Microphone recordings of each patient could be taken and displayed visually to the examiner in order to determine reading durations.

When the desk was used in conjunction with the Radner reading chart it proved to be a valid and reliable method of assessing both reading speed and patient distance [Dextl et al., 2010]. However, its suitability in a clinical setting is questionable. Distance measurement, using video-stereo-photometry, required the patient's nose to be marked with a small green dot, an inconvenient and invasive procedure. The size of the desk severely limits its portability and will require ample space within any clinical examination room.

2.1.6 Outcomes

Presently, reading texts have been used for assessing the effects of refractive correction or vision aids on low vision patients. Nevertheless, actual reading speed tests that assess reading performance over multiple text sizes are infrequently used in practice. This is despite recent studies having suggested that reading speed data can be used to discriminate or track disease in patients with certain pathological conditions [Kiss et al., 2006, Radner et al., 2004].

Researchers have responded to the need to measure reading speed performance by developing specialised reading speed texts [Trauzettel-Klosinski et al., 2012, Radner et al., 2002, Luebker and Lamay, 1989]. These generally try to control linguistic factors such as, lexical difficulty, text length and syntactic structure in order to produce several sentences or paragraphs that can be used in reading speed assessment. While it can be argued that longer paragraphs will reduce measurement error [Rubin, 2013], single sentence texts have proven to be reliable and require less testing time [Stifter et al., 2004a, Patel et al., 2011].

Both rapid serial visual presentation and drifting text have been used as presentation methods. Whilst they have proven to be effective [Fine and Peli, 1995], the methods themselves require both text size and presentation/drift speed to be varied. As such they are time consuming and thus impractical for anything other than research. Static testing methods are a more practical method of assessing reading speed within a clinical setting. By only varying text size, static testing methods require less time than those of RSVP and drifting text.

Computerised reading systems have been used for reading speed assessment [Mansfield et al., 1993], but display technology was not sufficiently advanced to be able to test a wide range of print sizes at typical reading distances. PC based charts also required ample room and could not be readily administered on the examination chair. Meanwhile, paper based static charts can be administered to patients on the examination chair and require only the chart itself and a stopwatch. This ease of use, portability, and time efficiency have spurred the development of several clinical paper based static reading charts (MNREAD, Radner Reading Test and IREST).

Currently, paper based static reading charts are the most appropriate for measuring reading performance in practice. They are not without their drawbacks however. Reading speed metric's are best determined by plotting reading performance data graphically, a laborious and time consuming process that would be better suited to a computer. Any metrics determined by the examiner can also be highly subjective; data around the CPS is particularly noisy and is likely to be misjudged [Cheung et al., 2008]. Manual time measurement that makes use of stopwatches is prone to measurement error [Trauzettel-Klosinski et al., 2012]. Implementation of the test is also cumbersome. The examiner is expected to handle a stopwatch, record time measurement, spot reading errors and supervise the patient and this makes the test impractical. Illumination can not be controlled easily.

Despite their physical size and low resolution screens, computerised reading tests still have many advantages over their paper rivals. Graphing data instantly allows examiners or even the computer itself to determine reading speed metrics quickly and objectively. Time measurement and voice recording become trivial on computer based charts [Dexl et al., 2010]. Furthermore, backlit displays on computer monitors can enable greater control over illumination conditions.

A quick and efficient reading test based on a mobile computing platform may be a viable alternative to current PC or paper charts. They can combine the utility of a computerised test with the portability of a paper based chart. High resolution

displays in products such as mobile phones and tablets can render very small text sizes perfectly at reading distances. Moreover, due to the compact size and long battery life, mobile devices can be readily used on the examination chair. Touch technology embedded in the screens can provide an intuitive and user friendly way for a patient or examiner to interact with a reading performance test. The increased processing power of mobile computing devices, as well as their plethora of inbuilt sensors, can enable reading metrics to be determined automatically. Therefore a smart-phone/tablet based reading chart could increase adoption of reading performance as a standard clinical test.

2.2 Aims and objectives

2.2.1 Aim

Current static reading speed tests take too long to complete and analyse, as well as being cumbersome for examiners and patients alike. Computer based reading systems have traditionally been too bulky, complicated and time consuming to use for anything other than research. The lack of an efficient and convenient reading speed assessment has hindered reading speed performance as a clinical measure of functional vision. A reading chart based on a mobile device will be able to offer consistent illumination, more accurate reading speed metrics and versatility than current paper based MNREAD and Radner Reading (RR) charts. The aim of this project therefore, was to develop a reading performance test on a mobile device that is better suited to a clinical setting than current paper and PC based charts.

2.2.2 Objectives

1. Selection of a suitable mobile device and computing platform.
2. Evaluate apt test chart designs.
3. Develop a suitable program to control the new mobile reading test.
4. Conduct a study to understand how a computer controlled mobile reading test compares to traditional static based paper charts.

2.3 Development-Aston Read

2.3.1 Mobile device and Software development kit

2.3.1.1 Criteria

Selection of an appropriate mobile device is critical to the development of an electronic test chart. The technical criteria for device selection were as follows:

Screen resolution Computer screen's capable of displaying a full reading test, must be able to present small characters at natural reading distances (40 cm). At this distance it would be ideal to render letters at -0.2 logMAR acuity (as Radner Reading and MNREAD charts do). The majority of patients in a low vision clinic however will not resolve to those levels at near distance so a minimum of 0.0 logMAR is acceptable. Using both Equation 2.1 and Equation 2.3, where s_{target} is the size of the target in cm, $D_{reading}$ is the working distance and P_{render} is the pixel height of a letter, ideal resolution would be 136 Pixels/CM and 86 Pixels/CM to reach a reference size of -0.2 and 0.0 logMAR at 40 cm.

$$S_{target} = D_{reading} * \tan(5/60) * 10^{LogMAR} \quad (2.1)$$

$$0.037cm = 40cm * \tan(5/60) * 10^{-0.2} \quad (2.2)$$

$$R_{screen} = \frac{P_{render}}{S_{target}} \quad (2.3)$$

$$136\text{pixels}/\text{cm} = \frac{5\text{pixels}}{0.037\text{cm}} \quad (2.4)$$

Touchscreen user input In order to record time accurately, any delay in user input should be kept to a minimum. Complicated controls and poorly designed user

interfaces are detrimental to both reaction times and usability. Fortunately many phones and tablets come equipped with touch sensitive screens. Touchscreens are already sufficiently common that many examiners and patients will be familiar with their use. A lack of external hardware also increases the portability of devices.

Portability There are two main aspects to portability, size and battery life. The device itself should be small enough to be held comfortably in the hand, much like a book or newspaper. Yet it has to be large enough for text to be rendered at large sizes; if the layout of sentences is not constant across all text sizes reading performance may be affected. Secondly the device must be able to last for extended periods of time in clinical use, so long battery life is a premium. A standby time of 9 hours, combined with a full working time of 3 hours would be more than sufficient to cover a clinician's day.

Software In addition to the physical requirement of the mobile device it is important to note the importance of any software development tools. Choosing a system with an integrated development environment, sufficient APIs (Application programming interfaces), and appropriate programming language can greatly reduce development time and increase the effectiveness of any reading performance test. Of the most popular smart phones and tablets, only two major operating systems exist; Apple's IOS and Google's Android. Both operating systems provide similar software development kits (SDK's) for app development. It can be noted however that the iOS SDK comes bundled with a more powerful interface builder and debug tools, all of which can be used to create a better user interface and reduce development time.

Internal sensors It is not only physical contact that is useful for user input. Additional sensors like microphones, cameras and accelerometers can enhance reading speed assessment. For example, a microphone can be used to record the patient, making it easier for an examiner to spot mistakes post-test. Inbuilt camera and accelerometers may also be beneficial for positioning the patient at a suitable reading distance. However, it is key that these sensors be inbuilt to the device. Any external sensors are detrimental to both portability and complexity.

2.3.1.2 Device selection: Apple iPad

Using the criteria mentioned in 3.1.1 the Apple iPad was chosen as the device to host the reading test. The relevant specifications and selection criteria are shown in Table 2.2. Size and portability requirements precluded the use of the majority of smart phones and laptops. E-readers with electronic ink displays were considered due to their sufficiently long battery life, but none could be found with suitable software development tools. Tablet computers were deemed the best choice for displaying a reading test. Most tablets offer a high resolution display, sufficient computing power and a plethora of sensors. Their size lends well to being used within a clinic. Typically no bigger than the size of an A4 page, they can be handled easily by examiners and patients alike, whilst still being able to render text at both large and small sizes. The Apple iPad tablet computer was chosen specifically as it offered the highest resolution display, an integrated development environment and was the most popular tablet computer since 2012.

Table 2.2: Criterion and iPad specification

Criteria	Ideal Value	Acceptable Value	Apple iPad Specification	Remarks
Screen size	H =22cm W =15cm	10cm<H< 30cm 10cm<W<22cm	H=19.71cm W=14.78 cm	Ideal based on standard paperback book size. Min values on largest text size to be rendered. Max no larger than A4 paper size.
Screen resolution	R>136 ppcm	R>86ppcm	R=104ppcm	Ideal on rendering a letter to -0.2 logMAR at 40cm. Acceptable based on rendering at 0.0 logMAR at 40cm
Overall dimensions	H<22cm W<15cm D<1cm	10cm<H< 30cm 10cm<W<22cm D<3cm	H=24cm W=18.7cm D=0.9cm	Ideal, slim paperback book sized. Acceptable, smaller than A4 paper size
Mass	M<0.5kg	M<1kg	M=0.6 kg	Must be comfortable enough for elderly patients to hold for several minutes.
Additional Sensors	Camera, microphone, distance.	microphone	Camera, microphone, gyroscope and accelerometers	Recording patients is a must for an enhanced test. Additional orientation and camera sensors useful for positioning but not essential.
Battery Life	Standby> 48hrs Working> 24 hrs	Standby>10 hrs Working> 4 hrs	Standby=72 hrs Working=10 hrs	Device must be able to get through a working day. Acceptable value standby of a full day with excessive 50% on time during work hours.
Human input	No additional hardware. Touchscreen	No additional hardware. Touchscreen.	Touchscreen, limited voice recognition.	Intuitive user interface demands touchscreen. Accurate voice recognition would be a feature for assessing reading accuracy.
Software development tools	N/A	N/A	IOS SDK, debug tools, user interface tools	Integrated development environment helps reduce development time and increase effectiveness of user interfaces.

2.3.1.3 Screen Calibration

In order to accurately determine the luminance of the device's pixels, screen calibration was carried out using a Minolta Ls-110 luminance meter. The brightness setting was set to its maximum and the grey scale values between 0 and 255 were all evaluated, see Figure 2.2a and Figure 2.2b. Pixel values could now be attributed to a specific luminance value which is critical when rendering contrast gratings.

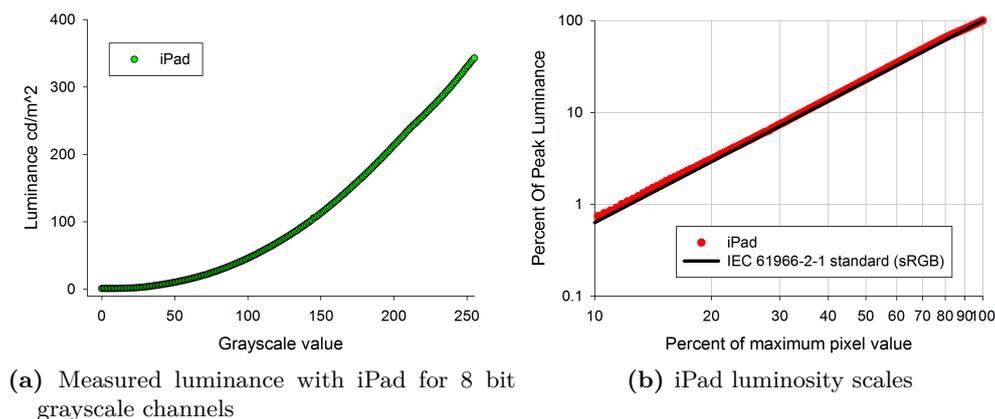


Figure 2.2: iPad screen calibration

2.3.2 Software

The application was designed to incorporate novel features that have not been, before now, featured in a reading test. Using the iPad's high resolution screen, small text sizes could be rendered at near viewing distances. Previous computer based reading test's incorporated low resolution LCD or CRT monitors. This forced examiners to either abandon testing at the lower reading acuities or position patients further away from the reading material [Luebker and Lamay, 1989]. Furthermore the size of the iPad lends itself to being handled much like a book or newspaper and can be used at natural reading distances. This greatly increases the practicality of a computerised chart as it can be administered on the examination chair much like today's paper based charts. Another major advantage of using an iPad is its touch screen; a user interface could be created that would be intuitive for both patient and examiner. The application could not only be designed for timing the patients in this way, relieving the examiner of having to manually operate a stopwatch, but the

patients' voices could be recorded to aid in mistake recognition once the test had been completed. Once reading had finished completed graphs could automatically be generated and displayed to aid the examiner in assessment. Using the back lighting of the iPad's screen to control illumination levels may also be beneficial.

The initial concept of the enhanced reading acuity test can be seen in Figure 2.3. The simplified flow chart for the program can be seen in Figure 2.4. The four distinct stages of the design are described below.

2.3.2.1 Starting view

The initial screen upon opening the application This includes the set-up options for the reading test. Two different sets of sentences can be used on the reading test. This was to ensure that if mistakes were made by either examiner or patient during the test, the patient could be re-examined with a fresh pair of sentences without bias to the results arising from remembering previously read sentences. Any instructions to the examiner are also shown on the starting view.

2.3.2.2 Reading Test

Once the examiner selects the chart and positions the patient at 40cm to the device using a length of string, the reading test shows. The sentences were replicated from the Radner reading chart (as described in subsection 2.1.5.2) as they were known to be reliable sentence optotypes [Maaijwee et al., 2008]. A single Radner sentence is displayed on the screen at any one time as well as two buttons; one allowing the patient to indicate they have read the text and the other for the patient to indicate that the sentence could not be read at all. The first sentence has a large text size of 1.0 logMAR. The computer starts a stop watch and records the patient as text is being read from it's microphone. Once the patient has read the text, the "text read" button is pressed, The stopwatch and recording will stop. The data is saved and another sentence appears at a smaller size (step sizes of 0.1 logMAR). Again the stopwatch and audio recording are started and the patient reads the text aloud. This process continues with each sentence. Once the patient indicates that they can no longer read the sentence, or all of the text sizes have been exhausted, the test finishes.

2.3.2.3 Evaluation

The evaluation screen displays the last whole sentence that was read by the patient. A button adjacent to the presented sentence is used to play the accompanying audio recording. Any words that were misread in the sentence can be pressed, and the associated syllables can be counted in order to more accurately determine reading acuity. Once the examiner is satisfied that all of the mistakes have been accounted for, the results screen shows.

2.3.2.4 Results

The results screen presents a graph and a table of the patient's reading speed based upon print size. This allows the examiner to determine the CPS accurately. Other metrics such as reading acuity can be displayed. Exporting the data to email or a printer can also be accessed off this screen.

2.3 Development-Aston Read

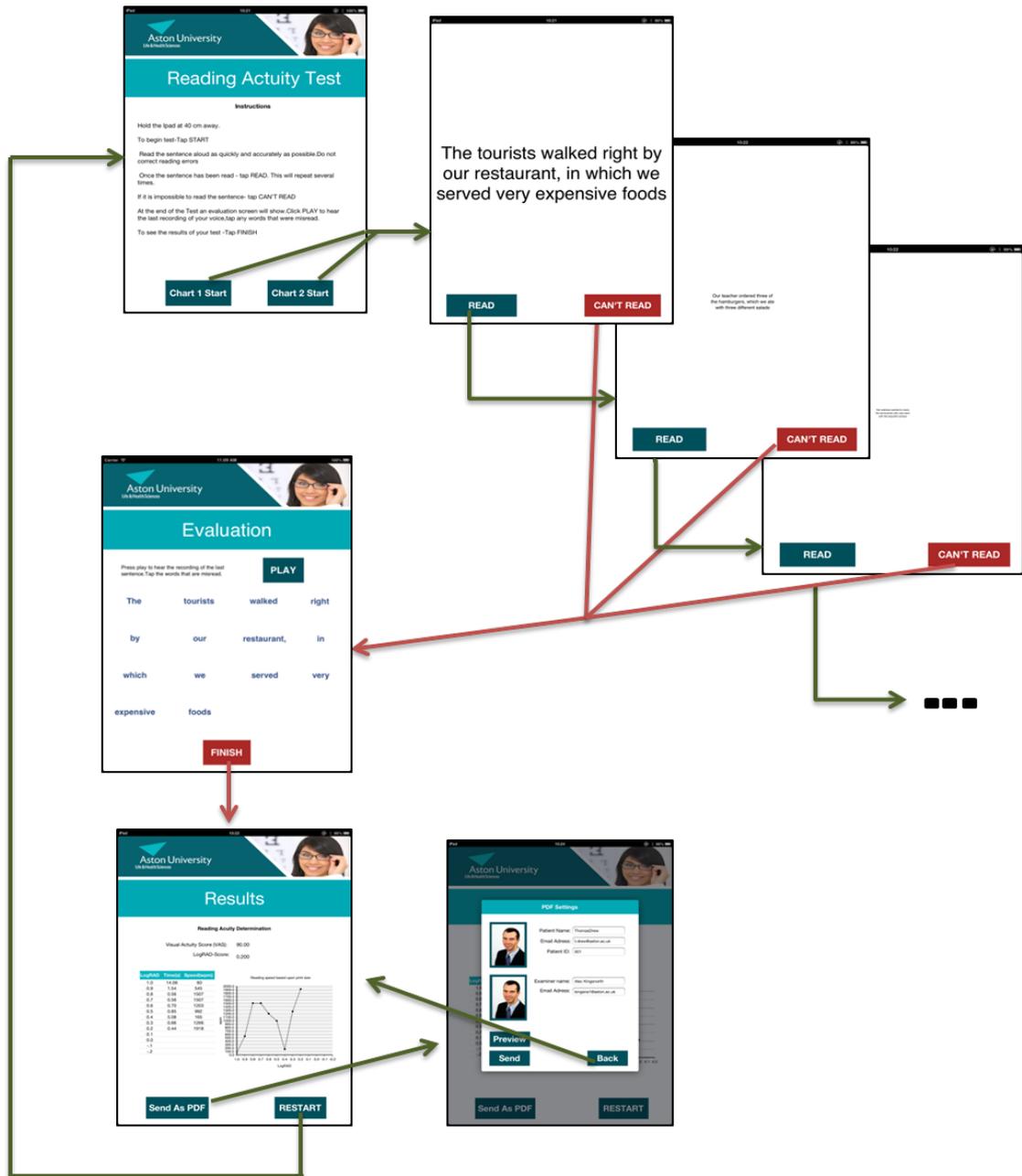


Figure 2.3: Aston Read - User Interface

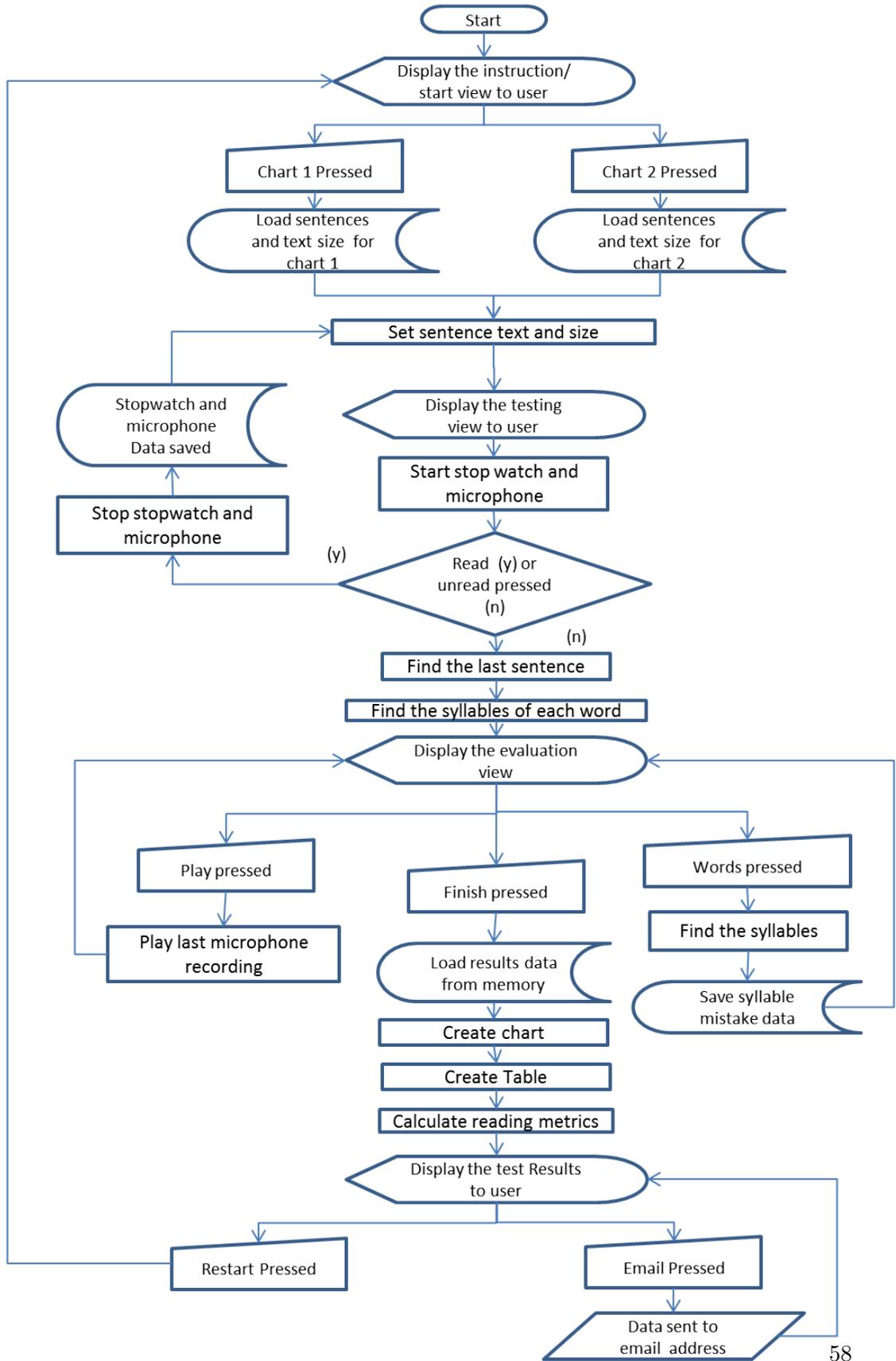


Figure 2.4: Aston Read - Programming flow diagram

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

2.4.1 Aim

The main aim of this study was to compare the performance of the enhanced Aston Read reading acuity test (as described in section 2.3) to those of a standard paper based reading assessment, the Radner Reading test in best corrected subjects with no prior pathology. Additionally intra-test repeatability of measurements made with the enhanced reading acuity test were examined.

2.4.2 Subjects

This experimental study comprised of 21 young-adult subjects with an age range of between 20-30 years old. 85% of the subjects were below 25 years of age and all were recruited from Aston University's student body or staff. Study approval came from the university ethics committee and all of the tenets from the declaration of Helsinki were followed. Subjects were required to speak English fluently, have a best corrected distance visual acuity <0.2 logMAR and to have no current ocular disease.

2.4.3 Protocol

Volunteers were asked to read 12 sentences of decreasing print size with the paper based chart (Radner reading), and two versions of the iPad based chart (Aston Read). Each chart was positioned at 40cm away from the subject, and the order in which they were taken was randomized. All the test sentences were selected from the Radner reading chart, and consisted of 14 words inside a relative clause as described in subsection 2.1.5.2. The time taken for a patient to undergo a full reading speed test was recorded using a manual timer.

Initially the Radner reading chart was covered with a sheet of paper. The patients were asked to uncover each sentence and read aloud as quickly and accurately as possible. Sentence text size would decrease as the test continued. The first sentence had an x-height text size of 1.0 logMAR and the following sentences decreased in step sizes of 0.1 logMAR to a minimum of -0.1 logMAR. From the unveiling of the

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

sentence to when the subject was first heard to stop, time duration was recorded manually on a stop watch. The test finished either when the whole chart had been completed or if the patient was unable to read a sentence within 30 seconds. Any mistakes in the reading test were noted manually by the examiner.

The Aston Read test was taken twice on the iPad. As described in section 2.3, the test presented the aforementioned Radner reading sentences one at a time in decreasing 0.1 logMAR steps starting at 1.0 logMAR. The computer timed the duration between when the text had appeared and the patient had pressed the “Read” button on the screen. As with the paper based chart, the test finished either when completed or if the patient pressed the “Cannot Read” button. At the end of the test a recording for the last sentence read could be played and mistakes counted by pressing any of the misread words.

2.4.3.1 Sentence presentation

To ensure that any variability between the Radner Reading charts and the Aston Read charts could be attributed to their respective testing methods and not the reading content, only Radner reading sentences were used. These all consisted of 14 words with 20 syllables spread over three lines, as shown in Figure 2.5a. The font used in both charts was Ariel, and the x-height of the text size was determined by the angle subtended within the eye at 40cm for logMAR values between 1.0 and -0.1 see Figure 2.5b. The logMAR formulae for determining x-Height size can be seen in screen resolution on page 50. Background luminance on the iPad based chart was set at 200 mcd (see subsection 2.3.1.3 for screen calibration). Michelson contrast between the white back ground and the black test sentences was 98.95%.

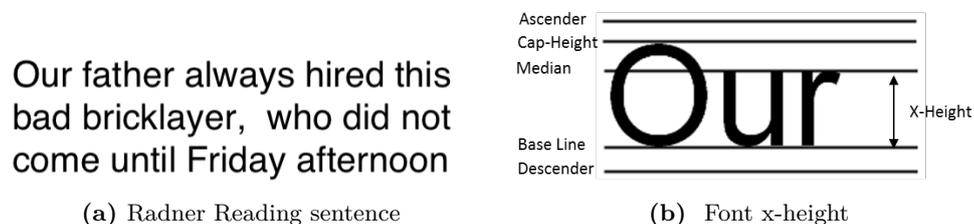


Figure 2.5: Sentence presentation

2.4.3.2 Reading Speed.

Reading speed was calculated using Equation 2.5 [Radner and Radner, 2008] where t is the time duration in seconds and $S_{reading}$ is the reading speed in words per minute.

$$S_{reading} = \frac{14 * 60}{t} \quad (2.5)$$

2.4.3.3 Reading Acuity

Each sentence consisted of 20 syllables and equates to 0.1 logMAR. To calculate reading acuity the logMAR value of the last sentence read is counted, and the number of syllable mistakes accounted for as shown in Equation 2.6 where R_{acuity} is the reading acuity, $LogMAR_{lastLineRead}$ is the reading acuity of the last full line read and $M_{syllables}$ is the number of misread syllables in the sentence.

$$R_{acuity} = LogMAR_{lastLineRead} + (0.005 * M_{syllables}) \quad (2.6)$$

2.4.3.4 Objective CPS

Reading speed measurements were ran through a custom curve fitting program created in MATLAB. [Cheung et al., 2008] found that exponential decay functions better fit reading speed data for objective CPS measurement. Thus an exponential model, as described in Equation 2.7 , with a least squares curve fitting algorithm was used to fit the raw reading speed data. A fitted curve from a single subject can be seen in Figure 2.6. At the larger text sizes the fitted exponential model remains at a constant gradient until the reading speed decreases and the gradient changes rapidly. The CPS point was defined when the fitted curve reaches 90% of its maximum value (see Figure 2.6 where A , B , C , D are the variables to be fit).

$$f(x) = (A * e^{bx}) + (C * e^{dx}) \quad (2.7)$$

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

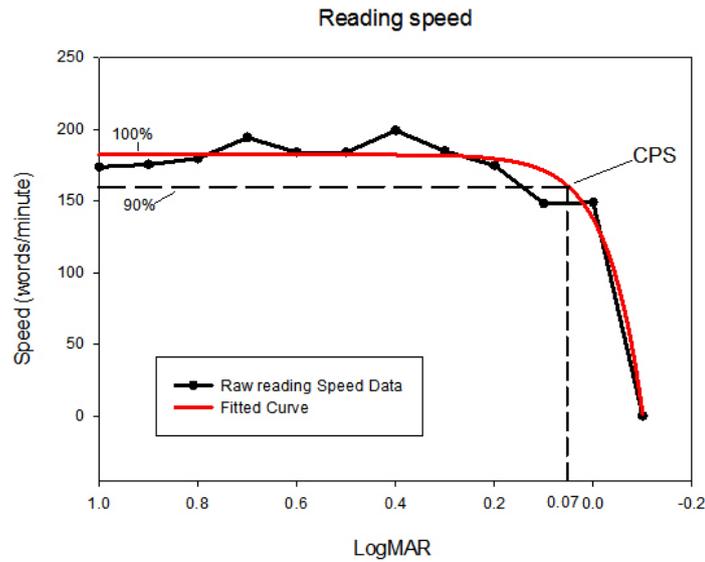


Figure 2.6: Raw reading speed data and objective CPS measurement

2.4.3.5 Optimal reading speed.

Optimal reading speed was calculated as the mean of reading speed values up to and including the subjective critical print size. Mean reading speed for an individual was calculated as the mean of reading speed values up to and including the reading acuity limit.

2.4.4 Statistical testing

Data was checked for normality using the Shapiro-Wilko ($P < 0.05$) test before parametric testing could begin. Repeated measures ANOVA's were conducted to find if there were significant differences between the app charts and the Radner reading chart for all of the reading speed metrics. If statistically significant effects were found, students t-tests between the charts were carried out to see which charts differed; C1 was the comparison between RR and chart 1, C2 between RR and chart 2, and C3 between chart 1 and chart 2 respectively. If non-normal distributions were found equivalent non parametric tests (Friedman and Wilcoxon signed-rank tests) were used.

The Bland-Altman method was used to analyse the agreement between the three reading speed charts for each metric. To determine correlation between the charts,

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

Spearman's or Pearson's correlation coefficients were calculated to assess the linear association between metrics on the two different app charts and the Radner reading chart. Inter-test repeatability of the Aston Read test was determined from the correlation scores for comparison C3.

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

2.4.5 Study results

2.4.5.1 Mean and Individual Reading speed

The mean reading speed values for each chart at each logMAR reading acuity are displayed in Figure 2.7. The mean reading speeds were $156 \pm 35wpm$ for the Radner Reading chart, $184 \pm 42wpm$ for chart 1 and $184 \pm 39wpm$ for chart 2 on the Aston Read chart respectively. Bland Altman plots using all the individual reading speed measurements are shown in Figure 2.8, Figure 2.9 and Figure 2.10. The mean time duration required to complete a full reading speed test were $178 \pm 22s$ for the Radner Reading chart and $107 \pm 13s$ for the Aston Read chart.

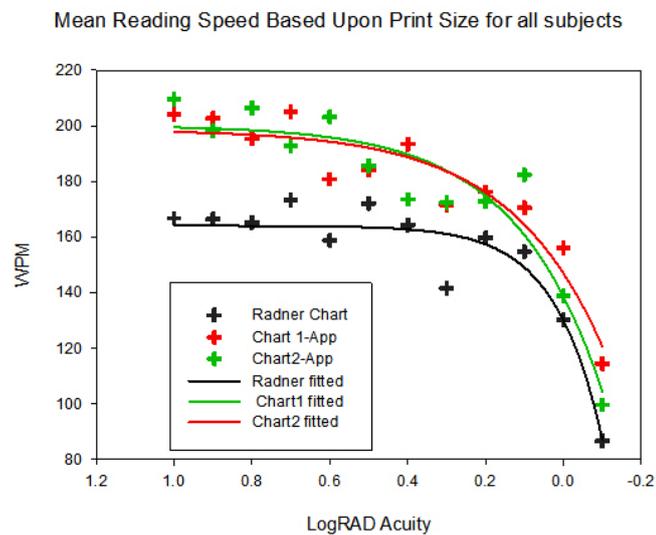


Figure 2.7: Mean reading speed results

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

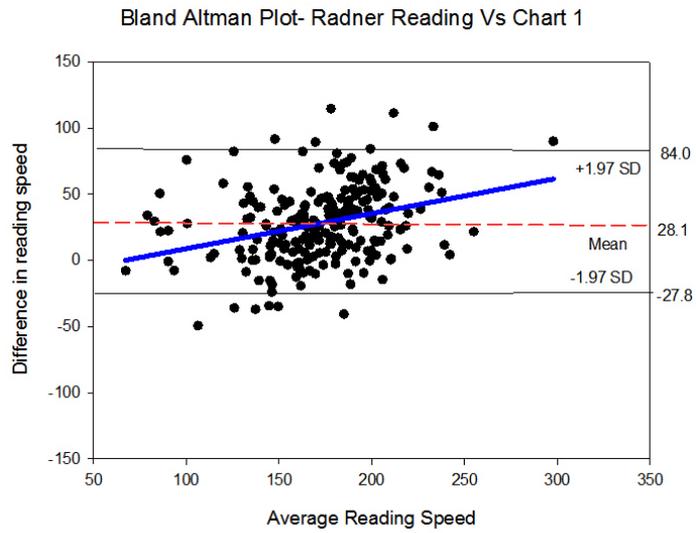


Figure 2.8: Difference plot RR versus Chart 1

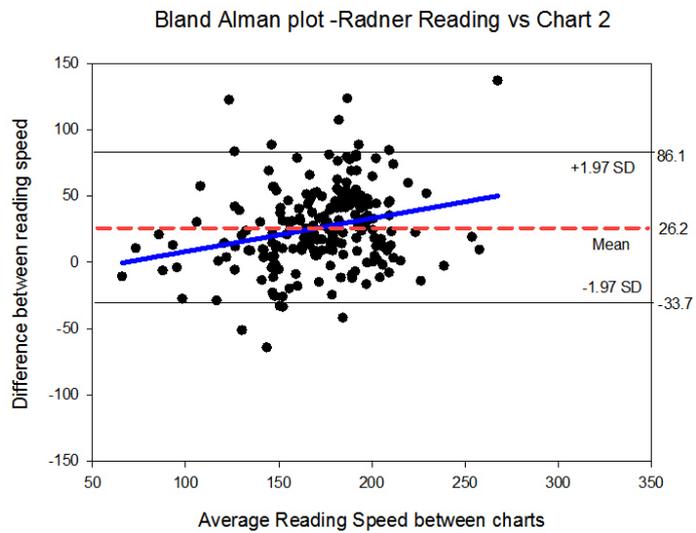


Figure 2.9: Difference plot RR versus Chart 2

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

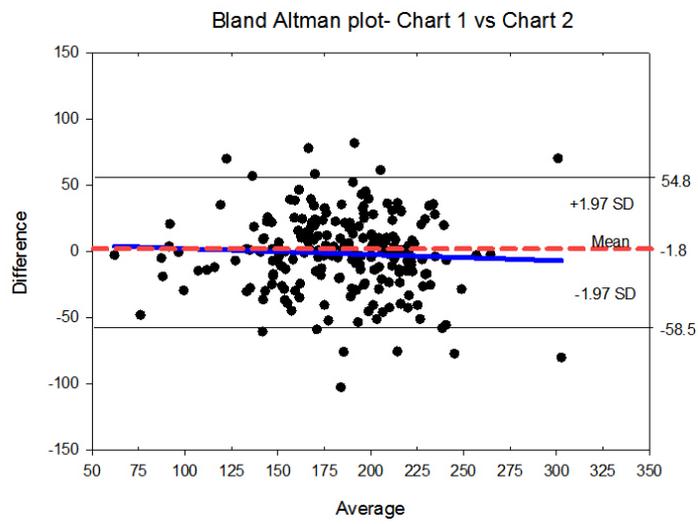


Figure 2.10: Difference plot Chart 1 versus Chart 2

2.4.5.2 Optimal Reading Speed

The mean optimal reading speeds for each chart were 166 ± 20 wpm for the Radner chart, 194 ± 29 wpm and 195 ± 25 wpm for chart 1 and chart 2 of the iPad app respectively. There was a statistically significant effect of chart type on optimal reading speed ($F = 57.000$, $p < 0.01$) with the difference being between the paper and app based charts ($p < 0.01$), but not between app based charts ($p = 0.89$). Pearson's correlation coefficients for the subjects tested on each of the charts can be seen in Table 2.3. Bland Altman plots of the ORS metric can be seen in Figure 2.11, Figure 2.12 and Figure 2.13.

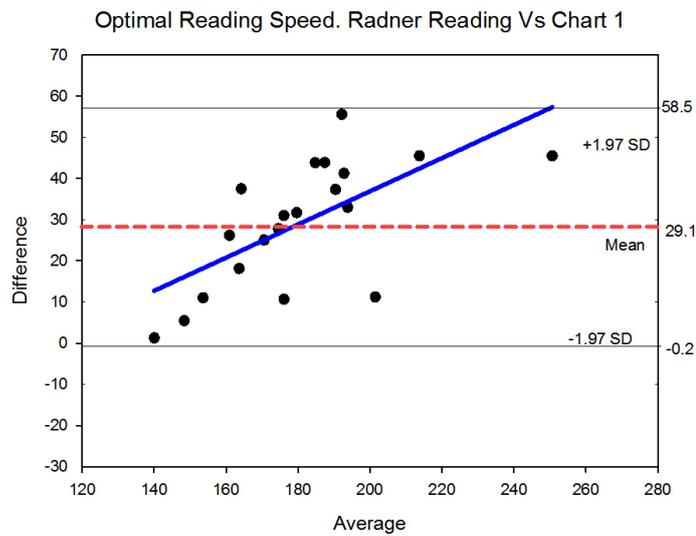


Figure 2.11: Difference plot RR versus Chart 1

Table 2.3: ORS agreement and correlation values

Subjective CPS	C1	C2	C3
Bland Altman			
Difference SD	0.17	0.17	0.08
Difference Mean	-0.08	-0.07	0.01
Pearson			
Correlation -r	0.614	0.534	0.917

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

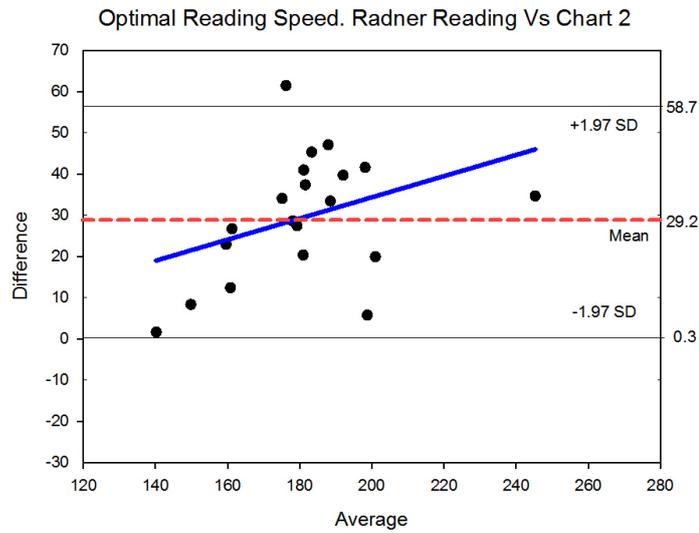


Figure 2.12: Difference plot RR versus Chart 2

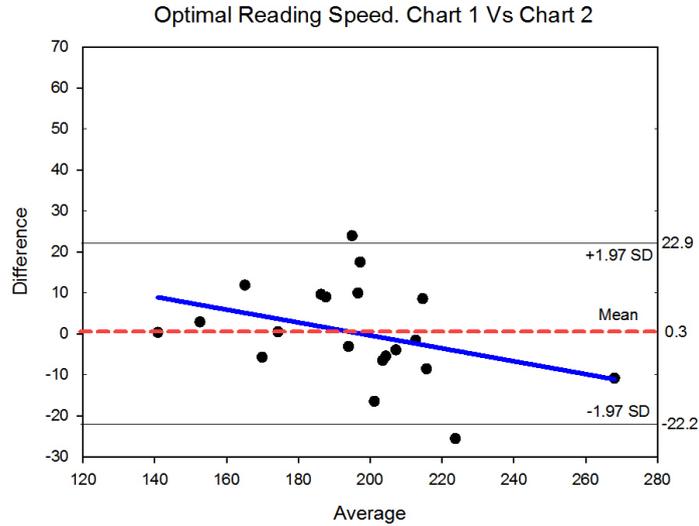


Figure 2.13: Difference plot Chart 1 versus Chart 2

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

2.4.5.3 Reading Acuity

The mean reading acuity thresholds for the subjects were $0.018 \pm 0.13 \log MAR$ for the Radner Reading chart, $0.007 \pm 0.10 \log MAR$ and $0.008 \pm 0.11 \log MAR$ for chart 1 and chart 2 of the iPad app respectively. Reading acuity data was found to violate the normality assumption so non-parametric testing was carried out. There was no statistically significant difference between charts for reading acuity ($\chi^2(2) = 0.76$, $p=0.68$). Spearman's correlation coefficients for the reading acuity can be seen in Table 2.4. Bland-Altman plots for the RA metric can be seen in Figure 2.14, Figure 2.15 and Figure 2.16.

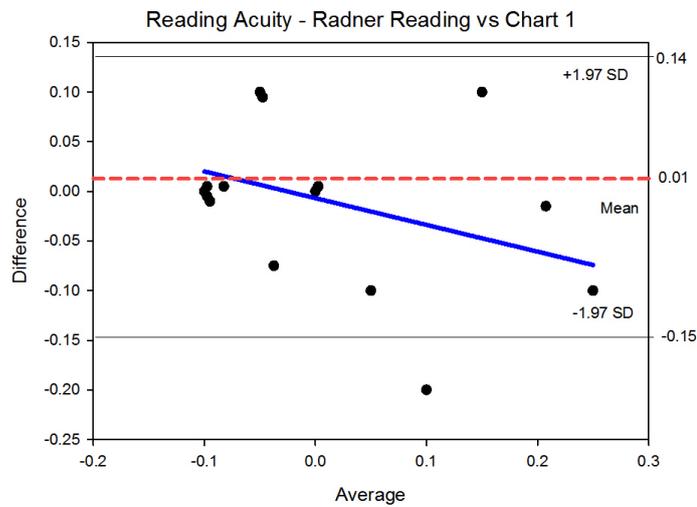


Figure 2.14: Difference plot RR versus Chart 1

Table 2.4: RA agreement and correlation values

Subjective CPS	C1	C2	C3
Bland Altman			
Difference SD	0.075	0.077	0.006
Difference Mean	-0.010	-0.010	-0.013
Spearman			
Correlation -rho	0.706	0.719	0.936

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

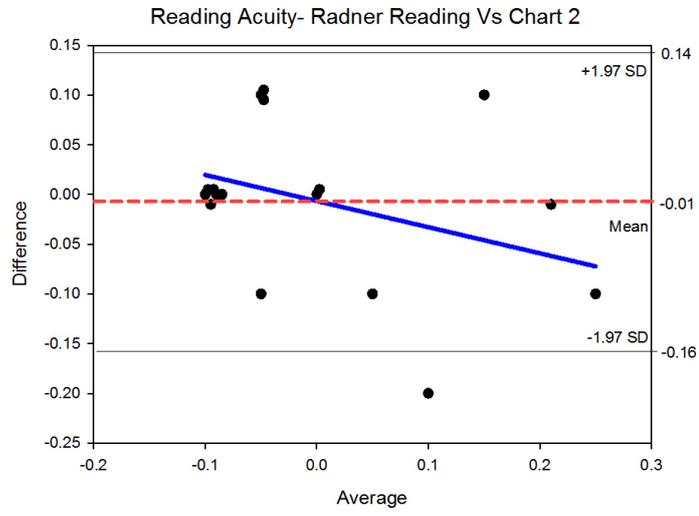


Figure 2.15: Difference plot RR versus Chart 2

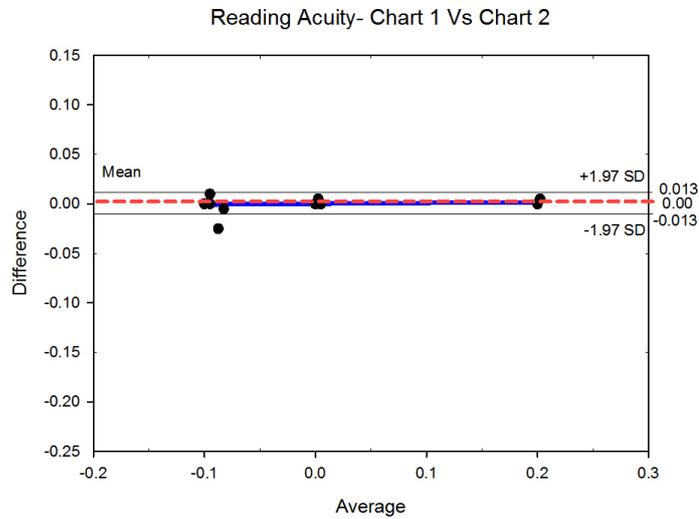


Figure 2.16: Difference plot Chart 1 versus Chart 2

2.4.5.4 Objective Critical Print size

The mean objective critical print sizes for the subjects were $0.27 \pm 0.25 \log MAR$ for the Radner Reading chart, $0.14 \pm 0.17 \log MAR$ and $0.18 \pm 0.19 \log MAR$ for chart 1 and chart 2 of the iPad app respectively. There was a statistically significant effect of chart type on objective CPS ($F=5.406$ $p<0.01$) with the difference being between the paper and app based charts ($p<0.01$, $p<0.01$), but not between app based charts ($p=0.89$). Pearson's correlation coefficients for the critical print size can be seen in Table 2.5. Bland-Altman plots for the objective CPS metric can be seen in Figure 2.17, Figure 2.18 and Figure 2.19.

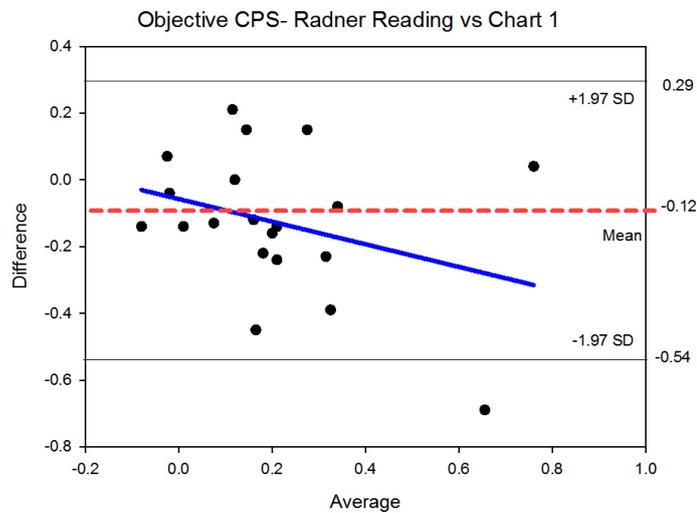


Figure 2.17: Difference plot RR versus Chart 1

Table 2.5: Objective CPS agreement and correlation values

Subjective CPS	C1	C2	C3
Bland Altman			
Difference SD	0.21	0.17	0.12
Difference Mean	-0.12	-0.08	0.04
Pearson			
Correlation -r	0.597	0.749	0.769

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

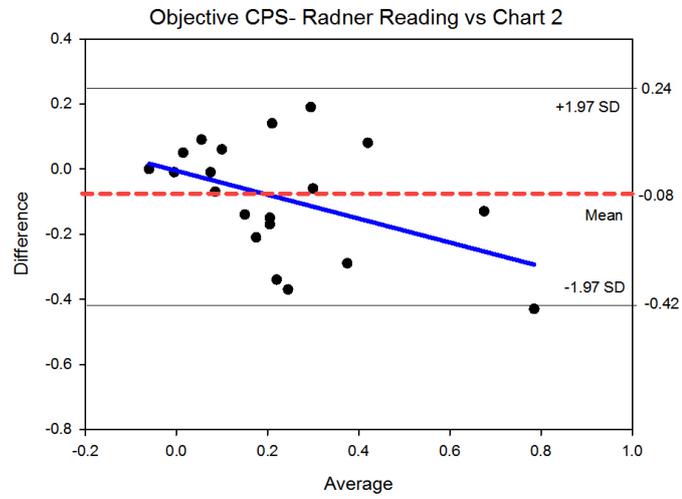


Figure 2.18: Difference plot RR versus Chart 2

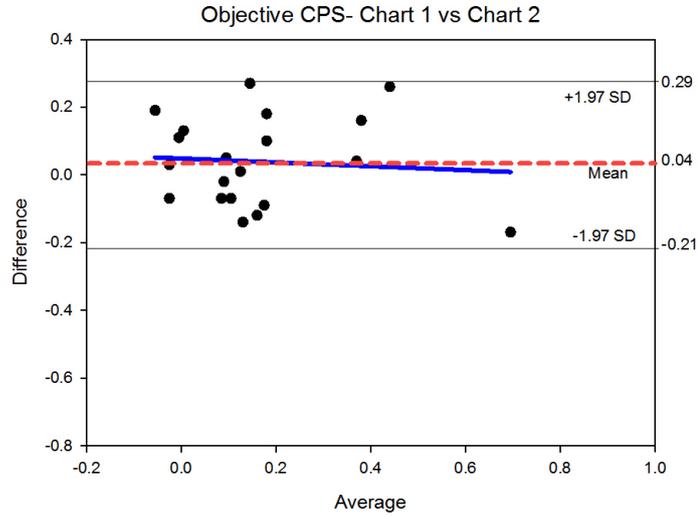


Figure 2.19: Difference plot Chart 1 versus Chart 2

2.4.6 Discussion and Conclusion

Correlation values between the Radner chart and the App for all of the reading speed metrics are summarized in Table 2.6. Strong correlation was found between all reading speed metrics taken with either the Radner chart or the App. One exception was CPS, where the correlation was less strong. This is to be expected, because reading speed values have previously been found to fluctuate more when approaching the CPS [Cheung et al., 2008]. A previous paper which examined the inter test repeatability of the Radner chart found similar CPS correlation [Stifter et al., 2004a] (see Table 2.7). The same paper listed the inter-test repeatability values of all the reading speed metrics and they show remarkable resemblance to the inter-test repeatability of the app itself, except for stronger correlation between objective CPS values with the app. Repeatability was better than that reported for a recently devised silent reading test which had limits of agreement (LOA) of ± 40 wpm [Ramulu et al., 2013] compared to ± 22.5 wpm with the mobile app. No data was given on the CPS.

Optimal reading speed values were demonstrably higher for the app based charts than for the traditional paper based chart by around 29 wpm. Even when comparing reading speeds across all of the acuities, as shown in Figure 2.7, the app based chart measured higher values. Previous studies using Radner reading sentences on young patients, found reading speeds (205 wpm, 209 wpm [Alió et al., 2008, Radner et al., 2002]) comparable to the apps (195 wpm, 194 wpm) but not to the Radner chart used in this study (166wpm). It can be assumed that the lower reading speeds were a result of an examiner offset error during time measurement, a known drawback of using manual timing methods and paper based charts. For example a timing offset of 500ms will greatly affect reading speed (-40 wpm) for a 3 second sentence.

The examiner offset time may also explain why strong positive trends are seen in Bland-Altman plots for optimal reading speed in Figure 2.11 and Figure 2.12. An offset of 500ms will have a much smaller effect for a 7 second sentence (-8 wpm.) than it did with a 3 second sentence. The app charts also had a lower critical print size, but similar threshold reading acuity, which is likely to relate to the 90% cut-off value of the fitted exponential allowing a greater reduction in reading speed before the criteria was reached. Whilst reaction times will play a part even in the app based charts, the automatic, instead of the manual unveiling of sentences will make time measurements more accurate. Ideally an objective method of determining sentence duration would be beneficial to remove examiner error.

2.4 Study: Comparing the performance of the new Aston Read computerised reading test to that of the traditional Radner Reading paper based test

Despite traditional paper based reading tests providing reliable and reproducible measurements [Stifter et al., 2004a, Mansfield et al., 1993], they have been infrequently used in practice. In a clinical setting the computerised Aston Read chart has clear advantages over paper based charts. The Aston Read chart has the ability to control illumination, scale easily for a patients reading distance and reduce testing time. Furthermore automatic scoring and displaying of patients data lessens the burden on the examiner and can reduce subjective bias. The results from this study demonstrate that while reading metrics between the app and the paper based charts are not interchangeable, they correlate highly and thus the app based charts have the potential to capture functional visual ability in research studies and clinical practice. The reader should take the findings above with caution as differences were not compared using true post-hoc tests and this risks type I statistical errors.

Table 2.6: Correlation of metrics between charts.

Metric - Correlation	RR and chart 1	RR and chart 2
Reading Acuity	0.706	0.709
Optimal reading speed	0.887	0.811
Objective critical print size	0.597	0.749

Table 2.7: Repeatability of the Aston Read and the Radner Reading test.

Metric - Repeatability	Radner Reading [Stifter et al., 2004a]	Aston Read
Reading Acuity	0.991	0.936
Optimal reading speed	0.967	0.924
Subjective critical print size	0.474	N/A
Objective critical print size	N/A	0.769

2.5 Further development of Aston Read

The previous study, as described in section 2.4, determined that the iPad based Aston Read chart could be used in place of the traditional paper based Radner Reading chart. However improvements to the software design could be made to improve the test procedure. The main criticisms were as follows:

- Test sentences were shown immediately one after another. No time was allowed for the subject to adjust to the next sentence. This may cause a fatigue affect in subjects after sentences have been read in quick succession, although this was speculative.
- It was noted that some patients had not finished reading aloud before pressing the “Read” button. Additionally, if this app was to be used with elderly patients they may struggle to tap the “Read” button quickly enough after they had stopped reading. With short test-sentences it is crucial that any disruption caused by reaction times can be minimized.
- Positioning the patient at 40cm required some sort of physical measuring device, separate of the chart itself. Once the distance is measured at the start, subjects could move towards the chart reducing the validity of further results.
- The evaluation screen only displayed and played the last sentence that was read. To be able to further increase accuracy, all of the test sentences should be available to evaluate.

2.5.1 Additions and improvements

A second prototype was created to address and validate the criticisms mentioned in the previous section .

2.5.1.1 Time delay between sentences

Before sentences are displayed to the user a two second delay is enacted. This serves two purposes. Firstly, this may address any fatigue issues experienced from succeeding sentences being displayed too abruptly. Secondly, the microphone can continue to record in the event that a patient may have touched the “read” button prematurely. The timings for the delay can be seen in Figure 2.20.

2.5.1.2 Audio Recording and Voice detection

To reduce the variability of manual time measurement it was decided an automated system of determining time duration could be used. Voice detection was incorporated into the app to time the exact moment the words were read. A diagram describing the workings of the voice detection can be seen in Figure 2.20. To begin with the microphone would start recording a second before the sentence is unveiled to the user. Once the sentence is unveiled, the microphone records the users voice until one second after the “read” button is pressed. The automated voice detection, works by a simple threshold method as seen in Figure 2.20. The threshold was set at 0.01V in order to stop false detection from background noise whilst also being small enough to detect unstressed vowels. Only the first and last samples from the microphone that were greater than the threshold were needed to find the duration.

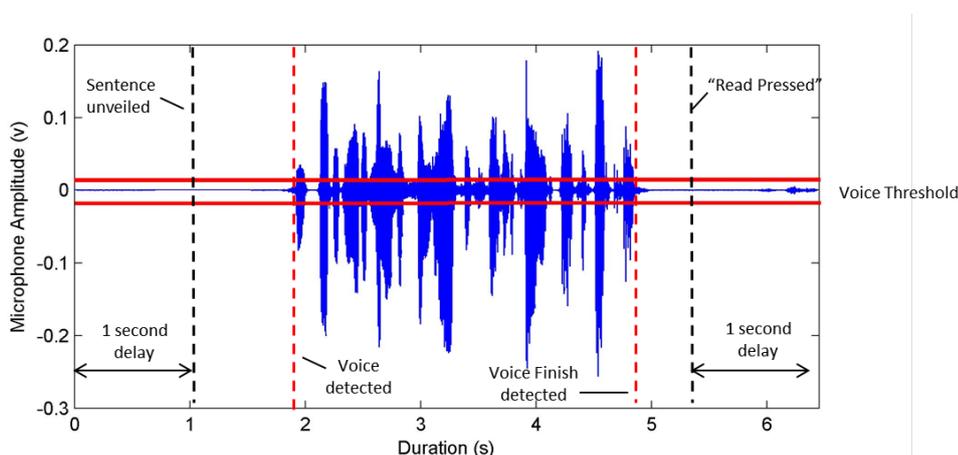


Figure 2.20: Audio sample of a single Radner sentence

2.5.1.3 Eye detection and distance estimation

Positioning patients at the correct distance from a reading test is crucial to its accuracy. Even an 8 cm deviation from a chart designed to be read at 40cm can alter reading acuity metrics by as much as 0.1 logMAR. A system was developed for the Aston Read test that could use the front facing camera and an eye tracking algorithm for primitive distance measurement. The system was to use an eye detection algorithm to detect the separation (in pixels) between the patients eyes, called pupillary distance (PD), from images produced by the iPad’s camera. When the iPad was held close to a patient the separation would be large; when it was

held further away the separation would decrease. Using an estimated value for PD, patient distance could be deduced.

A theoretical model was developed in order to understand the ideal accuracy of the camera based distance measurement. The mean pupillary distance value from the literature was $64.6 \pm 3.7 \text{ mm}$ for white European males [Dodgson, 2004], so the object height, h_{object} , was set at 64.6mm. The camera focal length, f , was 2.18mm and the image sensor resolution, r_{sensor} , was $2925 \frac{\text{pixels}}{\text{cm}}$. Therefore, using the single lens magnification formula (see Equation 2.8) the image height h_{image} at each working distance d could be found. The image pixel height, p_{image} on the camera's sensor could also be calculated using Equation 2.9.

With 95% confidence we can assume the actual PD of any patient is within $\pm 8.9\%$ of the mean therefore the pixel separation can also vary by the same amount. Figure 2.21a shows the inherent variability of any separation measurement based on an assumed pupillary distance and an eye detection accuracy of 100%. Hence using the mean PD separation curve to predict distance values would result in the positional accuracies described in Figure 2.21b.

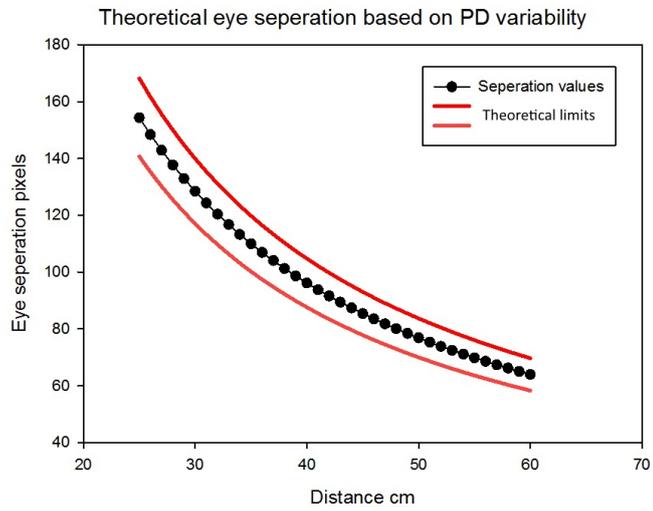
$$h_{image}(d) = h_{object} \left(\frac{f}{f - d} \right) \quad (2.8)$$

$$p_{image}(d) = r_{sensor} * h_{image}(d) \quad (2.9)$$

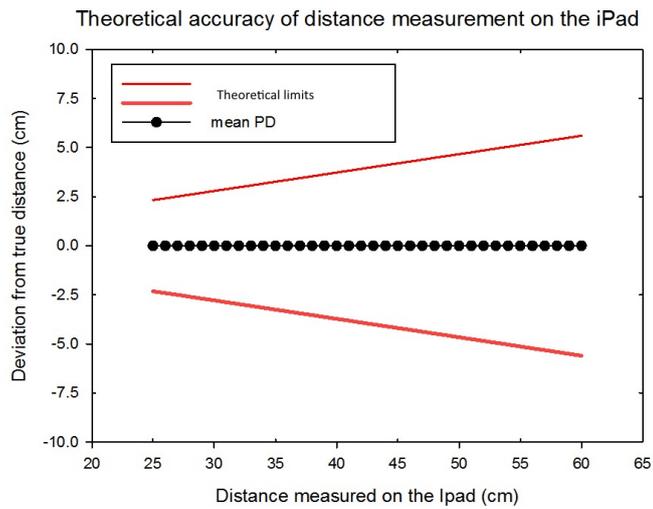
A Viola-Jones face detection algorithm [Viola and Jones, 2004] came supplied with Apple's extensive software development kit and was used for eye detection. Images were routed directly from the front facing camera and processed internally by the iPad's processor and GPU cores. During a reading test this would be put on a background thread invisible to the user. However, for calibration purposes, the image and pupil separation could be made visible, as shown in Figure 2.23. For testing purposes, a single line of string with increments of 2.5 cm marked along its length was created and attached to the iPad. Three subjects were asked to hold the string at their eye level and the pixel separation values were recorded for 15 set distances.

Results for each subject are displayed in Figure 2.22a as well as the theoretical limits.

Many of the subjects results fell within the theoretical limits but not all. As expected other minor sources of variability besides the PD are present. It can be reasoned that the eye detection algorithm contributes to variability. Measurement error using the string could also be present, as subjects were asked to subjectively hold the string at eye level. Although this experiment was limited by the use of three subjects, the theoretical model does give an indication as to the sort of accuracy that can be expected from a camera and PD based distance estimation (see Figure 2.22b).

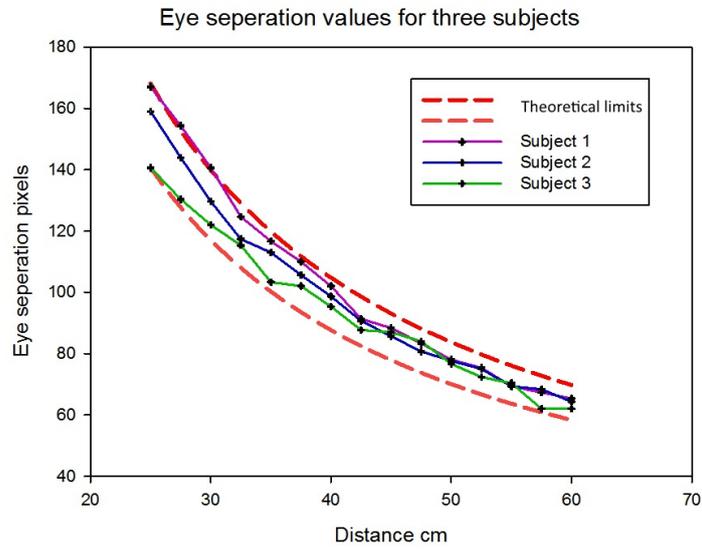


(a) Theoretical separation values for the distance estimation system.

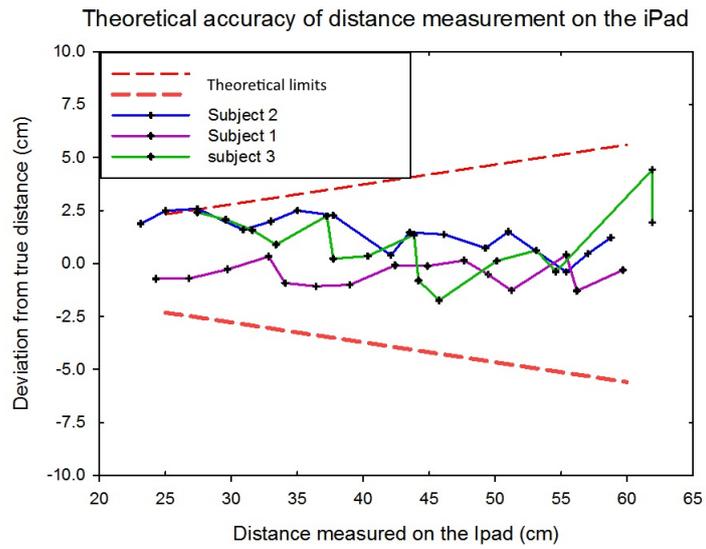


(b) Theoretical deviations from distance measurement

Figure 2.21: Estimated outcomes of distance measurement



(a) Measured separation for 3 subjects



(b) Measured distance variation for three subjects

Figure 2.22: Tested outcomes of distance measurement

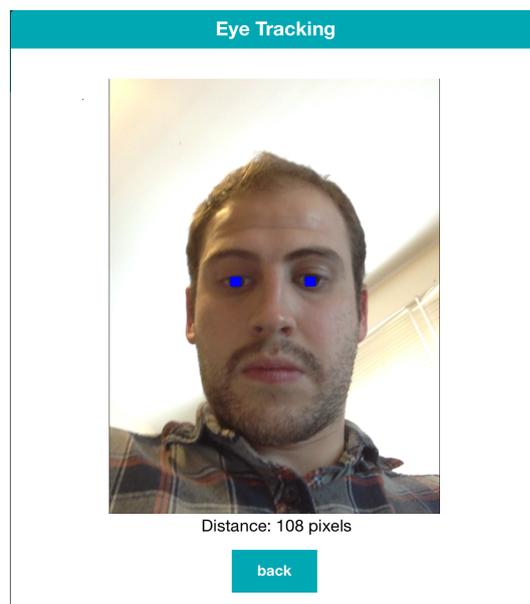


Figure 2.23: Eye Detection and pixel separation.

3 Development of a Mobile, Near and Distance Contrast Sensitivity Test.

3.1 Introduction

Many visual mechanisms use the relative, rather than absolute difference in luminance, to better distinguish objects from each other [Owsley and Sloane, 1987, Avidan and Harel, 2002]. This relative difference between two luminance values is called contrast, and the reciprocal of the contrast needed to identify separate objects in a static image is called the contrast sensitivity. By varying both spatial frequency and contrast of stimuli such as sine wave gratings presented to subjects, a contrast sensitivity function (CSF) can be deduced. One of the primary findings of psychophysical research over the last four decades has been the fundamental role the human contrast sensitivity function has in visual perception. The CSF is more closely related to the visibility of faces, road signs and common everyday objects than are traditional high contrast visual acuity measures [Owsley and Sloane, 1987].

Measurements of the CSF can be used when visual acuity alone may not fully characterize functional vision. CSF testing has been used effectively as a tool to identify aspects of visual function in early vision loss [Di Leo et al., 1992], cataract and refractive surgery [Superstein et al., 1999], glaucoma [Weinreb and Kaufman, 2011], amblyopia [Hou et al., 2010], multiple sclerosis [Balcer et al., 2000] and macular degeneration [Keane et al., 2010]. Despite the usefulness of CSF testing, it is not normally part of the battery of tests required during an optometric examination [Owsley, 2003]. This is due to the high cost and size of contrast testing equipment and the amount of time required to measure the CSF. Presenting hundreds of targets of varying frequency and contrast whilst soliciting patient responses, as is frequently carried out in psychophysics research, is often impractical and time consuming in the clinical setting.

Currently, paper based charts have offered the most convenient form of measuring

CSF but their reliability has been questioned [Pesudovs and Hazel, 2004, Reeves et al., 1991]. Hence a new method of measuring patients' CSF that is more time efficient and mobile than current computerised testing systems, with greater control over illumination, contrast and frequency levels than paper based charts may have clinical significance. This chapter therefore discusses the contrast sensitivity function, identifies current testing techniques and equipment and proposes a new, time efficient, near and far contrast sensitivity chart. More mobile and time efficient testing charts and methods may increase the future adoption of contrast sensitivity testing as a standard clinical test.

3.1.1 Contrast sensitivity function

Ground-breaking behavioral studies performed over 40 years ago by Campbell and Robson [1968] and Blakemore and Campbell [1969] laid the basis of both spatial frequency theory and the shape of the human contrast sensitivity function. In both of these experiments sinusoidal gratings of varying frequency (cycles per degree of visual angle) were presented to subjects. Subjects could control the contrast of the gratings by a knob and they were asked to find the threshold contrast at which the targets were barely perceptible. Once the threshold contrast had been determined the reciprocal, called the contrast sensitivity, could be plotted alongside its frequency.

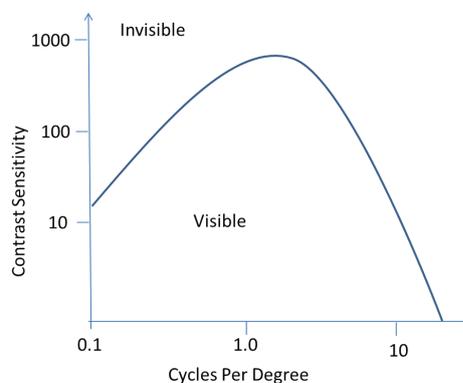


Figure 3.1: Human Contrast Sensitivity Function

An illustration of a typical human contrast sensitivity function (CSF) over a wide range of frequencies can be seen in Figure 3.1. Human contrast sensitivity is lowest

at both high and low frequencies but peaks at the mid range frequencies of 2-5 cycles per degree [Wright and Drasdo, 1985, Dorr et al., 2013, Patching and Jordan, 2005]. The cut-off of sensitivity at the high frequencies (18-60 cpd) is caused by both optical and retinal resolution limits [Campbell and Green, 1965]. For low frequencies (around 0.1 cpd) the CSF starts to plateau, possibly as a result of the grating no longer being detected primarily by the fovea [Peterzell and Teller, 1996].

Blakemore and Campbell's experiment not only established the CSF for all subjects, but also demonstrated that adaptation to spatial frequencies can occur. After an initial CSF measurement, subjects were exposed to a grating of a single spatial frequency for a prolonged period of time. Patients' CSF's were retested and it was found that the prolonged exposure of the grating caused the subjects visual system to adapt only at or near the spatial frequency of the grating. Additionally, adaptation was also found to be orientation sensitive [Bradley et al., 1988]. The extent of the adaption gave rise to the multiple channel hypothesis of human vision; the human visual system is composed of multiple, independent, overlapping, narrow band, spatial frequency "channels" each centered at a particular frequency and orientation, whose channel adapts accordingly to how sensitive it is to the frequencies contained within a stimulus.

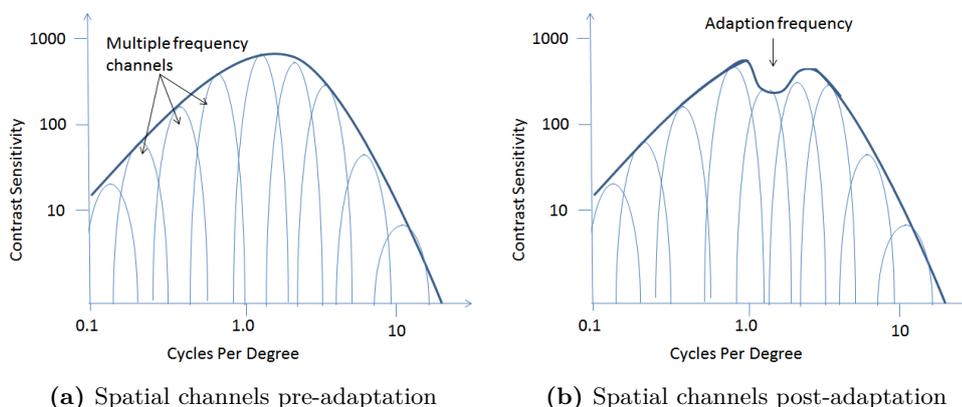


Figure 3.2: Multiple Spatial Channels and adaptation

A diagram illustrating both the adaption and spatial frequency channels can be viewed in Figure 3.2. Masking experiments, which used the addition of noise close and far to the centre frequency of a target, have deduced that the bandwidth of such channels is around 0.75-1.2 octaves either side of its center frequency [Stromeyer and Julesz, 1972] and that there are between 6-8 channels [Wilson et al., 1983, Watson

and Robson, 1981, Regan and Beverley, 1983]

3.1.1.1 Spatial frequency theory

Explanations for the results of psychophysical experiments are generally influenced by the framework of the spatial frequency theory of visual perception as first eluded to by Campbell and Robson [1968] and later refined by DeValois and DeValois [1988a]. In contrast to Hubel and Wiesel's [1959] theory that the building blocks of spatial vision are edges and lines, spatial frequency theory determines that it is sinusoidal spatial waveforms that are the building blocks of perception.

Fourier's theorem states that any continuous signal can be composed of a sum of sine and cosine terms that have a specific amplitude and phase. By making use of Fourier analysis, a 2-D luminance image, much like that which falls on a camera sensor or human retina, can be decomposed into sinusoidal signals each with a specific phase, amplitude and orientation. Interestingly, sinusoidal signals can be recomposed back into an image. No general theorem for lines and edges as primitives of image signals exist as they do for Fourier analysis.

Spatial theorists use Fourier analysis techniques to broadly explain experimental findings. For instance, the detectability of sine and square wave gratings with the same period shown to patients are broadly similar [Campbell and Robson, 1968]. Whilst being counter intuitive to theories that purport edge detection as the primary mechanism (the steeper luminance gradient of a square wave should aid in detection), spatial theorists would suggest that the major constituent sinusoid of a square wave is its first harmonic, hence the perception of both are broadly similar (not including a square wave's more minor harmonics).

The theory can also accurately predict results from spatial frequency experiments where prolonged exposure to a grating of certain frequency will cause a shift in the perception of gratings. Gratings with lower frequencies than the adapting grating appeared lower and gratings with higher frequencies appeared higher still [Blakemore et al., 1970]. This could only occur if cells were tuned to certain spatial frequencies and thus provides evidence for localized Fourier visual processing.

3.1.1.2 Underlying Physiology

Behavioural methods have indicated that a limited Fourier type analysis appears to occur within the visual system. Whilst psychophysics can explain the hypothetical

processing of the visual system, the psychophysical channels must be implemented somewhere within the structure of the nervous system. Macaque monkeys have been noted to have similar CSF's to humans [De Valois et al., 1974]. Consequently many experiments conducted to find the physiological determinants of contrast sensitivity used these monkeys in lieu of human subjects [Tootell and Silverman, 1988, Movshon and Newsome, 1996]

Physiological evidence of cells located in the macaque striate cortex (also called V1) found both simple and complex cells that were tuned to specific frequencies (0.5-16 cpd), with corresponding bandwidths of between 2.2 to 1.1 octaves [Valois et al., 1982]. A graph showing contrast sensitivity functions for six cells taken from Valois et al. [1982] can be viewed in Figure 3.3a. It also transpired that cortical cells are more responsive to sine-wave gratings than to square wave gratings [Albrecht et al., 1980]. Furthermore the receptive fields of V1 cells have a structure very similar to that of a Gabor function [DeValois and DeValois, 1988b] as do cat V1 cells [Jones and Palmer, 1987]. The Gabor like response from two V1 cells is illustrated in Figure 3.3b, where the firing rate relative to baseline is on the Y axis and the position of a narrow bar target across the receptive field of the cell along the X axis. Gabor filters, which can be used effectively for localized Fourier analysis in computer vision, may also have the same function in the human visual system.



(a) Contrast sensitivity functions for six cells in the striate cortex taken from [Valois et al., 1982]



(b) Striate cell receptive field reproduced from [DeValois and DeValois, 1988b]

Figure 3.3: Electrophysiology of macaque striate cells

3.1.2 Threshold Measurement

Determining contrast sensitivity thresholds broadly follows the psychophysical methods of adjustment, limit, and constant stimuli that were first standardized by Fechner [1860]. Generally these methods were designed when it was thought that sensory thresholds were like steadfast boundaries outside which a patient could not see. Figure 3.5a provides an illustration of this classical threshold theory. Although the classical threshold explanation has been invalidated [Perez et al., 2010], superseded by signal detection theory [Green and Swets, 1966], and its flaws identified (such as its inability to isolate decision criterion from sensory perception [Swets et al., 1961, Laming, 2013]), its influence on CSF testing methods has remained.

3.1.2.1 Method of Adjustment

In contrast sensitivity experiments that use the method of adjustment, patients are presented with either sine-wave gratings or Gabor patches of a set frequency. A dial or button provides feedback to a PC or analog circuit that directly affects the contrast of the stimuli. Subjects are tasked with controlling the contrast until they no longer perceive the stimulus (descending trial) or if the contrast is set low to adjust until they perceive the stimulus (ascending trial). The contrast at which the patient perceives the threshold measurement.

The method is not only subject to response bias [Woods and Thomson, 1995] but thresholds are higher than those that use the standard two alternative forced choice design (2-AFC) [Laming, 2013]. Hysteresis effects are also present. In ascending trials, the contrast sensitivities are higher than those with descending trials. Despite these shortcomings the early research that first found, classified, and theorized CSF used this method effectively [Campbell and Robson, 1968, Blakemore and Campbell, 1969] and it is demonstrably quicker for CSF assessment (<6 minutes to test at 5 spatial frequencies [Ginsburg and Cannon, 1983]) than discrete presentations of fixed stimuli. Hence it may have clinical use beyond research where time constraints prohibit fixed stimuli methods.

3.1.2.2 Method of limits and Method of Constant Stimuli

Rather than a continuous adjustment of contrast, the method of limits and method of constant stimuli seek to discretize the process of contrast sensitivity into many

trials. On each trial the stimulus will either be present or absent (called a catch trial). The patient is required to decide whether they have seen the stimuli in a YES/NO type fashion. As both frequency and contrast must be varied between presentations it can take many trials before the CSF curve is fully defined. By using and recording catch trials data, an estimate of discriminability independent of the criterion can be calculated (see subsection 3.1.2.4). Method of limits requires trials to be organized in either ascending or descending order of contrast, meanwhile the method of constant stimuli requires complete randomization of presentation to avoid hysteresis errors.

Criticisms of these methods include that full testing requires significant time to undertake compared to the method of adjustment, as in each trial both frequency and contrast must be varied. Response bias is also a factor.

3.1.2.3 Two-Alternative forced choice

While 2-AFC is a derivative of the method of constant stimuli, its popularity and widely adopted use requires further explanation. Temporal 2-AFC experiments involve two targets that are presented sequentially with a small time delay between presentations [Woods and Thomson, 1993, Legge and Foley, 1980]. Only one of the targets will have the grating stimuli. Patients are required to indicate in which presentation the grating stimuli can be seen. Spatial 2-AFC experiments meanwhile present both targets at the same time. Another variation is to present a target grating at a specific orientation; patients are required to select the orientation between the choices that are presented to them.

Forced choice methods in CSF testing reveal higher thresholds than the YES/NO method of constant stimuli and are less variable and more reliable [Vaegan and Halliday, 1982]. Patients can give better than chance answers even when they believe they are guessing. Additionally, it is assumed that the answers given in 2-AFC are unbiased [O'Mahony, 1992, Macmillan and Creelman, 2004]. This fits neatly with experimental evidence from signal detection theory see (subsection 3.1.2.4 for an explanation) and, as such, this method is the "Golden Standard" in psychophysical research. Using more alternative choices (i.e 4-AFC or 8-AFC) reduces the likelihood of randomly guessing correct answers which, in turn, reduces the number of trials needed to get an accurate assessment of CSF [Jäkel and Wichmann, 2006].

There is however temporal and spatial bias [García-Pérez and Alcalá-Quintana, 2011]

but with the proper procedure and repeated trials the effect is negligible [Jäkel and Wichmann, 2006]. Blackwell [1952] and Jäkel and Wichmann [2006] found temporal 2-AFC methods to be the best method for sensory determinacy and reliability in experienced patients. For naive patients like those seen within the clinic, spatial 2-AFC methods were also preferred.

Adaptive procedures With two dimensions (contrast and frequency) to vary for each discrete trial in 2-AFC procedures, finding a patient's CSF consumes considerable time. Limiting the number of frequencies studied can help but, typically, it still requires 50-100 trials per spatial frequency; with five frequencies studied 500-1000 trials can take around 30-60 minutes [Lesmes et al., 2010]. Unlike the classical 2-AFC approach, where all the trials are known far in advance, adaptive procedures will change the trial according to the responses of the subject.

Early attempts by Watson and Pelli [1983] looked at each spatial frequency separately and used a Bayesian adaptive procedure to optimize the contrast of the next trial. Lesmes et al. [2010] took this idea further and fitted initial trials to a function that approximates the shape of the CSF in order to create trials near to a patient's predictive sensitivity at the next frequency. They claimed that as few as 100 trials (10 minutes) are needed for an accurate determination of CSF using their method .

3.1.2.4 Signal Detection Theory

Signal detection theory (SDT) relates to the way human observers perceive stimuli under conditions of uncertainty [Tanner and Swets, 1954], see Figure 3.4. This theory suggests that two separate internal factors determine the outcome of CSF experiments; the underlying sensitivity of the visual system to sine wave gratings (Sensory process) and the psychological criteria used for deciding whether the sine wave gratings were detected (Decision process) [Green and Swets, 1966].

A psychometric function representing the number of correct responses when subjected to multiple Gabor patches of different contrasts at 2.0cpd can be seen in Figure 3.5a [García-Pérez et al., 2011]. Notice how there is not a definitive contrast threshold from which the Gabor patch stimuli can be detected with complete accuracy. Instead, an S-shaped curve arises which can be explained by SDT.

The sensory process is said to be subject to internal noise. The theory goes that if the same stimuli are presented to subjects repeatedly, the sensory response will

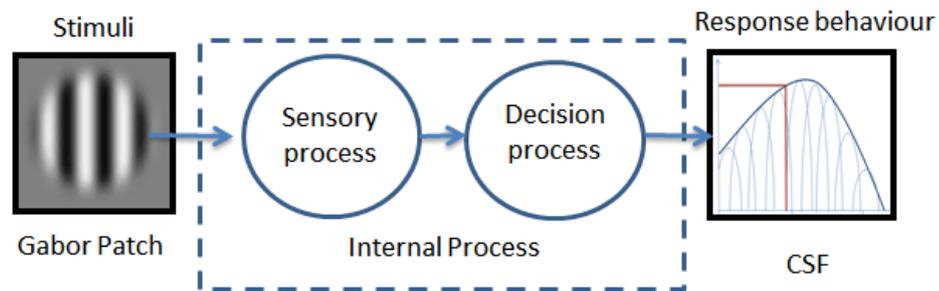
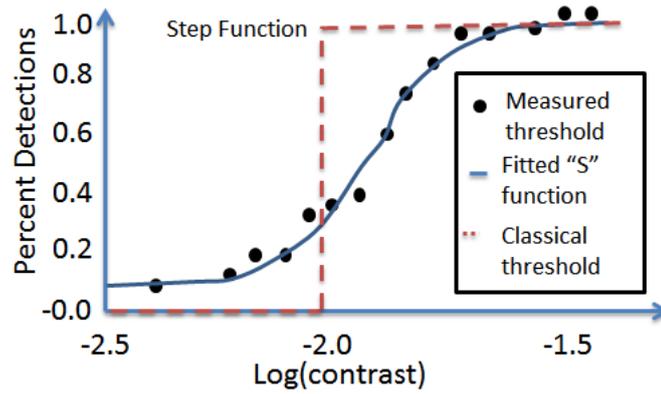


Figure 3.4: Detection theory

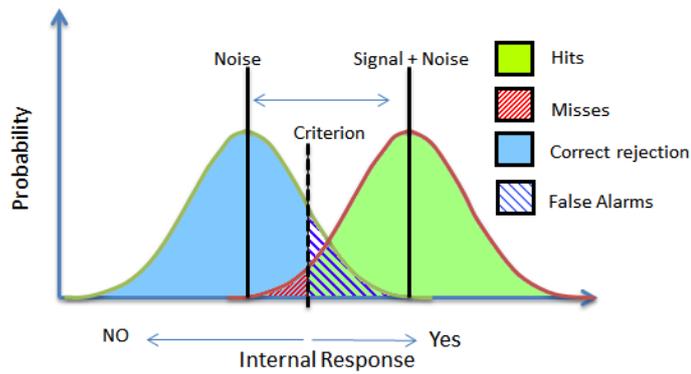
be different for each trial. Probability distributions of the noise-alone (when blank targets are presented) and the signal+noise distribution (when gratings are present) for signal strength are illustrated in Figure 3.5b. As shown in the same figure, there are four possible outcomes of YES/NO type trials; hits, misses, false alarms and correct rejections. It is the criteria and strength of those distributions that sets the boundaries for when these occur. The size and shape of the distributions are completely separate from the decision criterion. Validation of this theory came from assessing receiver operating characteristic (ROC) curves found in contrast sensitivity testing. The predicted ROC curves given by the equations that describe this theory matched those found in psychometric testing.

SDT can be used to compare 2-AFC methods to those of the YES/NO based methods [Macmillan and Creelman, 2004]. Ultimately, this theory promotes using methods for contrast sensitivity (such as the 2-AFC method) that fix the criterion and reduce biases. It also demonstrates the significance of catch trials (or alternative choices) in psychometric procedures in order to isolate sensory responses from decision making processes.

Variable decision criterion can subtly affect experimental results. For example, when contrast sensitivity was measured before and after exercise using a classical ascending method of limits paradigm a significant difference was found. However when the test was re-run using an 2-AFC method that fixed the decision criteria, there were no significant differences [Woods and Thomson, 1995]. The exercise had in fact not changed the sensory perception of contrast but had changed the decision criterion of the patients.



(a) Psychometric function for a patient presented with a gabor patch at a set frequency [García-Pérez et al., 2011]



(b) Responses for signal and signal+noise distributions

Figure 3.5: Signal Detection Theory

3.1.3 Charts and Instrumentation

3.1.3.1 Computerised tests

Early psychophysical experiments used cathode ray tube (CRT) oscilloscopes and analog function generators whose contrast was manually controlled either by examiners or patients [Arundale, 1978]. To incorporate more complicated stimuli and testing methods, contrast tests were computerised [Valois, 1977]. This allowed automated adaptive procedures to be implemented to reduce testing time [Hou et al., 2010, MacLeod et al., 2010].

A typical set-up will involve an cathode ray tube (CRT) monitor, an interface board and a desktop computer. The digital to analog (DAC) interface board and video attenuator will allow full control of the luminance of the CRT rather than the standard 256 (8-bit) step output by most computer graphics adapters. As such CRT offers greater dynamic range and gray level resolution than can be found with LCD panels [Colombo et al., 2009]. However, this comes at the price of portability. LCD screens, despite having only 8 bit gray levels, can increase their contrast by dithering pixels in a process called “bit-stealing” [Tyler and Chan, 1992]. Bit stealing increases the luminance resolution at the detriment of constant chromaticity. Initial evidence suggests that using the bit stealing technique can accurately mimic contrast testing done on laboratory grade CRT monitors [Dorr et al., 2013].

Tablet Based Tests In recent years a trend for more portable and convenient computing has spurred the development of contrast sensitivity tests onto tablet computers. The first from Dorr et al. [2013] used an adaptive 2-AFC method, previously described in subsection 3.1.2.3 by [Lesmes et al., 2010], whose targets were presented on an iPad screen. Subjects were required to fixate on the center of the screen and indicate whether a Gabor Patch was to the left or right of fixation. Despite suffering from the same contrast issues as all LCD displays, the CSF obtained on a mobile device was indistinguishable from a CRT laboratory setup. It was found that 120 trials (around 10 minutes testing time) was enough for an accurate CSF measurement at 6 spatial frequencies (0.5,1, 2, 5,10 and 20 cpd). Recently, open source tools have been developed to allow the easy configuration and running of psychophysics experiments on tablet computers [Turpin et al., 2014] but, as yet, no adaptive procedures have been implemented.

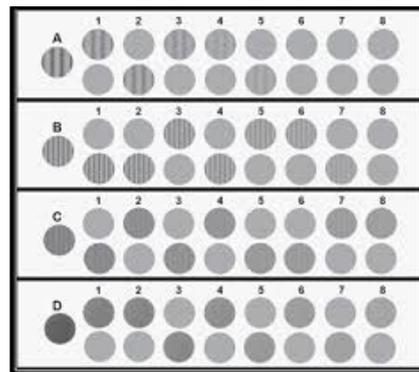
3.1.3.2 Paper Based Testing

It is apparent that computerised testing methods require large expensive equipment and substantial testing time, that make them impractical for a clinical setting. Charts were developed to fulfill the need for a test that could be administered quickly and efficiently. The CSV-1000, VCTS-6500 and FACT charts present sinusoidal Gabor targets with a range of frequencies and contrasts printed onto a single sheet either to be placed on a wall or behind a light box for more constant luminance. The most popular contrast chart, the Pelli-Robson chart [Pelli and Robson, 1988], presents letters of decreasing contrast, where subjects must identify the letters. Whereas computerized methods can render an infinite amount of trial combinations, charts, by their very nature, must limit both the frequencies and contrasts tested. Hence they have coarse steps in both frequency and contrast combinations. The CSV-1000 chart uses a YES/NO type method while the FACT and VCTS-6500 use a 2-AFC procedure. The Pelli-Robson chart uses letter identification with a criteria that 2 out of 3 letters per triplet must be recalled correctly. See Figure 3.6 for images of the paper based charts.

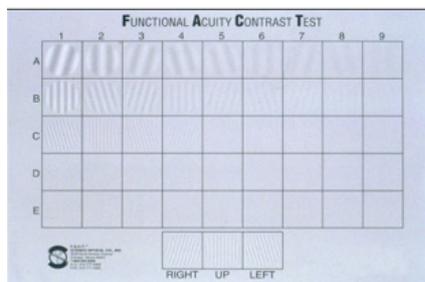
Paper based grating charts have come under intense scrutiny; several researchers found poor reliability with both FACT, VCTS6500 and CSV1000 charts [Reeves et al., 1991, Kelly et al., 2012, Pesudovs and Hazel, 2004] and they question the utility of these in screening patients. Despite their reliability being in question, paper charts have been continually used within research [e.g. Pérez-Santonja et al., 1998, Morales and Rocha, 2011, Wieder et al., 2013] probably due to their ease of use, reduced testing times and range of frequencies tested. The Pelli-Robson chart has demonstrably better reliability [Wender, 2006] and validity [Leat and Woo, 1997] but as all of the letter optotypes are the same size this enables only a single frequency to be studied.



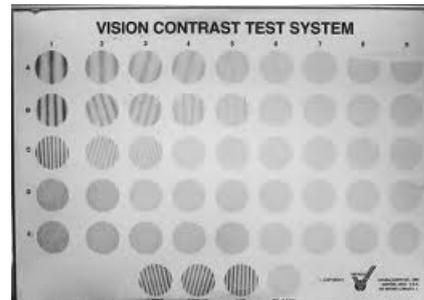
(a) Pelli Robson Chart



(b) CSV-1000 Chart



(c) FACT chart



(d) VCTS-6500 Chart

Figure 3.6: Paper Based Contrast Sensitivity Charts

3.1.4 Outcomes-What makes a good clinical test for contrast sensitivity?

From both signal detection theory and experimental evidence, 2-AFC methods which fix the decision making criterion show the greatest reliability and validity [Green and Swets, 1966, Vaegan and Halliday, 1982]. However the need to show discrete presentations of stimuli varying in both contrast and frequency, and to solicit electronic responses, comes at the cost of increased testing time. Adaptive procedures can reduce the time to approximately 10 minutes [Lesmes et al., 2010] but assumptions are made on the shape of the CSF curve which may not be valid for patients whose ocular health is unknown. From a clinical perspective even 10 minutes is long for a test of functional vision as it may take up a large proportion of an optometric exam.

Whereas reliability is essential for answering research questions, clinical utility also requires time efficiency to be taken into account. The classical method of adjustment provided enough validity and reliability for the studies that defined the CSF and spatial frequency theory [Campbell and Green, 1965, Campbell and Robson, 1968] and so may afford adequate performance in a contrast screening test. Additionally, by eschewing multiple discrete stimuli presentations testing duration can be lowered. Testing at multiple frequencies and contrasts within one presentation could be advantageous.

In terms of equipment, paper based grating charts have been discredited [Reeves et al., 1991, Kelly et al., 2012, Pesudovs and Hazel, 2004]. Lack of constant illumination, paper “Yellowing” and large contrast step sizes may be to blame. Letter based charts like the Pelli Robson are more effective [Leat and Woo, 1997]. However do not test a range of frequencies (different pathologies may effect different spatial frequencies). Furthermore, evaluation of an individual’s CSF function requires data to be plotted graphically, a time consuming and laborious process. Hence paper charts are unsuitable for robust clinical measurement of contrast.

A clinical test should be computerised to provide flexibility in both stimuli presentation, testing method and evaluation. CRT monitors with their increased control of contrast levels and gray scale resolution seem like the most appropriate option to display stimuli but their large size and decreasing availability prohibits their appeal. LCD monitors suffer from lower contrast levels, but with a “Bit-Stealing ” technique contrast sensitivity results are not significantly different from those of CRT monitor

[Dorr et al., 2013]. The slim size of LCD panels means they can be incorporated into tablet computers, further reducing cost.

A new time efficient contrast sensitivity test that uses the method of adjustment, tablet computers and LCD displays to display stimuli should be of clinical value. Whilst contrast testing on computer tablets is not novel, current systems have not used their plethora of sensors and touch screens to reduce testing times effectively. Therefore the author proposes the creation of a new tablet based contrast test that reduces testing time by presenting all spatial frequencies and contrasts in a single stimulus. It is hoped such technology will increase the adoption of contrast testing outside of research and into the clinic.

3.2 Aims and Objectives

3.2.1 Aim

To create a novel contrast sensitivity test on a tablet computer that is more reliable than current paper based charts, has reduced testing time in comparison to existing computerised systems, and is portable enough to be used on the examination chair.

3.2.2 Objectives

1. Understand the limits of contrast testing on tablet computers and LCD displays; this will include contrast resolution measurements.
2. Design a novel chart that renders contrast targets, records patient feedback and displays data to the examiner.
3. Implement the design onto a tablet computer.
4. Compare the new chart to current clinical tests of contrast sensitivity.

3.3 Development

3.3.1 Concept

Several options were explored with regards to the testing stimuli and overall design of the chart. The first involved using discrete targets generated on the iPad and an adaptive 2-AFC procedure. Patient responses could be recorded from a separate iPhone or controller, with the data being beamed back to the iPad over Bluetooth. However, the requirement for reduced testing time excluded the use of discrete stimuli that is normally associated with this type of contrast sensitivity testing. Additionally, work on a similar concept was published recently [Dorr et al., 2013]. The second, was to use the method of adjustment and present targets whose contrast could be adjusted until patients reported that they could no longer see the target, as used by early psychophysical experiments. Again, responses could be beamed over blue tooth from a custom “dial” like controller that would control the contrast. Whilst being notably quicker than a 2-AFC alternative, it would still require discrete presentations that varied in frequency.

The final concept was inspired by a chart designed by Campbell and Robson [1968], see Figure 3.7. Unlike other charts, all contrasts and frequencies could be displayed in one presentation using a swept frequency design. Patients could use the touch screen to trace the boundary between the visible and invisible gratings. This allows the full CSF to be determined rapidly. Repeated measures could easily be incorporated and the results displayed graphically to the examiner. It was also decided that separate near and distance contrast sensitivity tests would be constructed. The distance chart was to be displayed on a separate monitor. Patients would be required to move several sliders that are linked to targets on the monitor to indicate where the boundary between the visible and invisible regions of the chart are located.

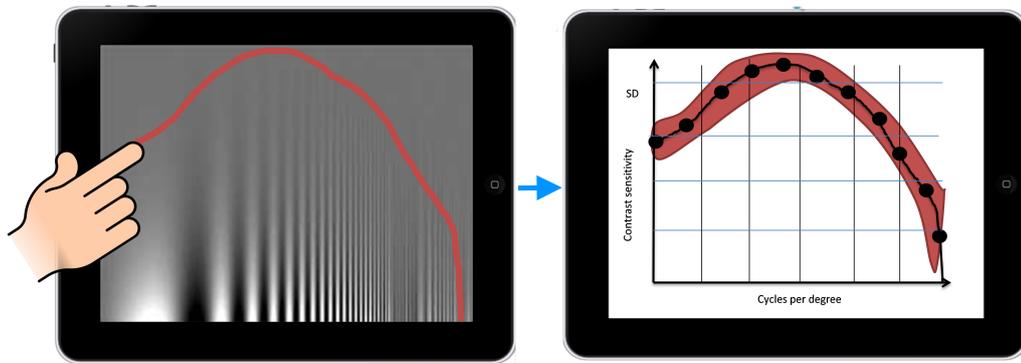


Figure 3.7: iPad concept adapted from charts first described by Campbell and Robson [1968]

3.3.2 Profiling and Calibration

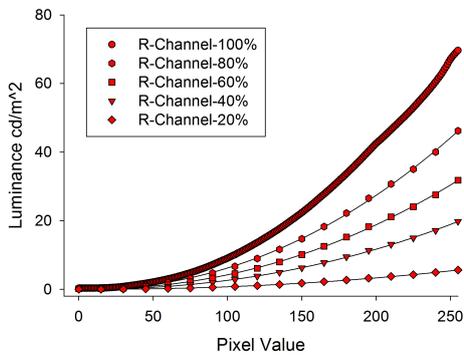
Before the contrast stimuli could be designed, an extensive calibration of the LCD screen on both the iPad and an external HDTV set were conducted. The aim of this testing was to ensure that the screens would be suitable for the enhanced contrast resolution that a “bit stealing” technique could offer. Characterizing each separate R,G and B channel’s luminance is useful when deciding what the overall luminance would be when any or all of the channels are dithered for the bit stealing routine [Tyler, 1997].

A Konica Minolta LS-110 luminance meter was attached atop of a tripod mount and placed 1 meter away from the screen to be tested. A custom program was created that allowed the separate R,G and B channels on the screen to be controlled remotely. Screens were set at full brightness in the software; starting at zero, each channel’s pixel value was incremented in steps of 1 and luminance recorded until the full range of values had been reached. As this technique was time consuming (full characterization for a single brightness setting requires 768 measurements), subsequent “brightness” levels set within software at 20%, 40%, 60% and 80% were tested in pixel steps of 15.

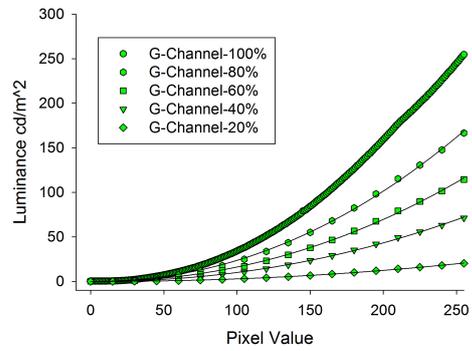
The luminance profile of the iPad and HDTV screen is represented graphically in Figure 3.8 and Figure 3.9. Recommended standards [ITU-R BT.709-5, 2002] for both computer and HDTV screens state that luminance should consist of 21% power from the red, 72% from the green, and 7% from the blue channel respectively [Poyn-ton, 2003]. The sRGB standard [IEC 61966-2-1, 1999] for screens such as the iPad declare a gamma value of $\gamma = 2.2$, whilst HDTV screen standards [ITU-R BT.1886, 2011] recommend a gamma value of $\gamma = 2.4$. A slight difference for the luminance coefficients and the expected gamma value with both the iPad and the HDTV screen were found in comparison to their respective standards (Table 3.1). In particular the HDTV screen does not maintain a constant gamma value (Figure 3.9d) over the entire pixel range but deviates at the low and high pixel values.

Table 3.1: Luminance Coefficients and Gamma Values

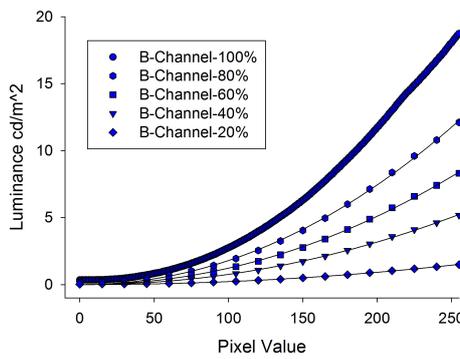
	n_R	n_G	n_B	γ
<i>ITU-R BT.709-5 standard (sRGB & HDTV)</i>	0.213	0.715	0.072	
<i>IEC 61966-2-1 standard (sRGB)</i>	-	-	-	2.2
<i>iPad</i>	0.199	0.741	0.060	2.17
<i>ITU-R BT. 1886 standard (HDTV)</i>	-	-	-	2.4
<i>HDTV</i>	0.237	0.711	0.052	2.49



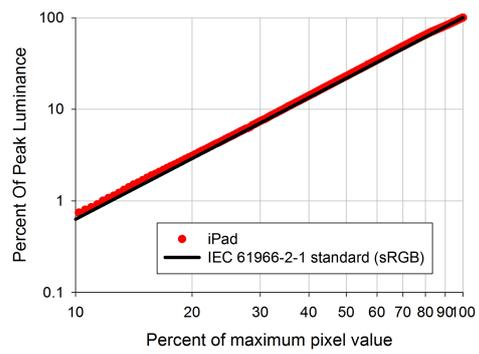
(a) Luminance Profile of the Red Channel



(b) Luminance Profile of the Green Channel



(c) Luminance Profile of the Blue Channel



(d) iPad Luminance Scale

Figure 3.8: Luminance profiles of the R,G and B channels for an iPad Screen

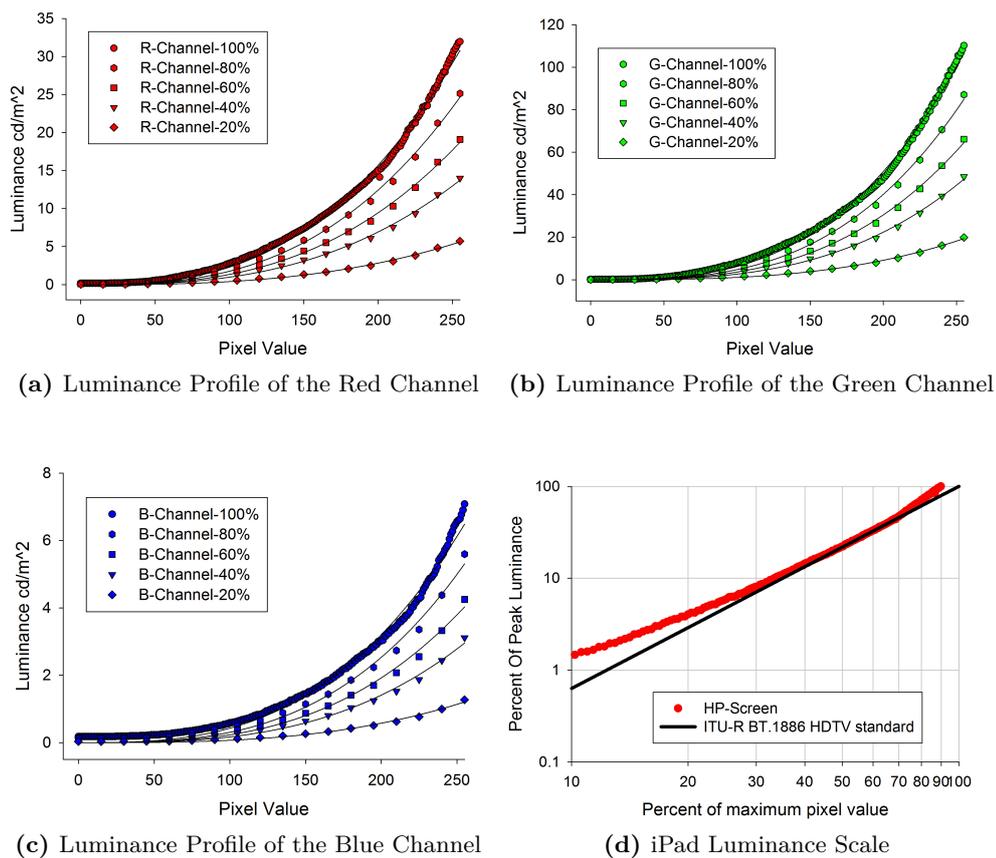


Figure 3.9: Luminance profiles of the R,G and B channels for the HDTV

3.3.2.1 Contrast Resolution

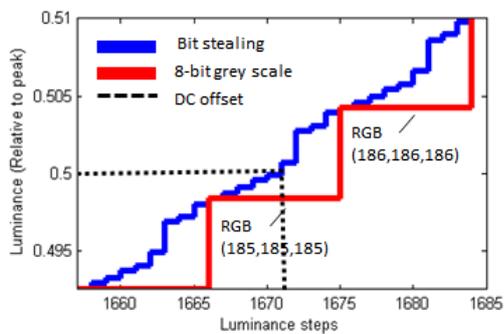
A bit stealing routine requires dithering the RGB sub-pixels to achieve greater luminance resolution at the detriment of a constant chromaticity within the contrast stimuli. Fortunately, it has been stated that humans are more sensitive to luminance changes than chromaticity so a bit stealing technique in contrast sensitivity testing is still valid [Tyler, 1997, Dorr et al., 2013]. The bit stealing steps used to increase luminance resolution are displayed in Table 3.2. Luminance coefficients n_r, n_g and n_b used to calculate the luminance Y in Equation 3.1, were experimentally derived (Table 3.1). All together there were 2294 luminance steps using bit stealing instead of the 255 that 8-bit gray scale offers. The majority of the chart however would be created from values centered around half of the peak luminance, with the iPad this is around the (R,G,B) value (185,185,185) while the HDTV with a larger gamma value will be (193,193,193). Figure 3.10 illustrates the luminance steps achievable with colour bit stealing, in comparison to 8-bit gray scale at 50% of the peak luminance for the iPad. Using the Michelson contrast formula in Equation 3.2 and taking the reciprocal, the maximum contrast sensitivity achievable using the bit stealing technique was 1.22×10^3 for the iPad and 0.99×10^3 for the monitor.

$$Y = n_R R + n_g G + n_b B \quad (3.1)$$

$$C_m = \frac{Y_{max} - Y_{min}}{Y_{max} + Y_{min}} \quad (3.2)$$

Table 3.2: Bit stealing steps

R	G	B	Y (iPad)	Y (HDTV)
0	0	0	0.00	0.00
0	0	1	0.06	0.05
0	0	2	0.12	0.10
1	0	0	0.20	0.24
1	0	1	0.26	0.29
2	0	0	0.40	0.48
0	1	0	0.74	0.71
0	1	1	0.80	0.76
1	1	0	0.94	0.95
1	1	1	1.00	1.00

**Figure 3.10:** Luminance steps available around 50% of the peak luminance

3.3.3 Stimuli design

With the red, green and blue channels fully profiled and the color bit-stealing steps defined, the swept frequency stimuli could be designed using the MATLAB programming environment. Initially a large empty m by n luminance matrix was created whose elements represented each pixel on the either the iPad's or monitor's screen. As illustrated in Figure 3.11, each element of the matrix was denoted by both i and j subscripts that represent their respective row and column. It was decided that the frequency would vary across the columns of the matrix, whilst the contrast would vary across the rows.

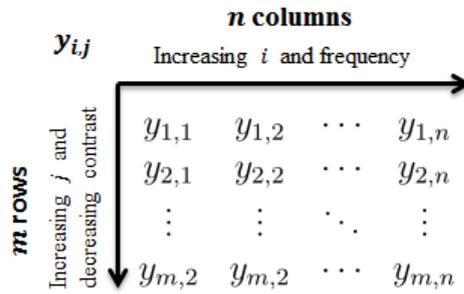


Figure 3.11: Luminance Matrix

The swept frequency profile was created using Equation 3.3, where f_0 is the starting frequency, k is the rate of exponential increase in frequency, $x(j)$ corresponds to the visual angle along the display, $c(i)$ the contrast and $y_{i,j}$ the luminance relative to peak. Equations of this type are called exponential chirp functions and are used extensively within the radio engineering fields.

$$y_{i,j} = \frac{c(i)}{2} \sin \left(2\pi f_0 \left(\frac{k^{x(j)} - 1}{\ln(k)} \right) \right) + \frac{1}{2} \quad (3.3)$$

The visual angle $x(i)$, could be derived for each element in the matrix by inputting the screen width in meters w , screen width in pixels n , distance from the patient d and the matrix column number j into Equation 3.4. Contrast $c(i)$ meanwhile could be computed by using only the screen height in pixels m and the matrix column number i using Equation 3.5.

$$x(j) = \frac{j}{n} \left(2 \tan^{-1} \left(\frac{w}{2d} \right) \right) \quad (3.4)$$

$$c(i) = \left(10^{-3}\right)^{\frac{i}{m}} \quad (3.5)$$

Calculating the rate of exponential increase k requires the starting frequency f_0 , the end frequency f_{end} and the maximum visual angle x_{max} to be known and placed in Equation 3.6.

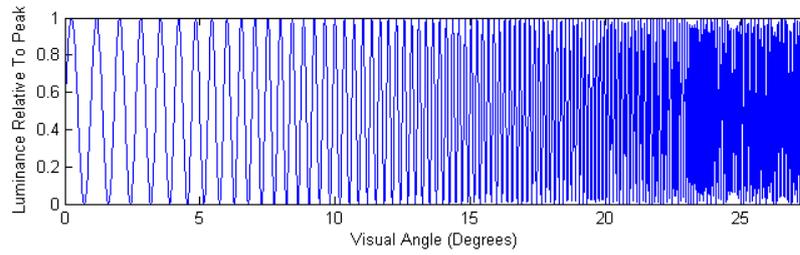
$$k = x_{max} \sqrt{\frac{f_{end}}{f_0}} \quad (3.6)$$

A list of the variables used with the iPad and the HDTV can be seen in Table 3.3. Some are dictated by the physical characteristics of the screen and its displacement from the patient, such as the screen height, width, range and maximum visual angles. Others were design choices such as the start and end frequencies which were chosen to include all of the major frequencies (1,2,8,16 and 30 cycles per degree) that require testing.

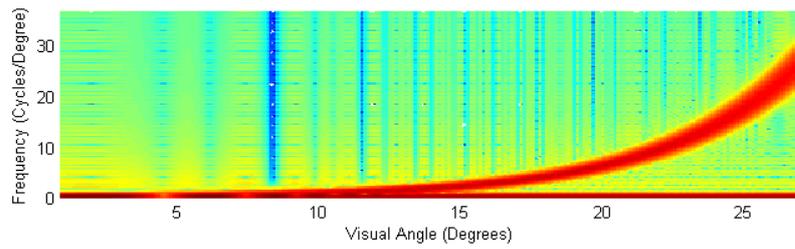
Table 3.3: Variables used in the contrast charts

Variable	iPad Near Chart	HDTV Distance Chart	Description
w	0.192	0.465	Screen Width (m)
d	0.4	2.4	Patient Distance (m)
m	1536	720	Pixel Height (px)
n	2048	1280	Pixel Width (px)
x	$\left\{\frac{1 \times 27}{2048}, \frac{2 \times 27}{2048}, \dots, 27.0\right\}$	$\left\{\frac{1 \times 11.1}{1280}, \frac{2 \times 11.1}{1280}, \dots, 11.1\right\}$	Range of visual angles ($^\circ$)
x_{max}	27.0	11.1	Max visual angle ($^\circ$)
f_0	0.1	0.1	Start Frequency (cpd)
f_{end}	30	30	End Frequency (cpd)

Figure 3.12a and Figure 3.12b illustrates both the swept frequency profile and its accompanying spectrogram for the near contrast sensitivity chart. The spectrogram plot demonstrates the exponential rate of change in frequency that can be seen along the rows of the luminance matrix. Contrast along the columns of the matrix was controlled by varying the amplitude of the swept frequency. Once the luminance matrix was filled, the values could be later mapped to their corresponding RGB pixel values by use of a look up table to produce an image, see Figure 3.13.



(a) Swept frequency profile



(b) Spectrogram of the swept frequency

Figure 3.12: Swept frequency characteristics for the near contrast sensitivity chart

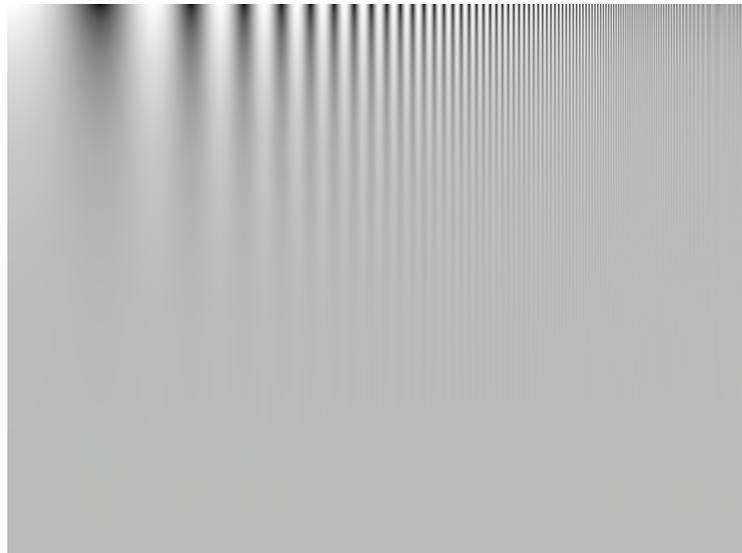


Figure 3.13: Near Contrast Sensitivity Chart

3.3.4 Software-Near Contrast Sensitivity

The near contrast sensitivity test was created for the iPad using the X-Code software development kit in objective-C. A storyboard of the apps user interface can be seen in Figure 3.14, while a basic programming flow diagram illustrating the main structure of the app can be viewed in Figure 3.15. Where as eye tracking can be used to establish approximate patient distance to the screen, as described in subsection 2.5.1.3 on page 76, the contrast app was designed to be held in landscape mode making it difficult for the on-board camera to get a head on image, thus distance measurement was omitted in the software.

Three distinct sections of the app can be seen in the user interface diagram, the introduction screen where an instructional video can be played, the testing screen where the swept frequency stimuli described on page 106 is presented and feedback from the user is collected, and finally the results screen where values are mapped onto a graph for examiner evaluation. Within the test screen, touch responses were recorded in a background thread; each touch event was fitted to a Bezier curve to smooth out sampling errors from the capacitive touch screen. The resulting Bezier curves were then presented to the user in the form of a thick green line rendered atop of the swept frequency stimuli. A large green button positioned at the bottom of the chart could be pressed when the user had finished tracing the curve, and the screen would clear to allow for another curve to be drawn until all repeated measures had been taken.

Once testing has been completed the results screen would appear. A graph presenting the mean contrast thresholds over the whole range of frequencies could be displayed as well as the standard deviation from the repeated measures. In addition a table listing out the contrast thresholds for the major frequencies could be rendered to aid the examiner. From the results screen options to communicate results via e-mail attachment or print services were included.

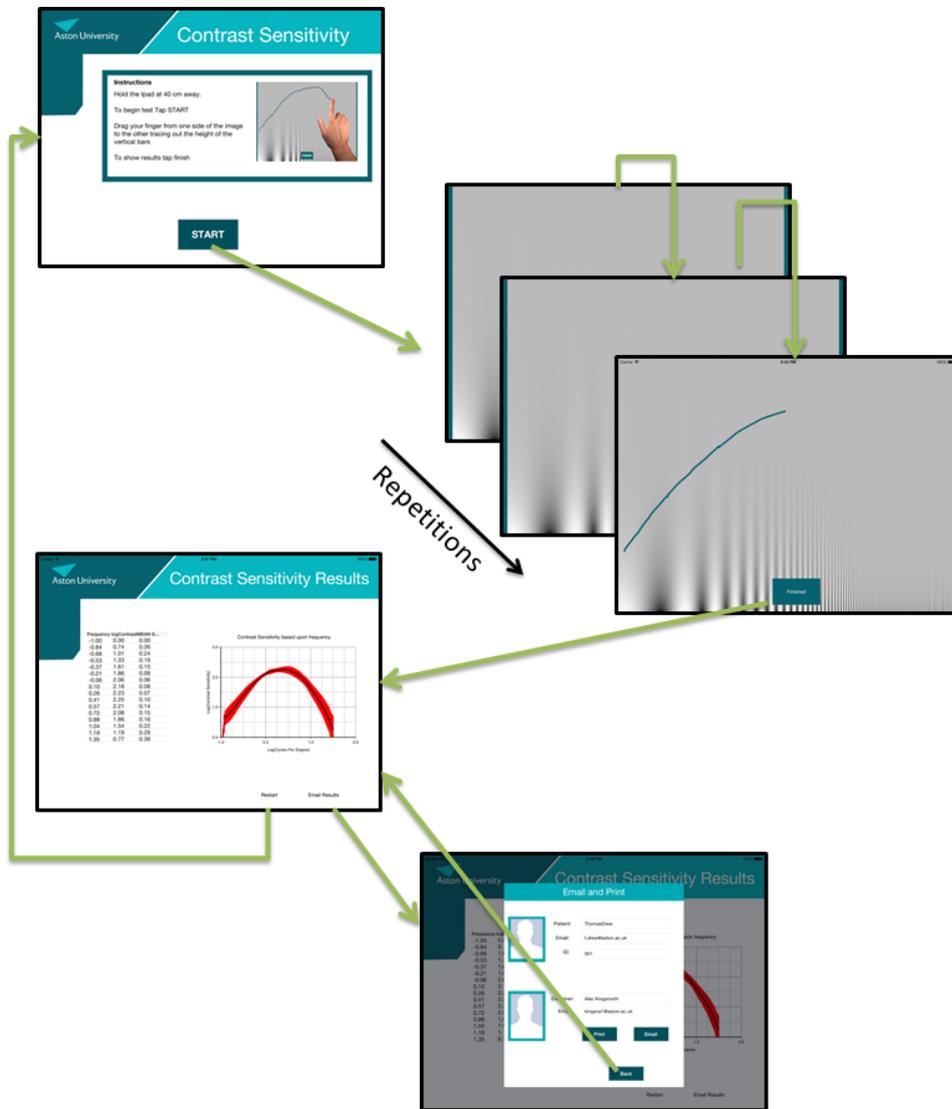


Figure 3.14: Near Contrast Sensitivity App User Interface

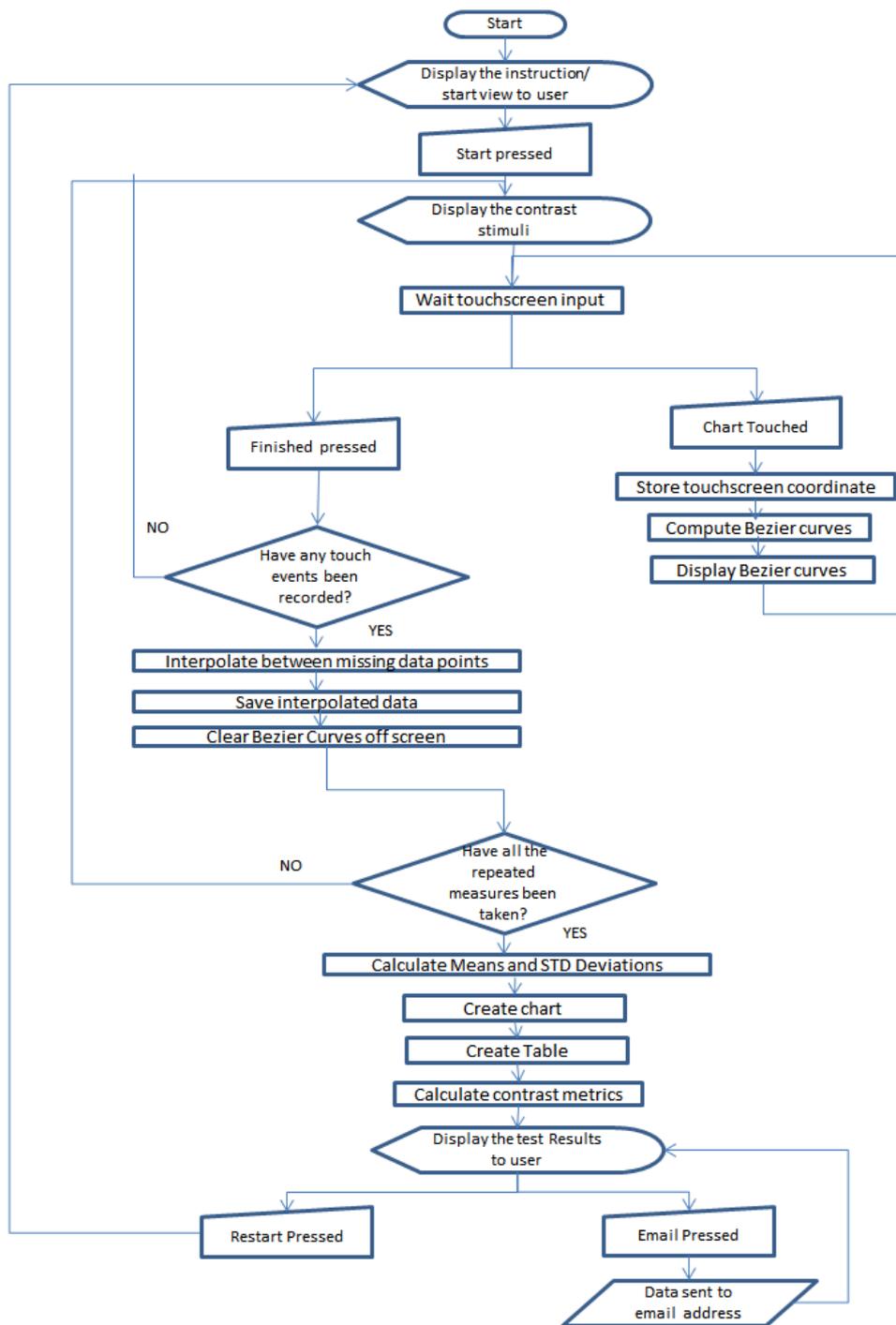


Figure 3.15: Near Contrast Sensitivity Programming Flow Diagram

3.3.5 Software and Hardware-Distance Contrast Sensitivity

3.3.5.1 Hardware

In order to create a dual screen app for the distance contrast sensitivity test, a customized hardware setup was created. A 22" HDTV from HP was connected to a consumer mini computer (Apple TV). The mini PC is commonly used for streaming video and internet content to a television but it is unique in that it can work as a second monitor when paired with an iPad. Communication from the iPad to the mini computer was handled wirelessly by the addition of a mobile WIFI router. Due to bandwidth limitations only a second screen of 1280x720 pixels could be rendered. All of the setup was placed on a custom mounting bracket to the rear of the monitor (see Figure 3.16).



Figure 3.16: Distance Contrast Sensitivity Hardware

3.3.5.2 Software

The dual screen app was designed to stream the swept frequency stimuli wirelessly to the mini computer to be displayed on the monitor. All of the processing required for the app was hosted by the iPad; communication to the mini computer was uni-directional and handled in a master-slave type relationship. API's that come bundled within the x-Code software development kit allow the monitor to appear in the software as a second screen from which to render graphics.

Similar to the near chart there were three main sections to this app, introduction, testing, and results screen. However, the main difference being that the iPad acted as a controller whose sliders controlled red squares superimposed on the contrast stimuli displayed on the monitor (Figure 3.17). Patients were required to move the red sliders to the position where they could no longer see the vertical bars. Once a subject was satisfied the task was complete, a finished button could be pressed and the test repeated. After three repetitions the app would jump to the final screen where both a table and graph of the results would be displayed. Again, options to extract data or repeat testing could be made. See Figure 3.18 for a simplified flow diagram.

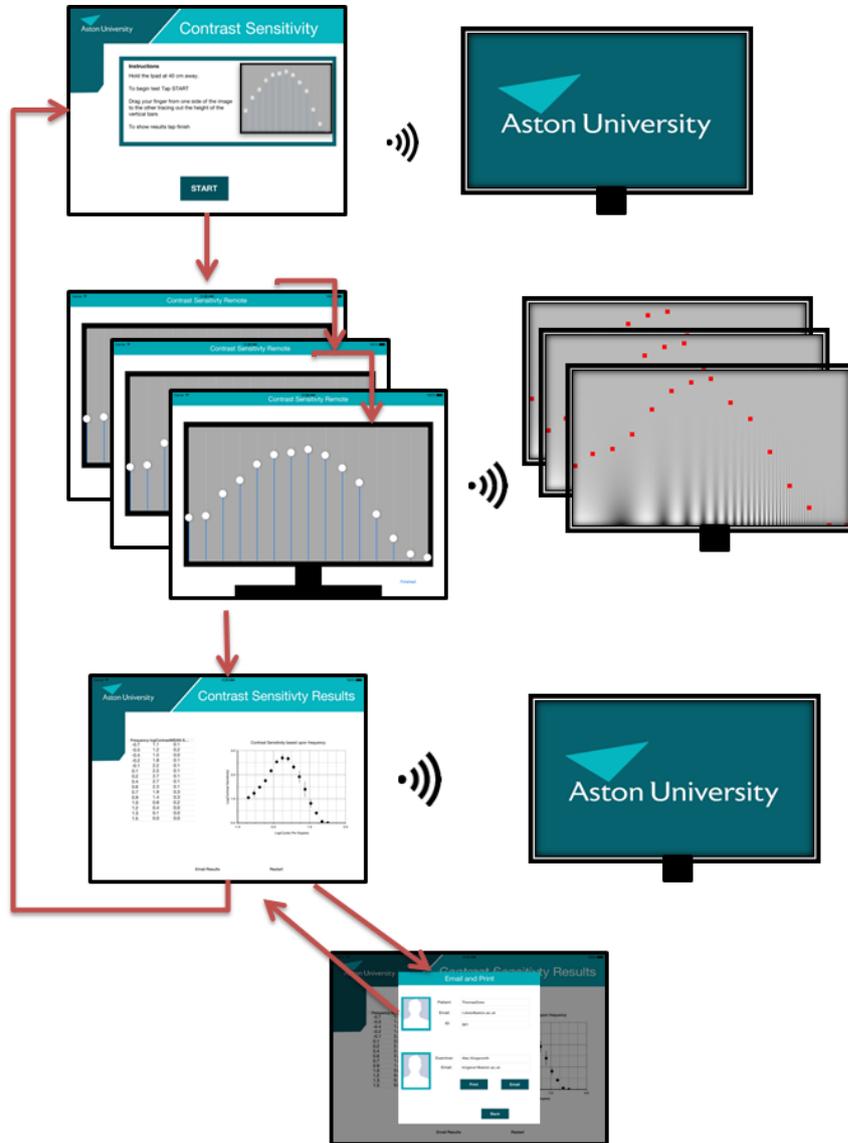


Figure 3.17: Distance Contrast Sensivity App User Interface

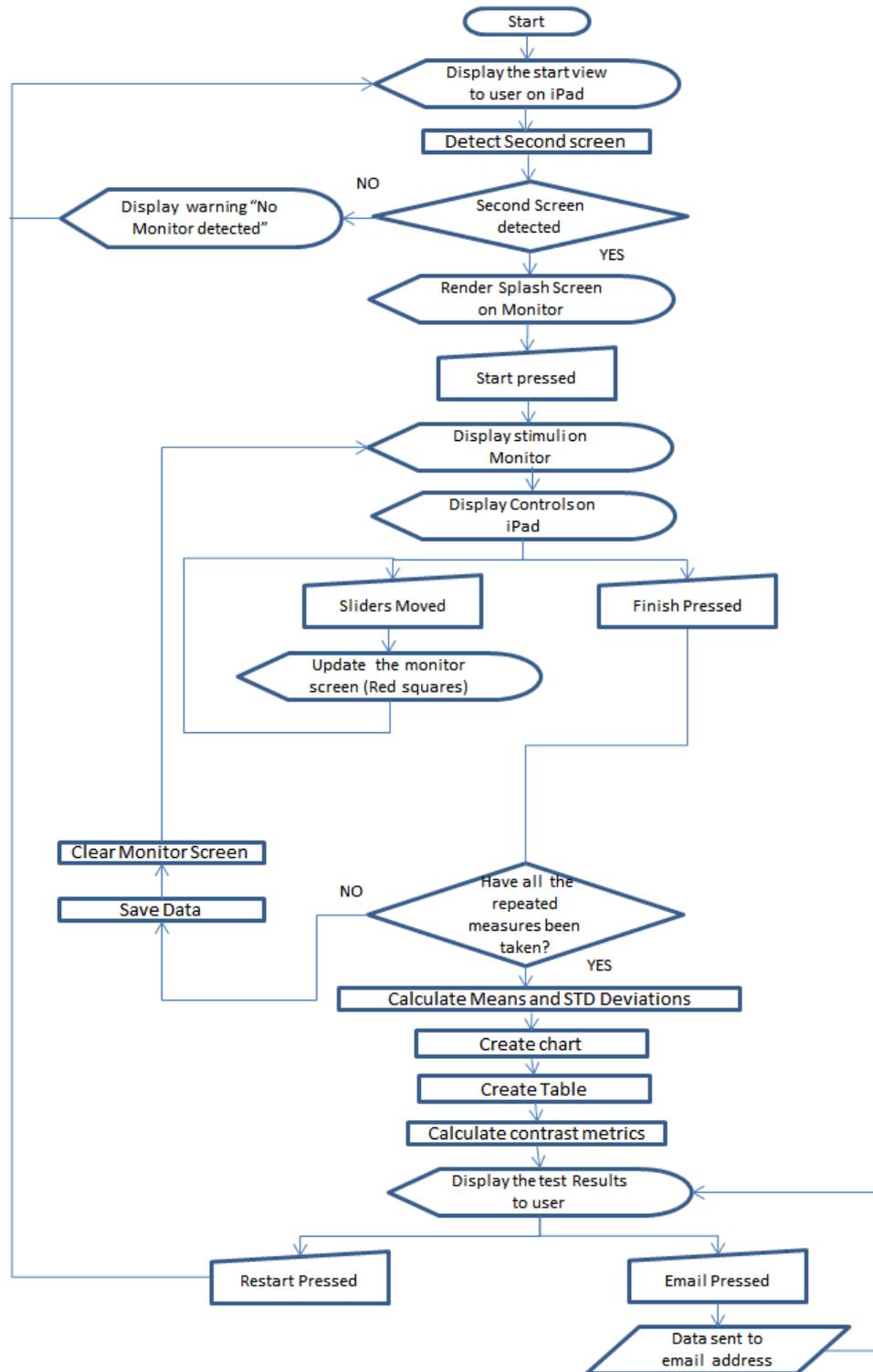


Figure 3.18: Distance contrast sensitivity programming flow diagram

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

3.4.1 Aim

The main aim of this study was to compare the novel near and far contrast sensitivity mobile apps to existing clinical methods of determining contrast sensitivity.

3.4.2 Subjects

Twenty subjects aged 19-28 years (average 23.4 ± 2.1 years) wearing their full refractive correction were recruited for the testing of the new near and far contrast sensitivity apps. All subjects were required to have a VA better than 0.2 logMAR and have no history of ocular disease. Informed consent was obtained from all participants and the study was conducted in accordance with the declarations of Helsinki. Ethical approval for the study was granted by the Aston University Ethics Committee.

3.4.3 Protocol

Four different contrast sensitivity charts were to be evaluated, the near and distance contrast sensitivity apps, the CSV-1000 back illuminated chart, and the Pelli-Robson letter chart. Selection of the traditional contrast tests reflected the two main types of contrast stimuli that can be used in a clinical setting, the letter and gratings based charts (see subsection 3.1.3 on page 93 for pictures and descriptions). Table 3.4 summarizes each contrast sensitivity chart's luminance (tested with a Konica Minolta LS-110 luminance meter) and characteristics which include required distances and frequencies tested.

Table 3.4: Chart Characteristics

Chart	Stimuli Type	Luminance cd/m^2	Distance m	Frequencies cpd
Near App	Swept-Frequency	90	0.4	1.5, 3, 6, 12, 18
Distance App	Swept-Frequency	88	2.4	1.5, 3, 6, 12, 18
CSV-1000	Gratings	71	2.4	3, 6, 12, 18
Pelli-Robson	Letter	78.	1.0	1.5 (Equivalent)

Once informed consent was taken, patients were required to sit for each contrast sensitivity test in turn. A tape measure was used to place patients at the appropriate distance to begin testing. Within each contrast sensitivity chart test, three levels of vision degraded by Bangerter occlusion foils were studied: the no foil condition, the 0.8 foil and the 0.2 foil condition respectively. These three conditions were chosen to simulate different levels of contrast deficiency within patients, with the 0.2 foil previously shown to be the most severe [Pérez et al., 2010]. Bangerter foils were placed inside a trial frame to be wore by the patient. In addition to testing each of the conditions, a repeated measure using the 0.8 occluding foil was taken so an assessment of the contrast charts inter-test repeatability could be undertaken. The repeatability using a 0.8 Bangerter foil would simulate the repeatability contrast deficient patients who are most likely to undertake a contrast sensitivity test. Both the order in which the charts were used and the occlusion conditions were randomized for each patient. All contrast sensitivity tests were conducted monocularly. The time taken to complete each contrast sensitivity test under the no foil condition, was recorded using a stop-watch.

Due to the various chart designs, the testing procedure for each chart was different. For the near app patients were asked to trace a line with their finger on the screen where they believe they could no longer see the height of the various bars, this was taken as the contrast threshold. Similarly, with the distance app patients were also asked to find the height of the bars, however they were required to move sliders on the iPad that shifted red markers on the contrast stimuli to indicate their choices. A modified 2-AFC procedure was used with the CSV-1000 chart; patients were presented with two patches and asked to identify which one contains a sine wave grating. If patients did not recognise any gratings they were asked to respond that both of the patches were blank. The contrast level of the last correct response was taken as the contrast threshold. The Pelli-Robson chart, whose groups of three letters varied in contrast across the design, required that two out of the three letters be identified incorrectly before the group's contrast was taken as the threshold measurement [Pelli and Robson, 1988].

3.4.4 Statistical Testing

Repeated measures ANOVA's with chart type, spatial frequency and occluding foil were ran to understand the global interactions between these independent variables on contrast sensitivity. Furthermore, repeated measures ANOVA's were used to

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

identify the effect of chart type and occluding foil at each individual spatial frequency. If significance was found between chart types paired t-tests and Pearson's correlation coefficients were calculated to understand which charts differed. Inter-test repeatability was assessed using Bland-Altman analyses and Pearson's correlation coefficients.

3.4.5 Results

Mean contrast sensitivities for twenty patients tested on all four charts are illustrated in Figure 3.19. Contrast sensitivities evaluated with the near app showed consistently higher values than the CSV-1000 chart across all frequencies and occlusion conditions. Using data between 3-18cpd there was a statistically significant affect of chart type ($F=29.6$, $P<0.01$), spatial frequency ($F=581.9$, $P<0.01$) and occlusion ($F=320.4$, $P<0.01$) on the measured contrast sensitivity. Moreover significant interactions between chart type and spatial frequency ($F=25.9$, $P<0.01$), and occlusion and spatial frequency ($F=16.293$, $P<0.01$) were found, but none between chart type and occlusion ($F=1.388$, $P=0.25$). The mean time duration to complete contrast sensitivity testing on the CSV-1000 was $124 \pm 37s$, Pelli-Robson $78 \pm 27s$ (one spatial frequency only), Near App $53 \pm 15s$ and the Distance App $107 \pm 36s$.

3.4.5.1 Individual Frequencies across Charts

1.5 cpd At a spatial frequency of 1.5cpd where the Pelli-Robson chart could be compared to the two new apps, there was a significance difference found with chart type ($F=121.8$, $P<0.01$) and occlusion ($F=181.0$, $P<0.01$), yet there was no interaction between them ($F=0.5$, $P=0.74$). Table 3.5 lists the Paired t-tests and Pearson correlation coefficients between charts. It was found significant differences between all charts at all occlusion conditions except for the near and distance apps whilst using the Bangerter 0.2 foil ($P=0.06$), similarly poor correlations were found between all charts (Table 3.5).

3cpd, 6cpd, 12cpd, 18cpd At 3, 6, 12 and 18 cpd where the new apps could be compared to the CSV-1000 both chart type ($F=95.9$ $P<0.01$, $F=12.7$ $P<0.01$, $F=5.8$ $P<0.01$, $F=16.2$ $P<0.01$) and occlusion ($F=235.2$ $P<0.01$, $F=227.5$ $P<0.01$, $F=231.6$ $P<0.01$, $F=214.6$ $P<0.01$) significantly affected results at all frequencies yet interaction between them was found only at 18cpd ($F=3.0$ $P=0.023$). Table 3.6 lists all the paired t-tests and Pearson correlation coefficients for each frequency and at each comparison. Most correlations between charts were weak although some mild correlations were statistically significant (seen in blue), additionally most t-tests found significance although there were a few exceptions (seen in red).

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

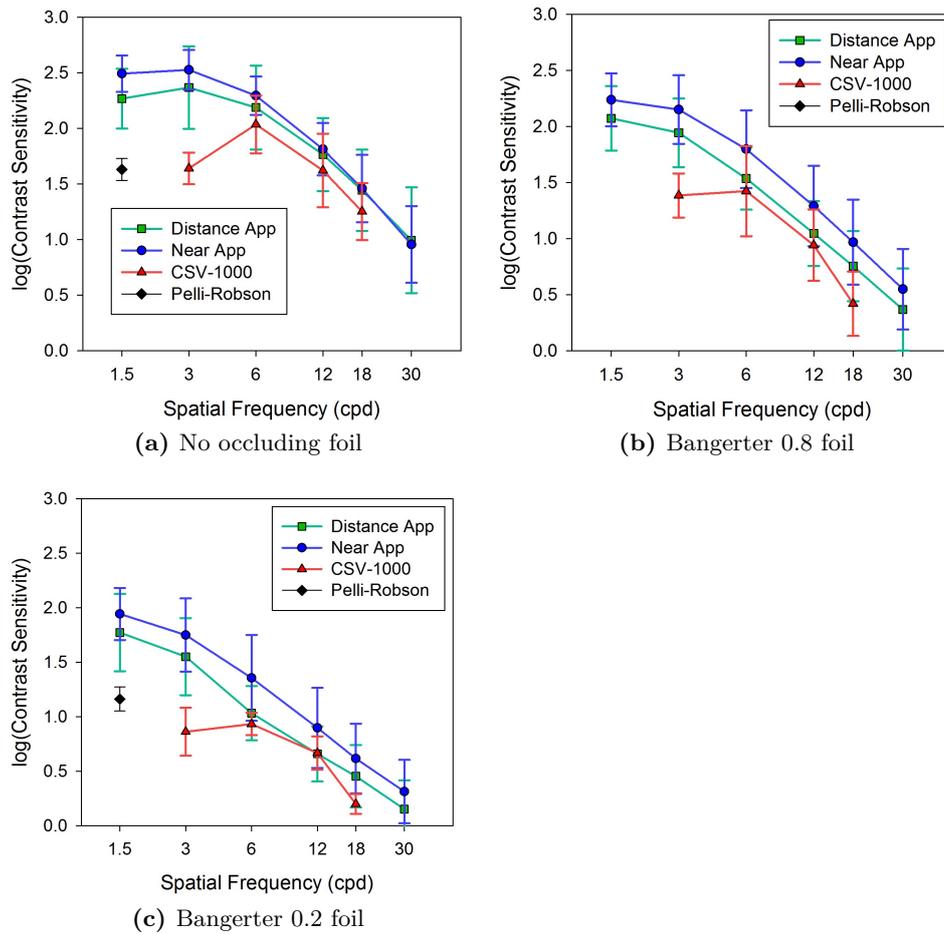


Figure 3.19: Mean results for twenty patients for all charts and occlusion conditions

Table 3.5: Paired t-tests at 1.5 cpd

	No Foil			Bangerter 0.8			Bangerter 0.2		
	N vs D *	N vs P **	D vs P ***	N vs D	N vs P	D vs P	N vs D	N vs P	D vs P
Paired t-test P	>0.01	>0.01	>0.01	0.04	>0.01	>0.01	0.06	>0.01	>0.01
Pearson's r	0.23	0.37	0.09	0.24	0.27	0.31	0.25	0.16	0.07

* comparison between the near and distance app

** comparison between the near app and the Pelli-Robson chart

*** comparison between the distance app and the Pelli-Robson chart

Table 3.6: Paired t-tests at 1.5 cpd

	No Foil			Bangerter 0.8			Bangerter 0.2		
	N vs D *	N vs C **	D vs C ***	N vs D	N vs C	D vs C	N vs D	N vs C	D vs C
3 cpd	Paired t-test P	<0.01	<0.01	0.04	<0.01	<0.01	0.02	<0.01	<0.01
	Pearson's r	0.41	0.314	0.28	-0.04	0.41	0.30	-0.06	0.38
6cpd	Paired t-test P	<0.01	0.04	<0.01	<0.01	0.11	<0.01	0.02	0.32
	Pearson's r	0.58	0.28	0.58	0.38	0.05	0.07	0.53	0.11
12cpd	Paired t-test P	0.56	0.09	0.15	<0.01	0.34	0.02	0.04	0.917
	Pearson's r	0.22	-0.35	0.20	0.44	-0.165	0.24	0.42	-0.266
18cpd	Paired t-test P	0.86	0.04	0.05	0.01	<0.01	0.06	<0.01	<0.01
	Pearson's r	0.18	-0.02	0.18	0.38	0.15	0.27	-0.15	-0.30

* comparison between the near and distance app

** comparison between the near app and the CSV-1000

*** comparison between the distance app and the CSV-1000

3.4.5.2 Repeatability

Bland-Altman plots for the Pelli-Robson, near app, distance app, and CSV-1000 using the repeated measures on the 0.8 Bangerter foil can be seen in Figure 3.20, Figure 3.21, Figure 3.22 and Figure 3.23 respectively. A brief summary of the findings including the coefficient of repeatability (COR) and Pearson’s correlation coefficients for the data can be found in Table 3.7. The Pelli-Robson chart had the greatest agreement (COR 0.14 logCS) between repeated measures at 1.5 cpd followed by the near app whose COR values averaged 0.31 logCS. The near app had a slightly higher average COR values across all frequencies (0.39 logCS), meanwhile the CSV-1000 demonstrated a robust COR of 0.30 logCS at 3 cpd but poor agreement at the higher spatial frequencies. Significant and strong correlation values were found with near app, distance app and Pelli Robson chart at all frequencies studied, however the CSV-1000 only showed strong correlation at 3.0 cpd (Table 3.7).

Table 3.7: Coefficients of repeatability (COR) and Pearson’s correlations using the 0.8 Bangerter foil

		Frequency (cpd)				
		1.5	3	6	12	18
Near App	COR log(contrast sensitivity)	0.26	0.28	0.34	0.37	0.30
	Pearson’s r	0.83	0.89	0.87	0.86	0.92
Distance App	COR log(contrast sensitivity)	0.34	0.41	0.44	0.39	0.39
	Pearson’s r	0.80	0.76	0.77	0.79	0.81
CSV-1000	COR log(contrast sensitivity)	–	0.30	0.87	0.93	0.86
	Pearson’s r	–	0.79	0.29	-0.01	0.07
Pelli-Robson	COR log(contrast sensitivity)	0.14	–	–	–	–
	Pearson’s r	0.76	–	–	–	–

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

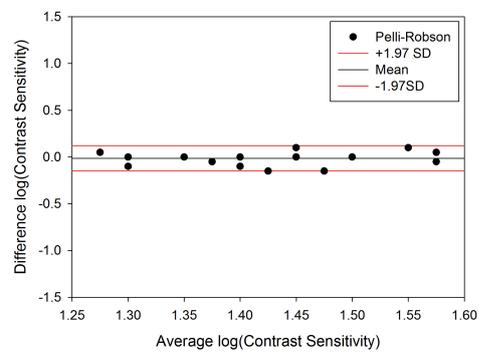


Figure 3.20: Pelli-Robson Repeatability Bland-Altman Plots, n=20

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

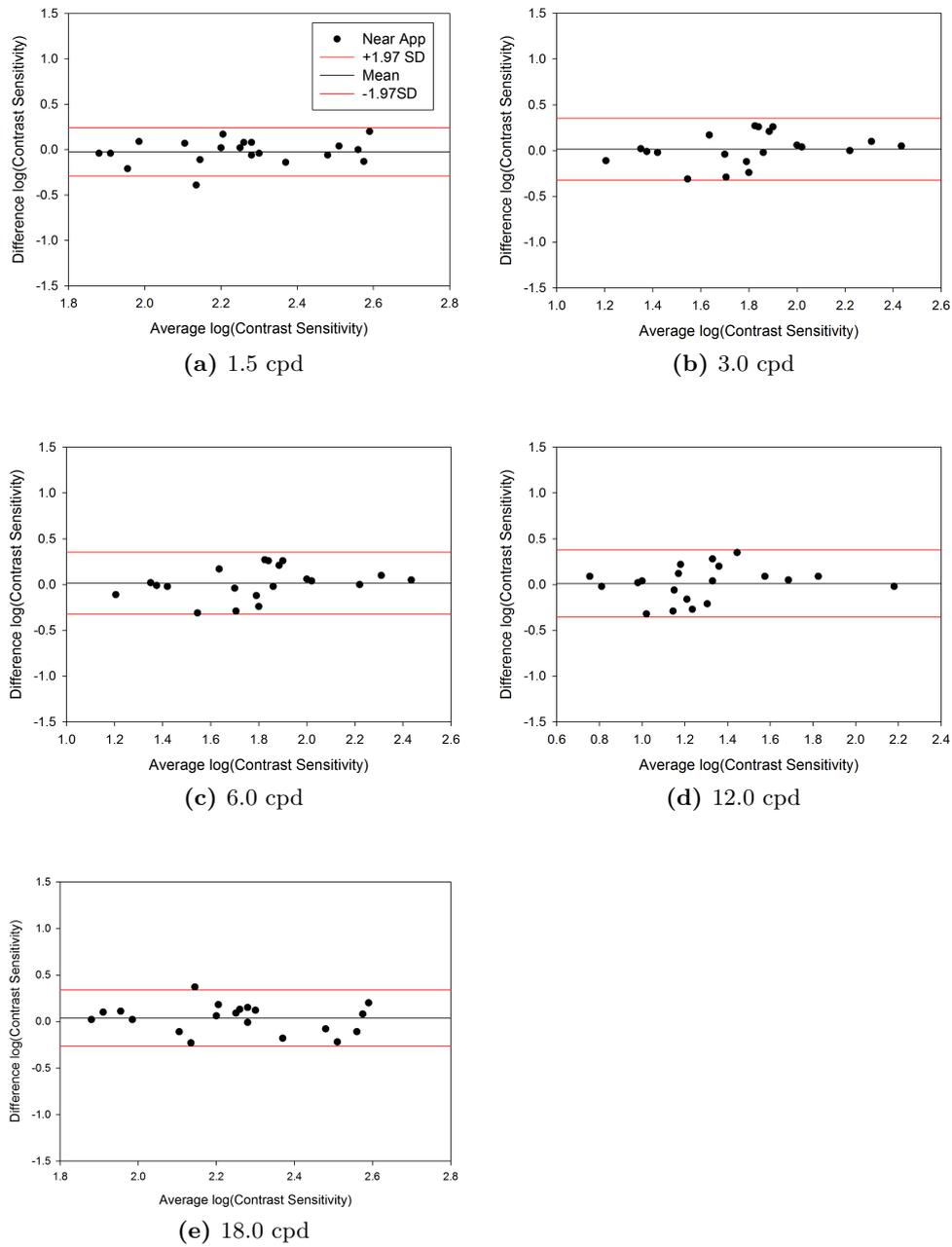


Figure 3.21: Near App Repeatability Bland-Altman Plots, n=20

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

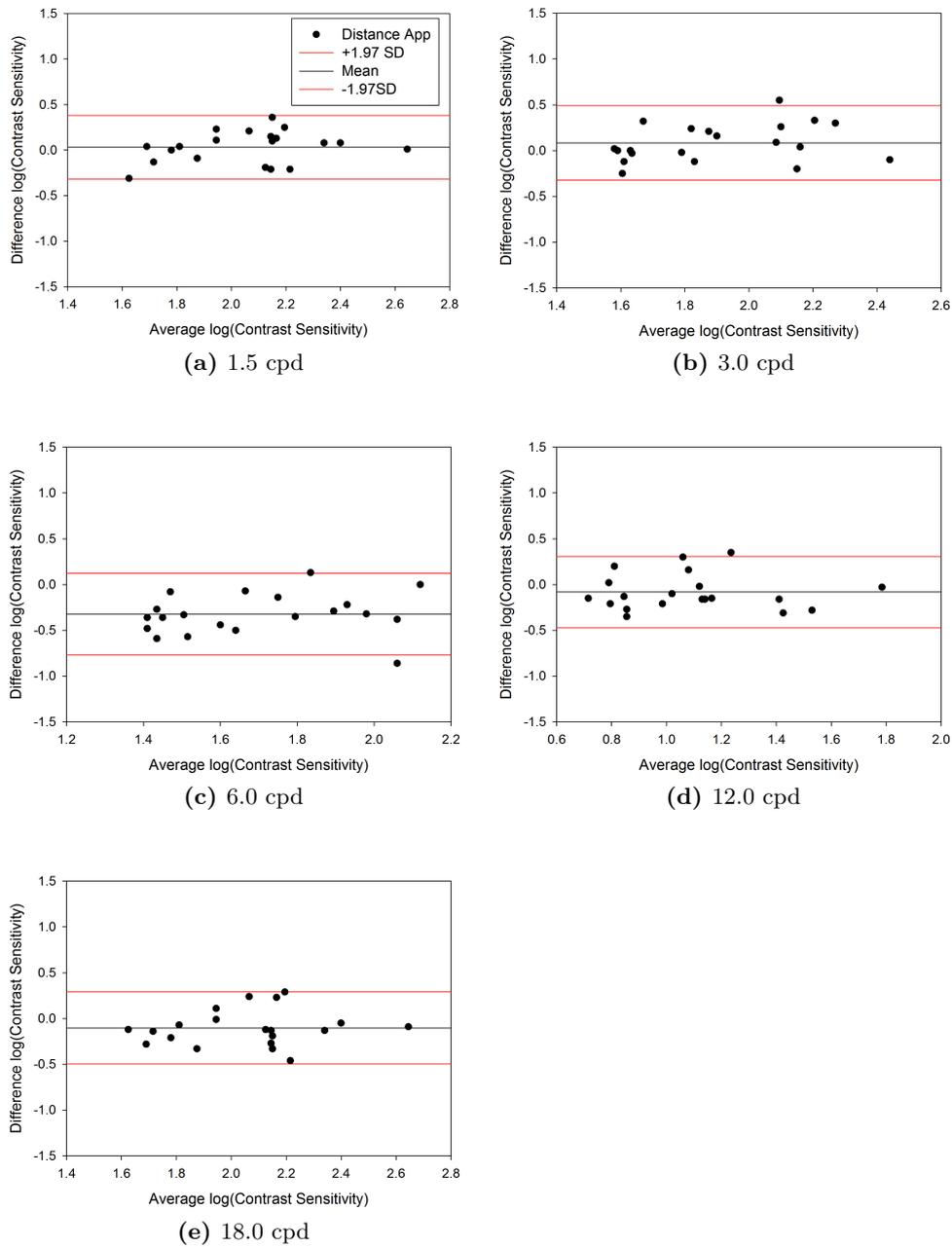


Figure 3.22: Distance App Repeatability Bland-Altman Plots , n=20

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

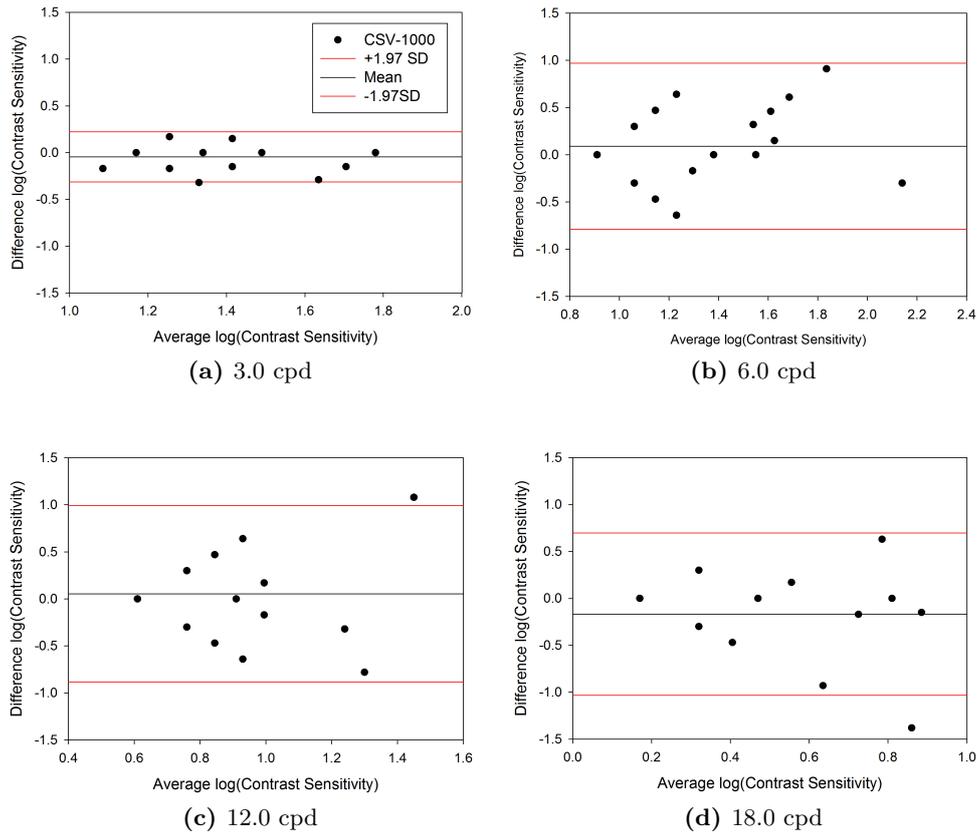


Figure 3.23: CSV-1000 Repeatability Bland-Altman Plots, n=20

3.4.6 Discussion and conclusion

This study aimed to create a novel contrast sensitivity test on a tablet computer that is more reliable than current paper based charts, has reduced testing time in comparison to existing computerised setups, and is portable enough to be used on the examination chair. The chart was developed for the iPad due to their built in battery and light weight, making them portable, the standardization of their monitors and the resolution possible with the newer versions. Despite the limited greyscales (256 levels) from the 8 bit system, an innovative technique of bit stealing (subsection 3.3.2) allowed the perception of a far greater range of contrasts to improve the sensitivity of the test. The test took an old concept of displaying a sinusoidal grating changing from low to high spatial frequency in the X axis and from low to high contrast in the Y axis [Campbell and Robson, 1968], but enabled tracing of the users contrast detection, hence ensuring rapid assessment at all (rather than selected) contrasts and spatial frequencies.

Studying the mean contrast results for all patients in Figure 3.19, it is clear that there are differences between the app and the paper based charts when it comes to the shape of the contrast sensitivity function when no occluding foils are present. While the CSV-1000 paper chart displays a definitive peak threshold at 6 cpd, the app charts shows similar sensitivity at 1.5 and 3.0cpd before thresholds reduce at higher spatial frequencies Dorr et al. [2013] found a similar shape to these app results for the contrast sensitivity function determined using a 2-AFC procedure on both the iPad and a traditional CRT laboratory setup, with no “peaks” seen but with a maximum contrast sensitivity around 1 cpd. Moreover, previous studies using research laboratory grade CRT setups on young adults found maximum contrast sensitivities to occur between 2-5 cpd [Ross et al., 1985, Derefeldt et al., 2009, Costa and Nogueira, 2013] which is in agreement with the app charts but not with the paper based chart. Artifacts or visual cues present within the CSV-1000 printed chart may be biasing patients responses and may explain why peak contrast sensitivity appears to be shifted towards the higher frequencies.

Few studies have measured distance and near contrast sensitivity in the same patients [Knorz and Claessens, 1993, Montés-Micó and Alió, 2003] and none of these used a chart of similar design allowing comparison of the thresholds across the range of spatial frequencies, whether there are displaced and/or symmetrical. Both the near and distance app found higher mean contrast thresholds at all frequencies in comparison to that of the CSV-1000 and the Pelli-Robson charts. In addition mean

thresholds were higher than those found by Dorr et al. [2013] who used a 2-AFC procedure on an iPad based contrast sensitivity test. These higher threshold values seen with the app charts may be a result of using the method of adjustment (patients were asked to find the height of the vertical bars). This is analogous with an adjustment based descending trial which has been shown to increase threshold measures in psychophysical experiments [Laming, 2013] and more specifically in contrast sensitivity testing [Higgins et al., 1988] compared to 2AFC methods.

As would be expected, chart type, spatial frequency and occlusion condition had a significant effect on contrast sensitivity thresholds. Significant interactions between spatial frequency and occlusion condition were unsurprising as Bangerter foils had been previously characterized by Pérez et al. [2010], who found that they reduced contrast more heavily at the higher spatial frequencies. No interactions were found between the chart type and occlusion, furthermore, when the data was analysed at each spatial frequency, thresholds at 1.5, 3, 6, and 12cpd showed no interaction between chart type and occlusion, with the only exception being 18 cpd. It appears that the effect occlusion foils have on the final threshold measurement does not depend on which chart was used, in general the occlusion condition affects all charts equally supporting the validity of the new app based charts.

As shown in Table 3.5 and Table 3.6, correlation between any of the contrast tests could be considered poor. Poor correlation between the sinusoidal type charts and the Pelli-Robson letter chart is unsurprising; McAnany and Alexander [2006] found differences in the shape of letter and gratings based CSF's and Herse and Bedell [1989] found that manipulating both gabor patches and letters in the same way produced different results. Even the near and distance apps, whose methods are broadly similar in their application, did not show highly correlated results. While the reason for this is unknown, it may be due to the limited age range of the patients within this study; there may not be a sufficiently wide spread of people with varying contrast sensitivity for correlations to be seen. Alternatively, the difference may be due to the “distractors”, the little red markers on the distance app that may be impeding contrast sensitivity testing.

This study was limited as there was a lack of a gold standard from which to judge all of the contrast tests. The gold standard should of been a CRT based contrast grating test however the equipment was unavailable for this study. Further studies that use the CRT gold standard could be used to determine whether the difference between distance and near contrast sensitivity at multiple spatial frequencies could

be diagnostic.

Within-session repeatability was poor for the CSV-1000 test at 6, 12 and 18 cpd, as the Bland Altman 95% confidence intervals were large (range: 0.76-0.93 logCS) and the correlation coefficients low ($r=-0.01-0.29$). These results are in agreement with Reeves et al. [1991] who tested the within session repeatability of a similar gratings chart (VCTS 6500) and found 95% confidence intervals of between 0.95-1.01 log CS. Gratings based contrast sensitivity charts have continually shown poor repeatability [Franco et al., 2010, Kelly et al., 2012], which is most likely the result of limited contrast sensitivity steps and the testing procedure; with only two choices for the target orientation patients have a 50% chance that they could guess the location correctly even when they are unable to see the target at all. With such large confidence intervals and low correlation values it can be concluded that the CSV-1000 paper based chart is not ideal to investigate contrast sensitivity within the clinic.

Greater repeatability was found with the distance contrast sensitivity app in comparison to the CSV-1000. Not only were the correlation values strong across all frequencies studied ($r=0.76-0.81$) but the 95% confidence intervals were significantly smaller (0.34-0.44 log CS). Meanwhile the near contrast sensitivity app performed better still with stronger correlations ($r=0.83-0.92$) and smaller 95% confidence intervals (0.26-0.37 logCS). The larger COR value for the distance app in comparison to the near app, may be caused by the distractors on the screen (red markers used to show the patient the contrast they have selected) which are present on the distance but not the near app. Pesudovs and Hazel [2004] suggested that a COR value less than 0.25 would be ideal to pick up subtle and small differences between patients. The new apps repeatability may not be high enough to meet that criteria but presumably it may still have clinical use if it can distinguish between those who are contrast deficient and those who are not. Future testing must be done to test this assumption.

The best within session reliability was with the Pelli-Robson chart with 95% confidence intervals of 0.14 logCS, which is similar to the 0.18 logCS found by Thayaparan et al. [2007]. However, high within session repeatability with the Pelli-Robson chart is a product of the testing method; not only is letter recognition on the Pelli-Robson chart equivalent to a 26 alternate forced choice method but patients could be considered experienced observers as letter recognition tasks are common place. While the Pelli-Robson chart outperforms both the near and distance contrast app in this

3.4 Study-Comparison between mobile apps and traditional clinical contrast sensitivity testing methods

study it is severely limited as it covers only one spatial frequency.

To conclude this study found both the near and distance app to have greater repeatability than those of the CSV-1000 based grating chart, but less than the Pelli-Robson chart. All of the charts did not correlate well with each other and the lack of a gold standard from which to measure limited the study. The clear advantages of reduced testing time, instant plotting of results and wide testing range of frequencies may make the near and distance app a suitable tool for evaluating the contrast sensitivity function of patients within the clinic.

4 The Development Of A Computerised, Mobile Test For Subjective Amplitude of Accommodation

4.1 Introduction

Accommodation is the ability of the eye to change the refractive power of the crystalline lens in order to focus objects at various distances on to the retina [Charman, 2008]. The maximum accommodative ability of the eye to focus between far and near distances is called the amplitude of accommodation (AOA). It is a measure between the far point of accommodation, when the crystalline lens is at minimum refractive strength, to the near point of accommodation where the lens is at its maximum refractive strength.

Figure 4.1a and Figure 4.1b describes the most widely held accommodative theory as proposed by Helmholtz [1865] and its relation to the near and far points of accommodation. It can be noted that moving an object beyond the near point, further towards the eye, results in blur as the crystalline lens can no longer focus light onto the retina. Like many biological systems, the eye's ability to accommodate does not remain fixed over time but slowly deteriorates with age. This is termed "Presbyopia", and is usually noticed around the age of 40 years where the inability to focus at near distances becomes problematic. Presbyopia will continue until no objective accommodation remains [Gilmartin, 1995]. A healthy young patient may view objects as near as 6cm clearly; meanwhile older presbyopic patients may only view objects as near as 50-100cm before refractive blurring occurs [Sterner, 2004, Sterner et al., 2004].

Hofstetter [1950] created a widely used formula that can be used to estimate the normal accommodative amplitudes of patients of various ages. In this way AOA measures can be used to discriminate between natural progression of presbyopia and

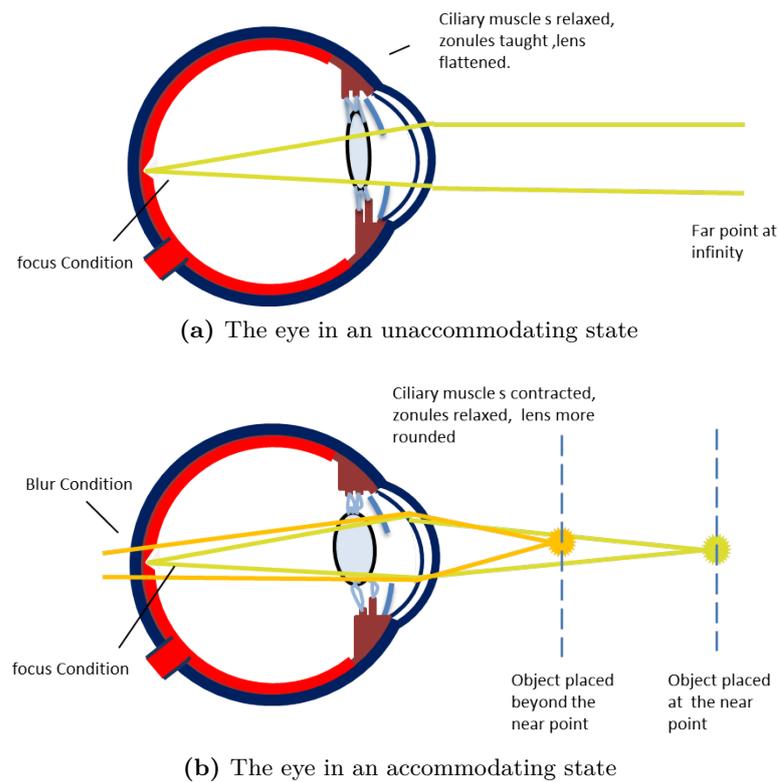


Figure 4.1: Near and far points of accommodation

ocular conditions of a more serious nature. Reduced accommodative amplitudes have been associated with accommodative insufficiency [Bartuccio and Taub, 2008], medicative side effects [Gilmartin, 1987], diabetes [Braun et al., 1995], ocular trauma, amongst various other systemic and ocular diseases [Elliott, 2013b]. As such AOA testing forms a major part of the battery of tests required to detect accommodative dysfunction (along with lead/lag, facility and cycloplegic refraction [Wick and Hall, 1987]). Hence it is important that any clinical measurement of amplitude of accommodation is precise, repeatable and accurate. This chapter explores the existing objective and subjective methods for determining AOA and proposes a new mobile test for amplitude of accommodation measures.

4.1.1 Objective AOA

Auto refractors [e.g Wolffsohn et al., 2011], aberrometers [e.g. Krueger et al., 2001], and optometers can all be used to take objective measurements of the eye's optical

power. By assessing the changes of optical power between open field conditions (so as not to induce proximal accommodation), and then with accommodative targets placed at various distances, the AOA can be deduced.

The main advantage of objective AOA measurement is that it can accurately determine true accommodation from pseudo accommodation caused as a result of pupil size [Wold et al., 2003] or a patients subjective assessment of blur. In effect it negates perceptual and depth of field effects associated with subjective methods from the change in optical power seen with accommodation. For this purpose objective AOA measurements have been repeatedly used in research to assess accommodating inter-ocular lenses (IOL's) post surgery to understand their efficacy at changing dioptric power [Wolffsohn et al., 2006, Nemeth et al., 2008, Dogru et al., 2005].

Objective AOA measurement within the clinic is not always used. In part, because it does not characterise the full effect of near vision, where depth of field effects, accommodation and perceptual factors all interact to create near vision [Wold et al., 2003]. In addition, the equipment needed is often expensive and bulky or objective techniques are time consuming in comparison to the relative simplicity of subjective methods. Even taking measurements may not be straightforward; auto-refractors can struggle to work with patients who have small pupils. For example, the typical minimum pupil diameters specified by one manufacturer (Shin-Nippon SRW-5000) is 2.9mm [Mallen et al., 2001] which can cause problems for those with senile miosis. Even when objective measures of the optical system and its aberrations are taken they may not necessarily correlate to the patient's own visual satisfaction as well as subjective measures.

4.1.2 Subjective AOA

Subjective AOA methods are commonly used in clinical practice. Unlike objective measures that measure an eyes optical power directly, subjective measures rely on a patient's perceptual tolerance to blur to infer a change in optical power. Figure 4.1b shows how moving an object past the near point of accommodation can result in a perceived blur on the retina. This blur is the visual clue in many subjective tests for AOA to determine the near point of accommodation. Conveniently, if the far point is at infinity, which is the case for emmetropes and those with optical correction for far distance, only the measurement of the near point is needed to determine AOA[Elliott, 2013b].

The inference that passing beyond the near point of accommodation causes noticeable optical blur (in the form of a larger circle of confusion) is not entirely accurate. Apparent depth of focus at near distances is not just regulated by the accommodative elements of the eye, but also by the pupil acting as an aperture stop [Wang and Ciuffreda, 2006] (see Figure 4.2), multi-focality of corrective lenses or the cornea [Fukuyama et al., 1999], myopia [Elder et al., 1996], higher order aberrations [Wolffsohn et al., 2010] and a subjects own tolerance of blur [Rosenfield and Cohen, 1995b]. Mistaking depth of field and perceptual effects for accommodative ability has been termed “pseudo” or “apparent” accommodation by researchers [Nakazawa and Ohtsuki, 1984, Chen, 2012]. For these reasons subjective methods can over estimate accommodative amplitude by as much as 2.0-4.0 D [Ostrin and Glasser, 2004]. Despite these disadvantages, subjective methods of determining AOA remain widespread in practice; it remains clinically significant to understand a patients “real world” vision where real and pseudo accommodation interacts with subjective accommodation measures [Atchison et al., 1994].

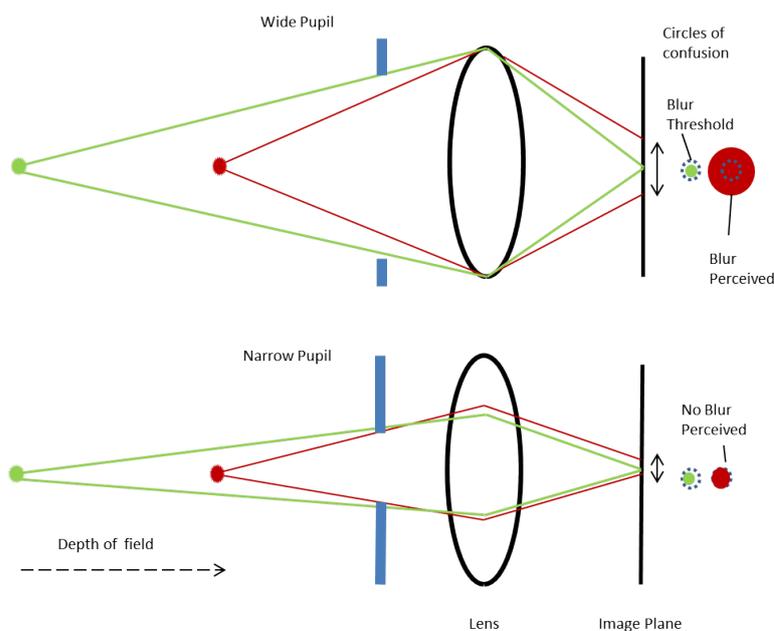


Figure 4.2: Pseudo accommodation caused by pupillary constriction

Two main methods exist to test subjective AOA clinically: minus to blur [Sheard, 1920] and the subjective push-up/pull-down test [Donders and Moore, 1864]. The minus to blur technique is performed monocularly with a target located at 40cm. Additional lenses are added in -0.25 D steps until the patient reports blur. The

value of the lens addition at the point of blur plus the + 2.5D for working distance compensation results in an accommodative amplitude for that eye.

Using the minus lens procedure partly negates any effects of a static targets increasing retinal size on perceptual blur that occurs when using the push-up and pull-down method, although image magnification/minification effects still occur. It has been speculated that this is main reason for lower AOA amplitudes and an increased repeatability in comparison to the push-up/pull-down techniques [Antona et al., 2009]. Additionally, minus to blur is in better agreement with objective measures [Ostrin and Glasser, 2004], see Figure 4.3. As minus to blur can more likely mimic objective measures, it appears that maintaining a targets size on the retina may have greater impact on AOA than pupil size effects. Despite positive characteristics, the test can only be performed monocularly and can take a longer time to perform correctly than the push-up technique. So to use minus to blur as a quick yardstick to measure functional near vision may be less practical than that of the push-up technique.



Figure 4.3: Mean accommodative amplitudes for subjective and objective methods taken from [Ostrin and Glasser, 2004]



Figure 4.4: R.A.F. near point rule

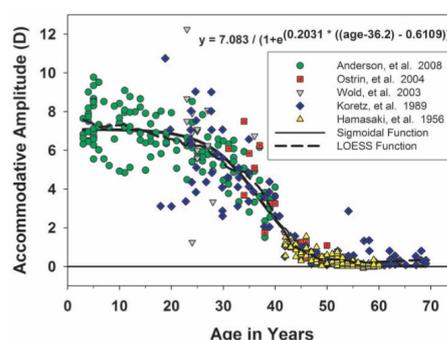
Despite its limitations, the subjective push-up test [Donders and Moore, 1864] remains one of the most popular methods for assessing AOA in practice [Scheiman

and Wick, 2008]. In this method, a target, usually consisting of fixed letters or reading material, is placed at a far distance to the patient (typically over 50cm) on a sliding rule. The target is then slowly brought closer to the eyes until blur is reported by the patient. If a patient's near point is either too far or too close for the rule to measure, additional positive and negative lenses may be used to extend the testing range. The distance on the rule is recorded and then expressed in dioptres to determine the amplitude of accommodation (minus any additional lens power added to extend the testing range).

A variation on this test is called the push down test whereby the target is placed very close to the eyes and pulled back until clearance is reported. Higher values of AOA have been found to occur with the push-up technique when compared with the push down method [Wold, 1967, Fitch, 1971, Antona et al., 2009]. Antona et al. [2009] attributed this to the differing endpoints of each test; push-up measuring the first blur and the push down measuring the first clear point. It appears patients may be better at establishing when an object has passed from blurred to clear, than from clear to blurred. It has also been suggested that the push-up test causes an enhanced proximally induced accommodative response as the target approaches the subject [Rosenfield and Gilmartin, 1990] that is absent with the pull-down and minus to blur techniques [Rosenfield and Cohen, 1996].

As eluded to previously, the actual measurement of AOA using this method is compromised in several regards. First and foremost, angular sub tense of the target does not remained fixed as the target is moved down the rule. A 6/6 target at 50cm becomes a 6/12 target at 25cm. This has severe implications for a patients tolerance to blur. Both Atchison et al. [1994] and Rosenfield and Cohen [1995a] found subjective AOA measurements increased with target size. It has been theorised by Somers and Ford [1983] that when viewing larger targets the borders between the various elements of the target are further apart and there is less overlap of their blur distributions. Thus as the target's angular sub tense increases whilst it travels down the rule there may be a delay in the patient first appreciating the presence of blur. Hence the over estimation of AOA measurements with the push-up method when compared to the fixed retinal size target of the minus to blur method [Antona et al., 2009, Ostrin and Glasser, 2004]. To counter this, a method of instructing the patient to view differing size letters on the near point rule as it advances towards them, can maintain a constant visual angle. This was shown to reduce subjective AOA measurements by 1.4 D [Atchison et al., 1994]. However it requires the exam-

iner to instruct the patient what line to read at a certain distance. Atchison et al. [1997] found that greater sensitivity to defocus is found with letters that are close to the acuity limit and thus recommended using the smallest possible letter that can be seen to assess AOA using the near point rule. Although this presents a problem, with a series of fixed targets that can be seen on instruments such as the RAF rule (see Figure 4.4), even when the smallest line is selected it cannot match the near visual acuity that can be experienced by young patients.



(a) Subjective AOA and Age taken from [Duane, 1922] (b) Objective AOA and age taken from [Anderson et al., 2008]. Data pooled from: [Ostrin and Glasser, 2004, Wold et al., 2003, Anderson et al., 2008, Hamasaki et al., 1956, Koretz et al., 1989]

Figure 4.5: Comparison of objective and subjective measures of AOA and age.

The near triad of responses when viewing a near object, especially pupillary constriction has shown to affect push-up amplitudes [Wang and Ciuffreda, 2006]. Fukuyama et al. [1999] found that pupillary diameter had a significant positive correlation with push-up amplitudes in pseudophakic eyes. As the pupils constrict when the target becomes close, the depth of field of a patient increases [Atchison et al., 1997] (see Figure 4.2 for explanation). The push-up test can only measure apparent accommodation and cannot differentiate between depth of field (DOF) effects and that caused by a dioptric change in the power of the crystalline lens. However this is not to say that the test produces fallacious results; the main component of pseudo accommodation in the young is objective accommodation. Well cited work by Duane [1922] using the push-up method showed the characteristic accommodative decline

with age that can be attributed to presbyopia, this mimics the curve produced by objective results albeit with higher amplitudes (see Figure 4.5a and Figure 4.5b). Part of the continued popularity of the push-up may even be that it can provide a measure of functional vision for patients where all the elements of near vision interact. The push-up test can be used binocularly in natural viewing conditions where looming targets regularly appear and the proximal cues for accommodation are present [Rosenfield and Gilmartin, 1990].

Minor criticisms of the push-up test also exist. Typically the test is performed clinically with a RAF rule, with a manual target the test relies on constant lighting of the test room to illuminate the target. An over-illuminated room will constrict the pupils and increase DOF affects in measurement. Pupil size changes aside, it has been found that a reduction in luminance or contrast has little affect on accommodative response or DOF [Tucker and Charman, 1986]. Secondly, near point rules require that one end of the rule rests intrusively on the patients cheeks. While only moderately uncomfortable it still is an undesirable trait. Lastly, near point rules have been used for measurement of near point convergence where it has been found the rule gives higher readings than free space methods [Adler et al., 2007]. Adler et al. [2007] attributed this to the conflicting cues to the vergence response from the rule acting as reference point in space from which the target can be judged. Presumably, as the near triad is neurally linked these conflicting cues may also affect accommodation measurements.



Figure 4.6: Compact accommodometer taken from [Ide et al., 2012]

Ide et al. [2012] attempted to address concerns with the near point rule by creating a new instrument see Figure 4.6. The compact accommodometer consisted of a fixed static paper target attached to a box that included an ultrasonic distance sensor and digital read out. Patients were required to test themselves by moving the

accommodometer closer until blur was detected. The study found that the compact accommodometer tended to over estimate amplitudes, but that it still showed the decline of accommodation with age. The designers claimed that it is easier to use and requires less space than other near point measuring devices. However it did have one severe limitation; it could only measure between 20-60 cm. Extra lenses were used on younger patients to measure their respective near points. Another accommodometer instrument called the D'ACOMO [World Optical Corp, Kyoto, Japan] has been used in several papers to measure subjective AOA [Kamiya et al., 2008, Ide et al., 2012]. Here the target is moved along a motorised linear actuator so that the speed of approach can be controlled. However, there is no attempt to control angular sub-tense of a target.

4.1.3 Background summary

Addressing the concerns with the push-up test whilst maintaining the near rule's positive attributes remains a challenge. Any subjective method would have to be quick to administer, require little skill or training for the examiner, measure functional vision and provide adequate AOA readings. Considering the current limitations of the near point rule, and the previous attempts to address concerns with AOA measurement, a new instrument could be created to measure subjective AOA. A hand-held electronic device based on a smart phone would make the ideal candidate. Not only are smart-phones portable, require little space and have ample computing power, many clinicians will already be familiar with their use. Multiple potential usage will also decrease the cost of the device. Distance feedback from a sensor could provide enough information to dynamically resize targets on the mobile's high definition display. Thus a new practical device may enhance the clinical measurement of subjective AOA.

4.2 Aims and objectives

4.2.1 Aim

The primary aim of this project was to create and evaluate a novel, hand-held, mobile accommodometer to measure subjective amplitude of accommodation. The secondary aim of this project was to assess the efficacy of dynamically resizing targets for AOA measurement.

4.2.2 Objectives

1. Selection of a suitable mobile platform and any electronics or sensors that can be used for distance measurement.
2. Design and fabricate the AOA device.
3. Develop a computer application that will interact between the mobile device, external hardware and the patient.
4. Create and implement experimental methodology to test the new device in comparison to the near point rule.

4.3 Development-Aston Accommodometer

Due to the popularity of the iPhone [Apple inc, Cupertino, US] and the functionality of its software development kit, the Aston Accommodometer was developed on Apple's iOS platform. Many of the clinicians and subjects who would use the device would already be familiar with the look and feel of smart-phone apps. The Aston Accommodometer would be able to make use of the high resolution display's that are built in to the iPhone to present a dynamically resizing target.

4.3.1 Distance measurement

Measurements of the accommodative near point require a distance reading to be made. Three methods of using the inbuilt sensors were explored: pupillary distance estimation, reference object estimation and dead reckoning, but were deemed unsuitable.

Pupillary distance estimation, as described and used in subsection 2.5.1.3 for the Aston-Read reading test, is a method whereby the distance between the patients pupils are tracked using image processing algorithms. By assuming that a patient's pupillary distance (PD) is a certain size and knowing the optics of the camera being used you can estimate the horizontal distance from the eyes. However this method is only an approximation; the mean pupillary distance for male patients is $64.6 \pm 3.7 \text{ mm}$ [Dodgson, 2004]. Calculations that include the standard deviation of PD, as shown in Figure 2.21a, indicate that at 25cm the measurement may be as much as $\pm 2.5 \text{ cm}$ from the correct distance. This even assumes that the image tracking algorithms are 100% accurate and the camera is always in the same horizontal and vertical plane as the centre of the two pupils, so inaccuracy may be much higher than this model suggests. Using a reference object of known size and shape in the field of view of the camera could be used to calculate the distance in lieu of tracking pupillary distance. Image processing could determine the size of the reference object on the image sensor plane and relate that to the objects known dimensions to calculate the horizontal displacement. This will reduce the margin for error that is inherent with the pupillary distance estimation, however it would require a reference object to be placed in the same plane as the eyes. The inconvenience of having to place a reference object close to the eyes coupled with the reduced response times and accuracy of real time image processing for distance measurement ruled out any camera based method.

Inbuilt accelerometers and gyroscopes in theory could be used to produce a dead reckoning device for displacement. At the most simplest level double integration of linear acceleration data would result in a displacement value. Due to the double integration, any noise or bias in the system can have a dramatic effect on the estimated displacement. Commercial accelerometers and gyroscopes as used within the iPhone neither have the accuracy nor the sample rate to reduce drift to an acceptable level, but they may still be useful for orientation of the device.

An external ultrasonic ranging sensor was therefore selected for the device. The manufacturer selected [Devantech, Attleborough, UK] quoted an accuracy of $\pm 1\text{cm}$, a ranging time of 65ms and a larger range (3cm-3m) than the sensor used by Ide et al. [2012]. For a dynamically resizing target the responsiveness, accuracy and range made the ultrasonic sensor the preferable option.

4.3.1.1 Ultrasonic distance sensor calibration

A test rig was set up using the ultrasonic sensor and a model human face to determine the accuracy of ultrasonic distance measurement (see Figure 4.7). Actual displacement was measured using a membrane potentiometer attached to a linear slide. A wiper would sweep down the potentiometer and the resulting voltage signal was fed into a micro-controller [Arduino, Ivrea, Italy] that would display the distance on an electronic LCD display. A Devantech SRF08 ultrasonic sensor was mounted on top of the wiper and ultrasonic distance data was fed into a laptop for recording. Above the ultrasonic sensor a laser pointer was attached, the reflection of the laser from the model head was used to place the linear rule parallel to the model face. Ultrasonic values were recorded at 1cm intervals from 9-40cm for both a flat acrylic plate and model human head representing accommodative values of between 2.5D-11.1D. The test was repeated three times, the results can be seen in Figure 4.8.

Figure 4.8a demonstrates that accurate measures can be taken with the sensor when used with the flat plate. Meanwhile in Figure 4.8b the measures do not follow the ideal measurement but tend to deviate from the ideal measurement more widely. Figure 4.8c and Figure 4.8d show the plots of measurement error over distance for the plate and head respectively. The mean error was $-0.08 \pm 0.29\text{ cm}$ for the flat plate and $0.09 \pm 0.89\text{ cm}$ for the model head between the three test runs.

Figure 4.8a can be explained in terms of the beam width and changing attenuation of the sensor hitting various elements of the face as it is being brought in or pulled

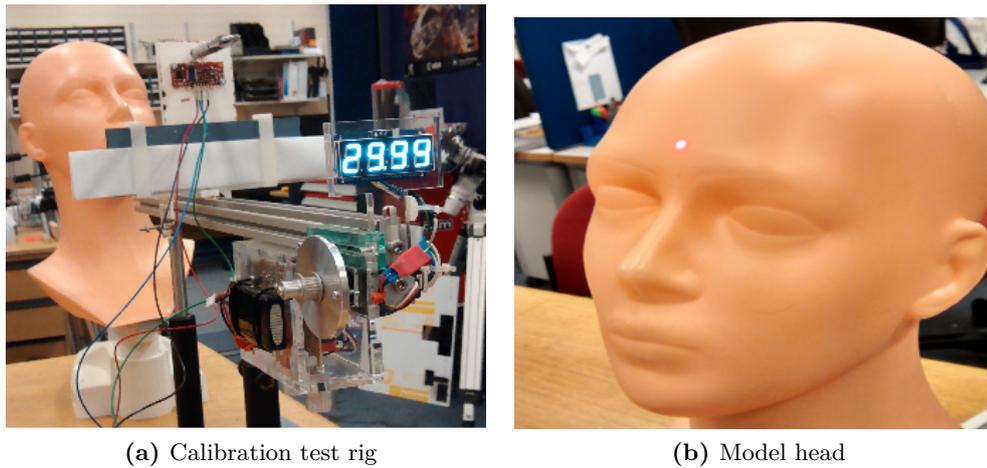


Figure 4.7: Calibration test rig and model head

away. All repeated measures for the model head had this unique displacement profile. For example, with a beam width of 60° at 5cm, a 27cm^2 area of the forehead may be covered, once moved out to 10cm this would increase to 105cm^2 which will incorporate other facial features such as the nose and eyes. These may be detected despite the sensor still being centred on the forehead. Additionally, due to the varying topography of the face this can also affect the attenuation of any signals received by the transducer. Figure 4.8c indicates that the ultrasonic distance sensor can be accurate at distance measurement, however controlling where the measurement is taken from may be more difficult.

The greater fluctuation from the ideal measurement may also explain the higher standard deviations seen with the head compared to the flat plate. If the distance was out by even a small amount between tests, the attenuation and acoustic properties could change dramatically on the face in comparison to a flat surface and hence the higher standard deviations observed between tests.

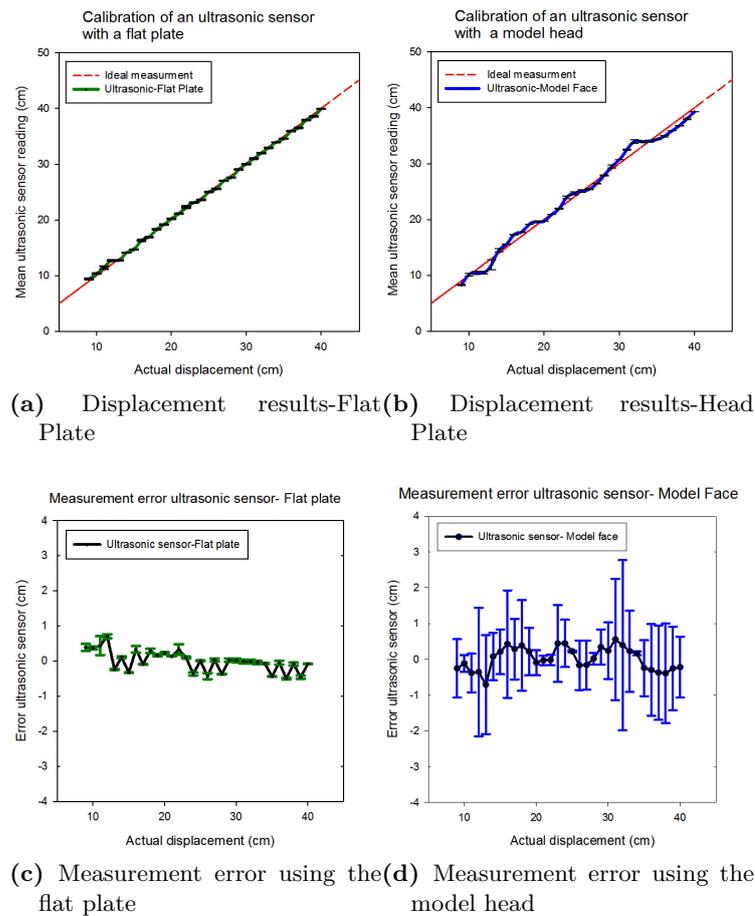


Figure 4.8: Calibration results

4.3.2 Concept

A diagram showing the operating principles of the Aston Accommodometer can be seen in Figure 4.9. The mobile device is placed at arms length then slowly brought forward. Feedback from the ultrasonic sensor updates the size of the target on the mobile phone’s display to maintain a constant visual angle with the subject.

It has been theorised by Somers and Ford [1983] that when viewing larger targets the borders between the various elements of the target are further apart and there is less overlap of their blur distributions. Thus as the target’s angular sub tense increases on the retina patients may be more tolerant to blur. It may be ideal to fix the patients tolerance to blur at all working distances through maintaining a constant visual angle. Once the subject notices blurring of the target for AOA, the

touch screen is tapped and the final reading from the ultrasonic sensor is taken as the accommodative near point. To overcome any problems with the tilt of the device affecting measurement readings, accelerometer data will be used to check the orientation of the device. If the app detects it is not being held vertical an error message will appear telling the user to reorientate the device. In order to remain practical any external hardware was to be fit into a custom phone case and powered by the mobile phone itself.

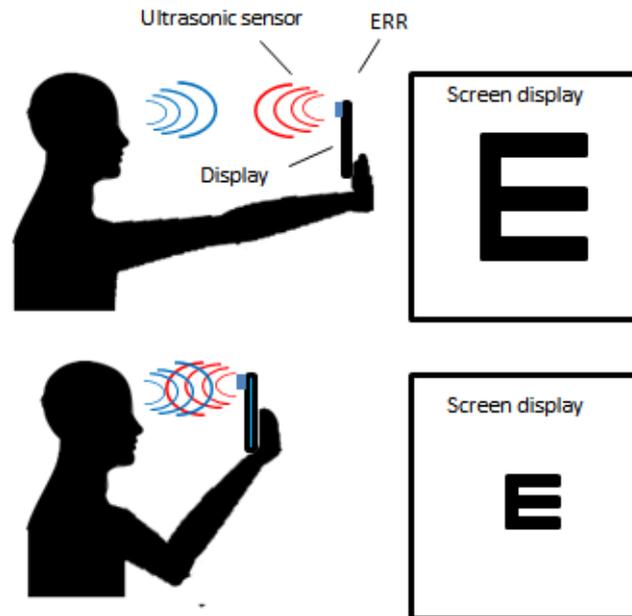


Figure 4.9: Operating principles of the Aston Accommodometer

4.3.3 Hardware

The pinout circuit diagram for the Aston Accommodometer's hardware can be seen in Figure 4.10. The major components of the electronic hardware were: (a) iPhone dock connector, (b) 3.3 to 5V step-up board, (c) Ultrasonic sensor, (d) Arduino development board with ATmega168 [Atmel, San Jose US] microcontroller, (e) Audio modem, (f) Headphone jack. A description of each component's relevant specifications and the rationale behind their selection can be seen in Table 4.1.

Using the internal battery of the iPhone, only 230mA of current at 3.3V could be drawn to power all the components. As the microcontroller and sensor needed 5V and approximately a 100mA, a step-up board was placed in between the battery

and the electronics. The ultrasonic sensor was interfaced to the microcontroller by an I^2C master-slave bus. However, interfacing the resulting distance data to the iPhone was not as simple. Application programmable interfaces (API's) that are used throughout the app development process are locked down when trying to communicate to external sensors/devices unless you are an approved manufacturer of iOS accessories. To work around this restriction an audio modem coupled with a microcontroller was used (first proposed by [Kuo et al., 2010] for interfacing sensors to mobile phones). Using the microphone port as an analog input and the speaker port as an analog output, communication to the device can be established with Frequency Shift Keying (FSK). Within the micro-controller, binary data was encoded at two frequencies $f_1 = 4900 Hz$ and $f_2 = 7350 Hz$ for 0 and 1 respectively, sent through to the iPhone's microphone port by use of an audio modem and demodulated using a software library supplied by [SwitchScience, Tokyo, Japan]. The library used two bandpass filters centred at each frequency to decode the signal back into binary. The encompassing iPhone case was rapidly prototyped on a stereo-lithography machine and the front face was laser cut from black acrylic sheet and cemented onto the case. The finished prototype can be seen in Figure 4.11.

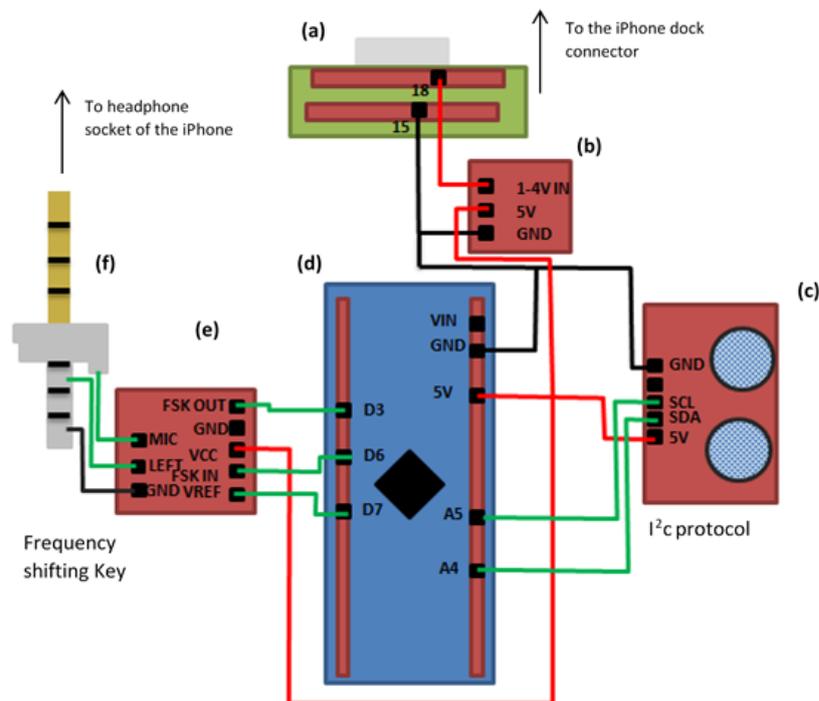
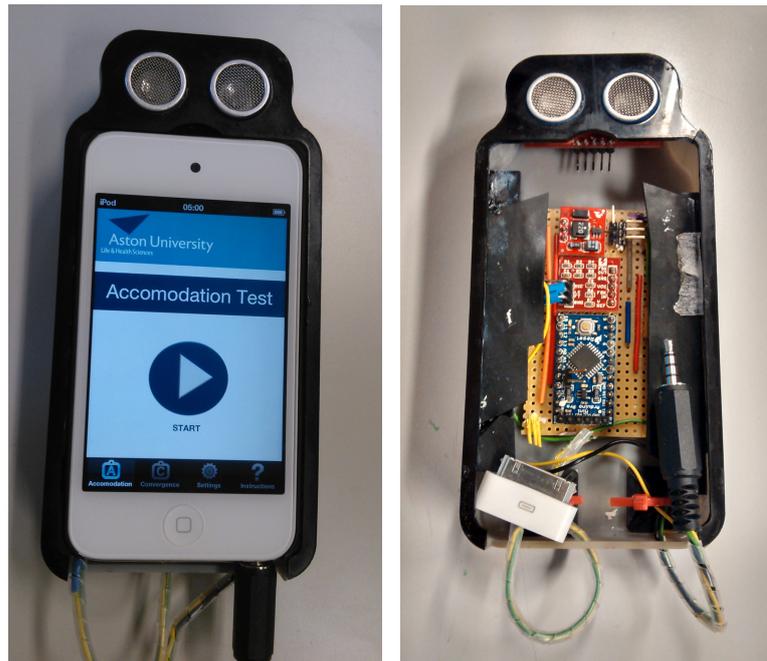


Figure 4.10: Component and pin out diagram of the Aston Accomodometer

Table 4.1: Component specification and rationale

Component	Specifications	Rationale
iPhone connector	<ul style="list-style-type: none">• Standard 30 pin iPod dock.	To deliver power to the hardware. Pins 18 and 15 provide 3.3V supply
NCP1402-5V step-up board	<ul style="list-style-type: none">• 3.3 to 5V DC voltage converter• 200 mA sourcing.	Steps up the voltage to power the ultrasonic sensor and microcontroller.
Deventech SRF08 Ultrasonic sensor	<ul style="list-style-type: none">• 0.03m to 11m range in increments of 0.01 m (manufacturers calibration).• 55 degree beam width at 40kHz.• 65ms ranging time.• 50 mA current consumption.	Within the range and accuracy needed for AOA. The low frequency is able to reflect off soft surfaces like human skin. Quick ranging time allows measurements to be taken at 15Hz.
Arduino Pro mini 5v development board	<ul style="list-style-type: none">• Atmel 168 16 GHz 8bit microcontroller.• 50mA current consumption• 14 digital and 6 analog I/O. I2c and FSK compatible	To take data from the ultrasonic sensor (implementing i2c protocol) convert it to frequency data using FSK which can then be sent through the audio port of the iPhone. Simple and well supported programming environment for the i2c and FSK communication protocols
AudioJack modem	<ul style="list-style-type: none">• N/A	Interfaces the data being sent from the microcontroller to the audio Jack
Audio Jack	<ul style="list-style-type: none">• Standard four pole 3.5 mm audio jack.	Connection to the audio socket of the iPhone.



(a) Aston Accommodometer

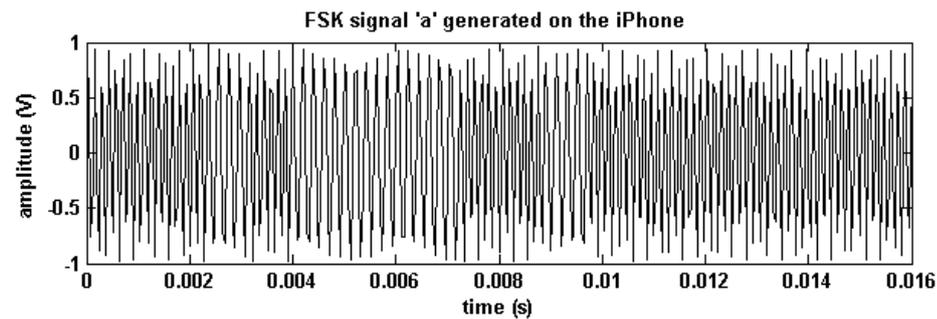
(b) Smart phone case with electronics

Figure 4.11: Aston Accommodometer Prototype

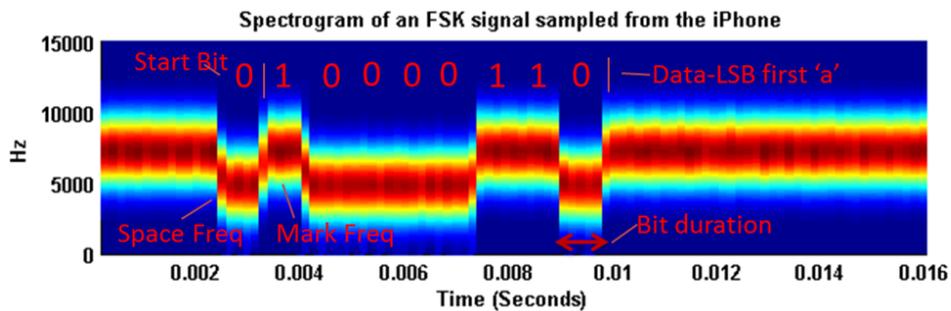
4.3.4 Software

4.3.4.1 Arduino micro-controller program

Interaction between the iPhone and the microcontroller was accomplished by using a simple protocol. The iPhone would request a reading by encoding the relevant binary data into FSK to send over the speaker port and into the micro-controller which would decode the instruction. If the instruction was to request a distance reading, the microcontroller would instruct the ultrasonic sensor to take a reading over the I^2C master-slave bus. Readings typically take at least 65ms to complete so the microcontroller would wait and then request that the completed ultrasonic reading be sent to the microcontroller. Once the reading was received it was encoded into FSK (0 and 1 were encoded at 4900Hz and 7350Hz respectively) and sent out to the microphone port of the iPhone. The programming flow diagram can be seen in Figure 4.13. The program was written in C and used Arduino software libraries. The actual FSK signal as sampled from the iPhone's microphone port can be seen in Figure 4.12.



(a) Raw sampled FSK modulated data



(b) Spectrogram of the FSK modulated data

Figure 4.12: Oscilloscope sampling of the iPhones microphone port for binary data modulated to FSK

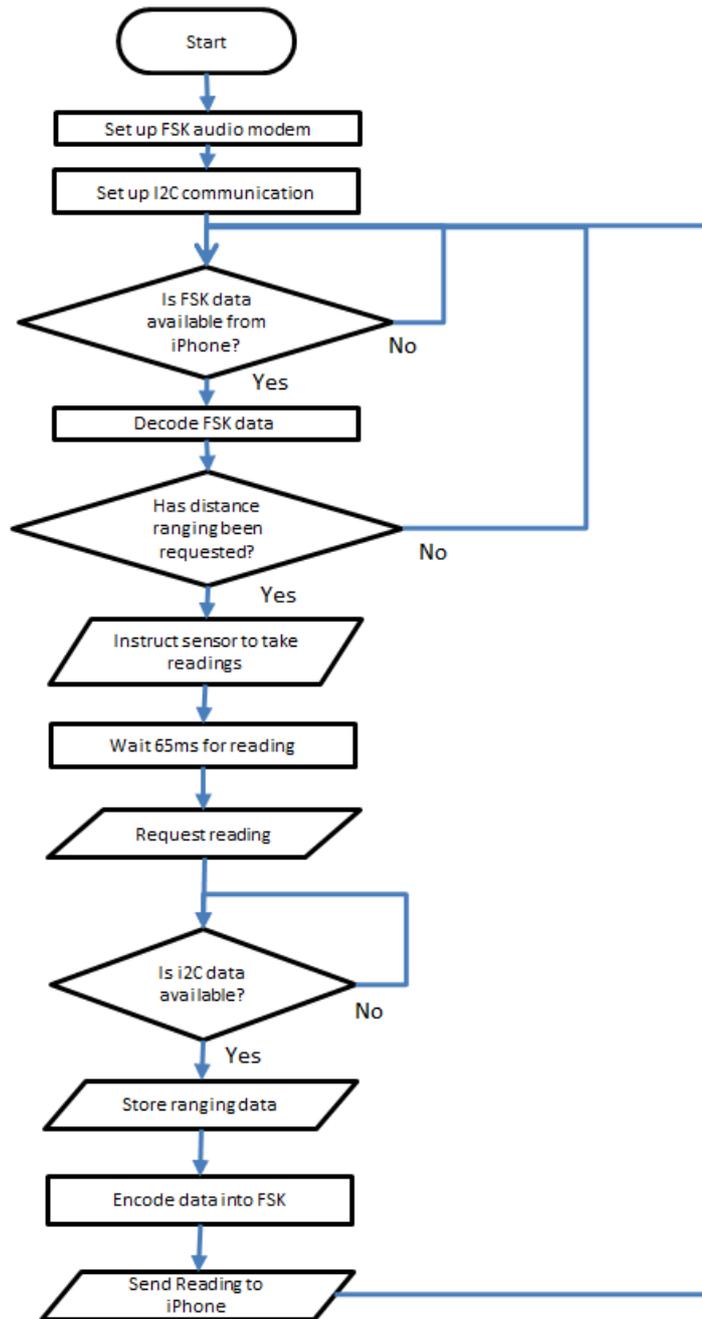


Figure 4.13: Microcontroller flow diagram

4.3.4.2 iPhone application

The purpose of the iPhone app was two-fold. Primarily it was to be used to display the target and change its size in response to the distance readings, but its secondary purpose was to act as the human interface device and react when the user had decided that the target had blurred or if the target was being held incorrectly. The major advantages of using an app to interact with a user are that its simple design, touch interface and ease of use would enable subjects to take their own readings with this device and the resulting measures could be rapidly displayed back to the user. Figure 4.14 displays the initial user interface concept. The main programming flow diagram can be seen in Figure 4.15.

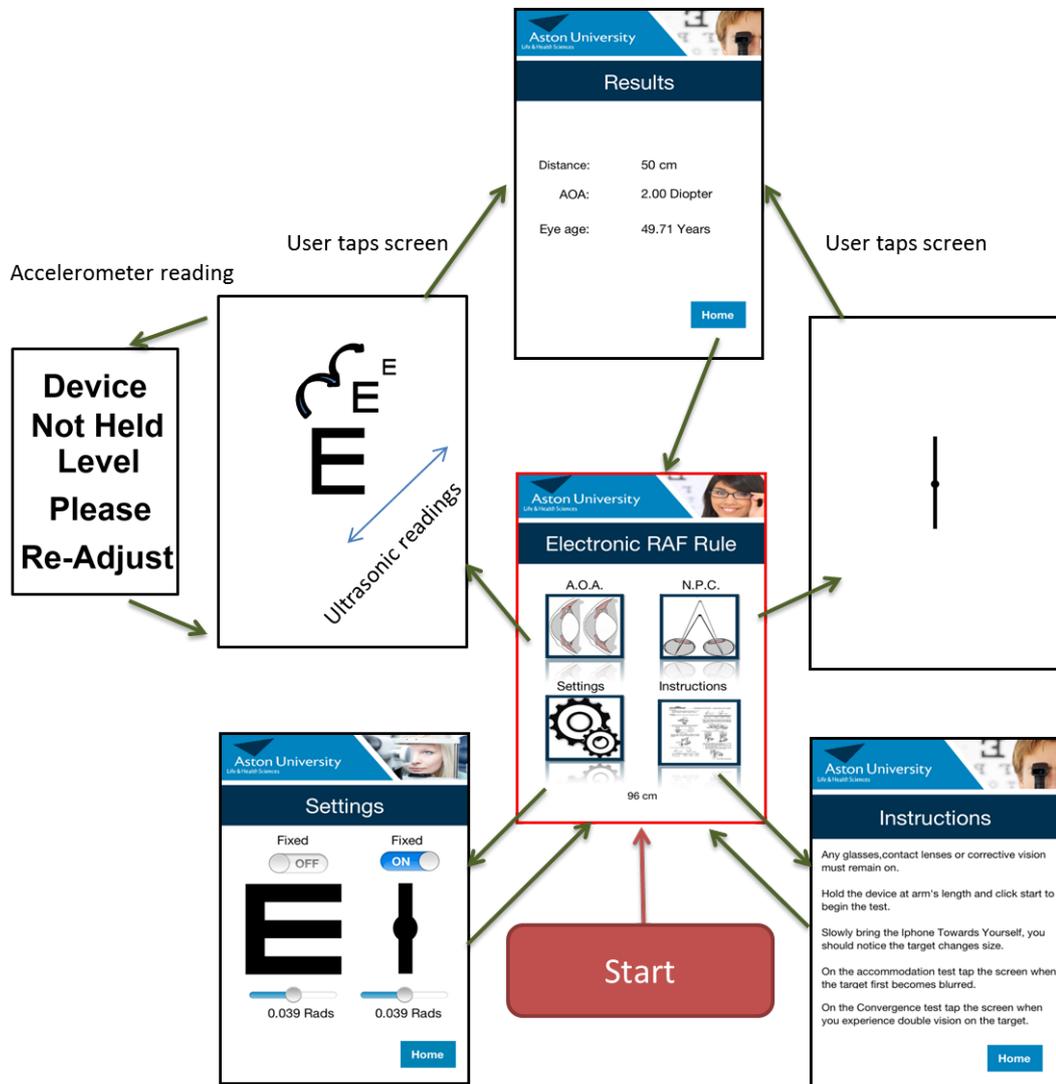


Figure 4.14: User interface for the Aston Accommodometer

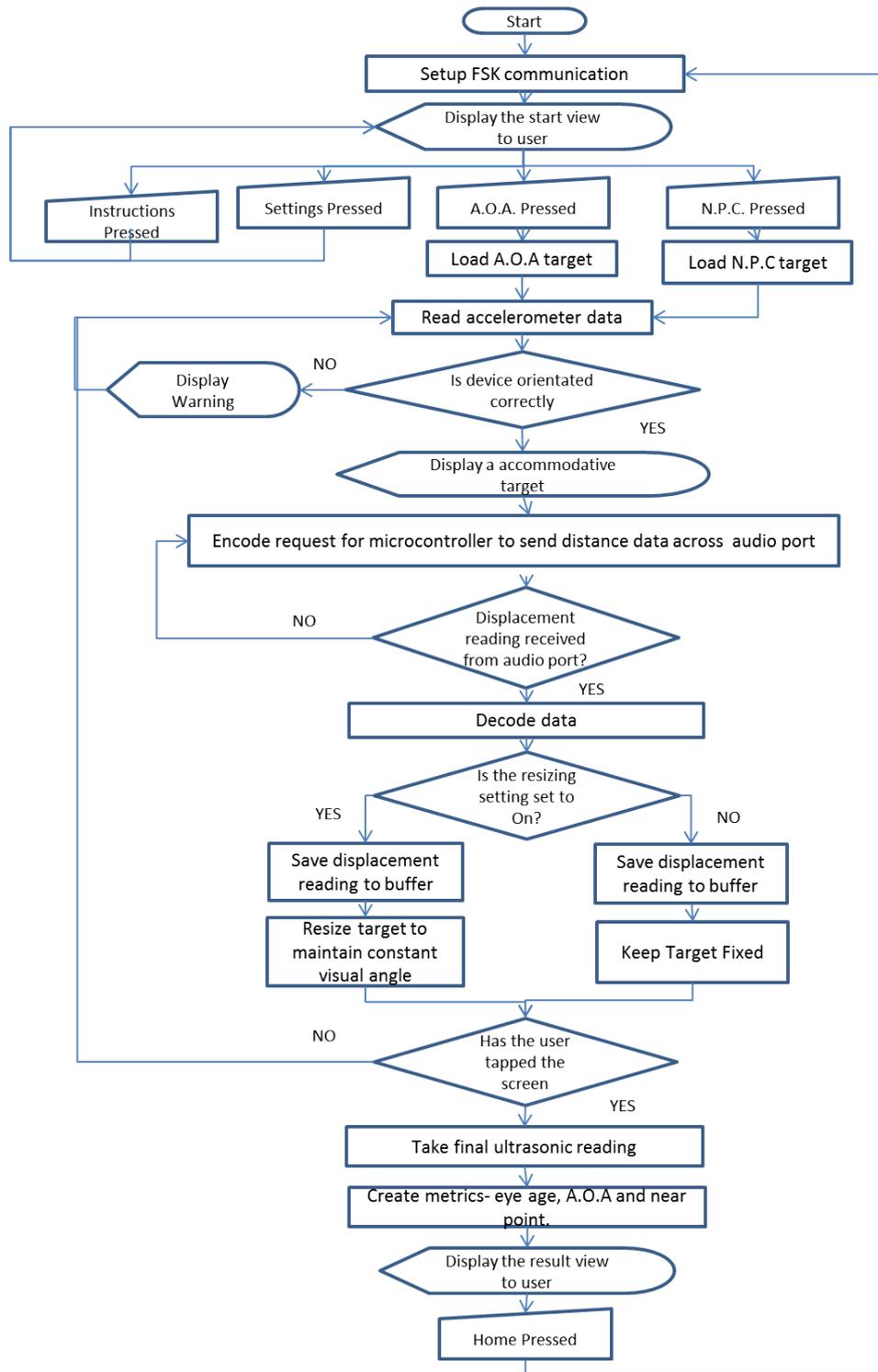


Figure 4.15: iPhone programming flow diagram

4.4 Study - Comparison between the Aston Accommodometer and the RAF rule

4.4.1 Aim

The main aim of this study was to compare the Aston Accommodometer with the traditional near point RAF rule for subjective amplitude of accommodation measurements for the push-up and pull-down methods. The secondary aim was to compare the amplitude of accommodation measurements when using either a resizing target or a fixed target on the Aston Accommodometer. This would show whether the method of distance measurement or the resizing target affects subjective amplitude of accommodation measures.

4.4.2 Subjects

Thirty four subjects with a mean age of 23 ± 3.9 years were recruited for this study from the optometry course at Aston University. Subjects were asked whether they had any form of ocular pathology, abnormality or corrective surgery and were excluded on this basis. All of those that participated were best distance visually corrected to at least 6/6 and could see the target on the RAF rule at 50cm. Approval was received from the university ethics committee and followed the tenets of the Declaration of Helsinki.

4.4.3 Protocol

Volunteers undertook three repeated measures each for the push-up and pull-down method on the RAF rule, the Aston Accommodometer with a fixed target and the Aston Accommodometer with a resizing target. All measurements were taken in a single sitting. The instrument used and the push-up or pull-down procedure sequences were randomized for each patient to negate any fatigue effects from the statistical analysis. The target for both the RAF rule and the Aston Accommodometer consisted of a short sentence with a text size of 0.4 logMAR at 40cm written in Times New Roman font. When the resizing target was to be tested, feedback from the ultrasonic sensor could maintain the 0.4 logMAR text at all distances up to 10cm. Luminance on the screen was calibrated to 100 mcd and Weber contrast to 98% by use of a Minolta Ls-110 luminance meter.

For push-up measurements with the RAF rule, the cheek plate was rested on the subjects face and they were asked to focus on the attached target and report when blur was first perceived. The target was moved slowly by the examiner towards the patient until the first sustained blur and the distance recorded. Meanwhile for the pull-down method the target was placed as near to the patient as the rule would allow and slowly brought back by the examiner until the blur was cleared, again relying on the patient to report when this had occurred.

When using the Aston Accommodometer patients tested themselves for both the push-up and pull-down technique. They were instructed to hold the instrument at eye level and slowly bring the device from arms length towards themselves whilst holding the device as upright as possible. If the Aston Accommodometer detected that it was not being held upright a warning on the screen would show and ask the patient to re-adjust their hand position accordingly. To indicate that blur had occurred on the push-up, or that clearance was reported on the pull-down, the subject could tap the screen to end the test. Mean results from the last 5 ultrasonic sensor readings before the screen was tapped, were recorded as the distance measurement for that test.

4.4.4 Statistical testing

All data was checked for normality before parametric testing could be applied. Bland Altman analyses for comparing the difference between the three measurement methods were plotted as well as the mean values and 95% (1.97SD) confidence limits. Repeated measures ANOVA's were used for both the push-up and pull-down methods on all three instruments to determine whether the measurement method had a statistically significant effect on subjective AOA readings. If an effect was found then pair wise comparisons (Paired t-test) were conducted to assess which means differed from each other. Inter-test repeatability was assessed using Bland-Altman analyses of the first and third repeated measure, and the pooled standard deviation for all patients was calculated.

4.4.5 Results

4.4.5.1 Push-Up Technique

Mean scores for the RAF-rule, Aston Accommodometer with a resizing target, and the Aston Accommodometer with a fixed target were $8.7 \pm 1.7D$, $8.8 \pm 1.9D$ and $9.1 \pm 2.2D$ respectively for the push-up technique. Bland Altman plots comparing the three experimental methods can be seen in Figure 4.16a, Figure 4.16b and Figure 4.16c. Repeated measures ANOVA applied to the data found that no significant difference ($F=1.893$, $p=0.16$) between the three measurement methods for assessing subjective push-up amplitude of accommodation. Pearson's correlation values and the 95% Bland-Altman confidence intervals between each method are shown in Table 4.2.

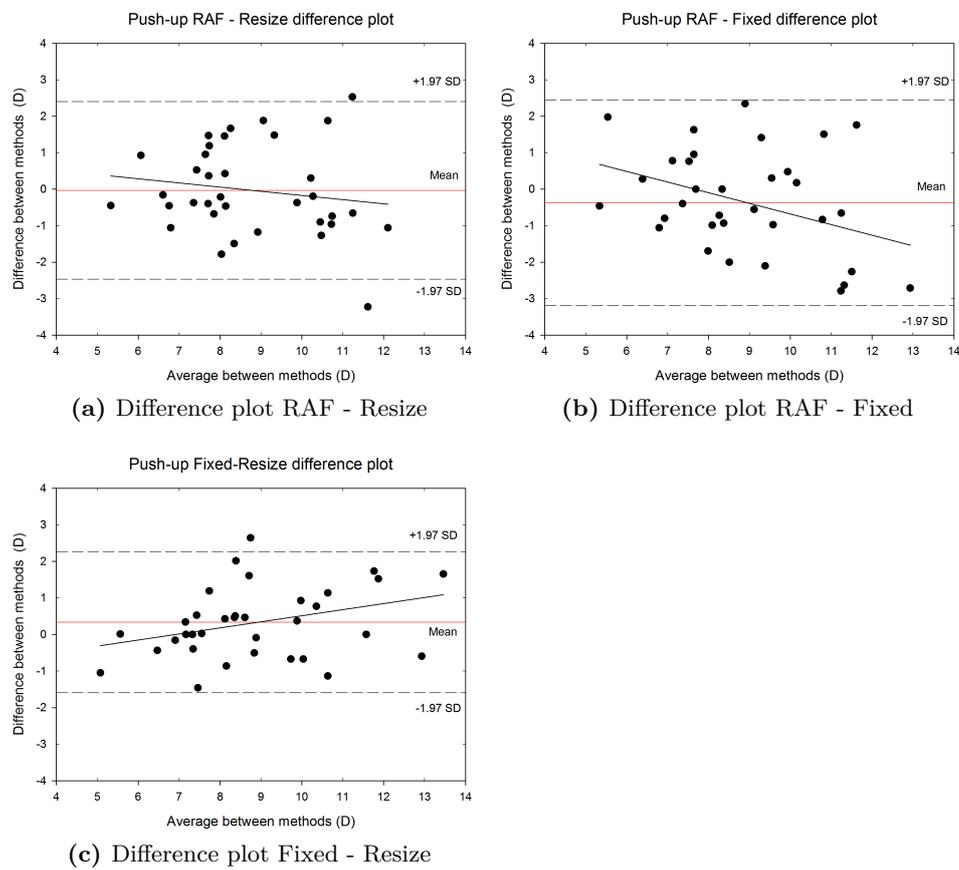


Figure 4.16: Bland Altman analyses for the push-up test

Table 4.2: Push-Up Correlation table

Comparison	Pearson's correlation	BA-mean (D)	BA-95% CI (D)
RAF-Resize	0.77	-0.03	2.40
RAF-Fixed	0.76	-0.37	2.77
Fixed-Resize	0.89	0.33	1.90

4.4.5.2 Pull-Down Technique

Mean scores for the RAF-rule, Aston Accommodometer with a resizing target and the Aston Accommodometer with a fixed target were $8.3 \pm 1.4D$, $7.5 \pm 1.4D$ and $8.1 \pm 1.7D$ respectively for the pull-down technique. Bland Altman plots comparing the three experimental methods can be seen in Figure 4.17a, Figure 4.17b and Figure 4.17c. Repeated measures ANOVA applied to the data found that a significant difference ($F=9.069$, $p<0.0005$) between the three measurement methods for assessing subjective pull-down amplitude of accommodation existed. On further analysis using paired T-tests, no significant difference ($p=0.44$) was found between the RAF rule and the fixed target Aston Accommodometer. However significant differences were found between both the RAF rule and the resizing target Aston Accommodometer ($p<0.001$) and between the fixed and resizing target Aston Accommodometer ($p=0.004$). In both cases the resizing target had significantly lower measures for the pull-down technique. Pearson's correlation values and the 95% Bland-Altman confidence intervals between each method are shown in Table 4.3.

Table 4.3: Pull-Down Correlation table

Comparison	Pearson's correlation	BA-mean (D)	BA-95% CI (D)
RAF-Resize	0.62	0.82	2.24
RAF-Fixed	0.78	0.15	2.17
Fixed-Resize	0.69	0.68	2.49

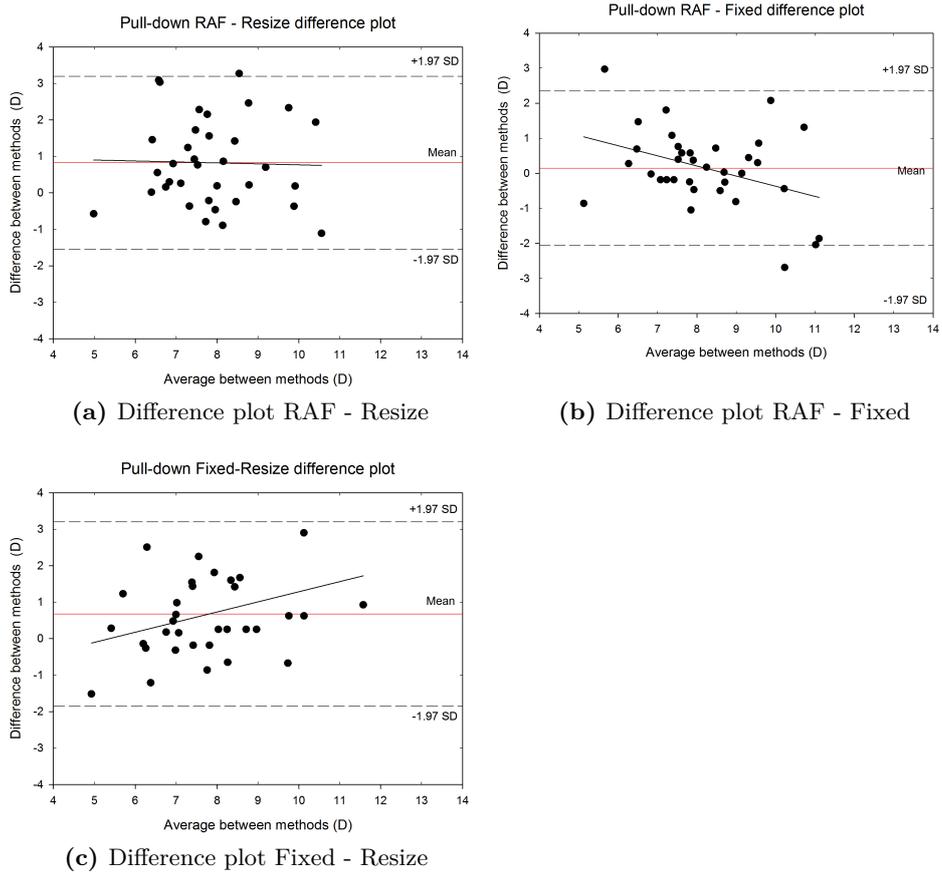


Figure 4.17: Bland Altman analyses for the pull-down test

4.4.5.3 Push-Up and Pull-Down comparison

Paired t-tests found a significance difference for the RAF rule ($p=0.005$), resizing Accommodometer ($p<0.001$) and fixed Accommodometer ($p<0.001$) between push-up and pull-down technique. A Bland Altman plot examining the difference between the push-up and pull-down technique for each experimental method can be viewed in Figure 4.18. The mean difference and 95% agreement intervals show that the push-up up technique provides slightly higher amplitudes of accommodation in all the methods, however the difference was greater when the Aston Accommodometer with a fixed or resizing target was used. As would be expected strong correlation values were found between techniques except for the resizing target where only mild correlation was found, see Table 4.4.

Push Up - Pull down difference plot for each measurement method

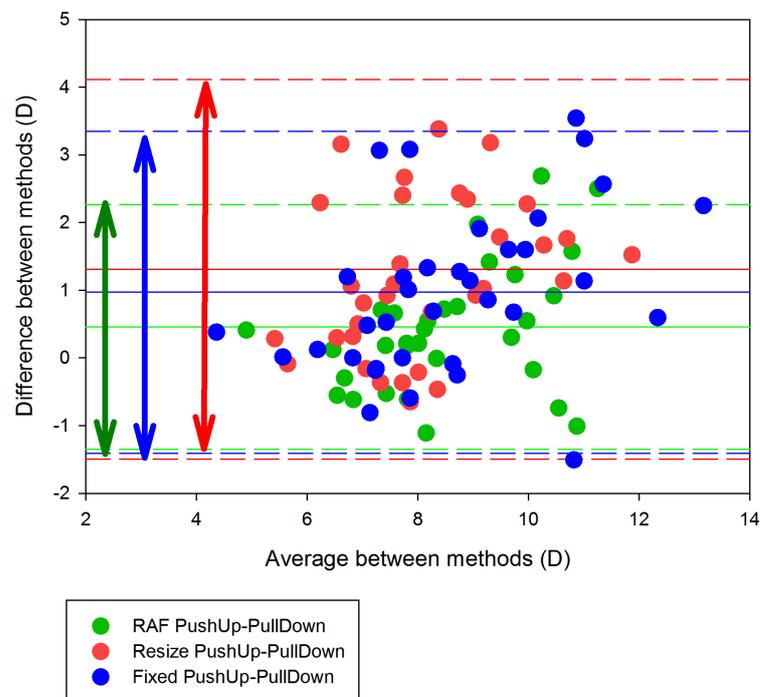


Figure 4.18: Difference plot between the push-up and the pull-down technique for all testing methods

Table 4.4: Push-up and Pull-Down Correlations

Comparison	Pearson's correlation	BA-mean (D)	BA-95% CI (D)
RAF Push-up,Pull-down	0.84	0.45	1.78
Resize Push-up, Pull-down	0.66	1.30	2.76
Fixed Push-up,Pull-down	0.85	0.97	2.35

4.4.5.4 Intra-Session Repeatability

Pooled standard deviation of the repeated measures taken on each instrument for both techniques are shown in Table 4.5. The RAF rule showed the least variance for both push-up and pull-down techniques, where as larger variability was experienced with the Aston Accommodometer. A Bland-Altman difference plot between the first and the third measurement for each method can be seen in Figure 4.19 and Figure 4.20. All 95% confidence intervals were larger for the both the resize and fixed target Aston Accommodometer compared to the RAF rule for both the push-up and pull-down technique. Moreover with all three testing instruments, differences increased as the mean accommodative value increased. Both the pooled standard deviations and confidence intervals were larger with the push-up than the pull-down technique.

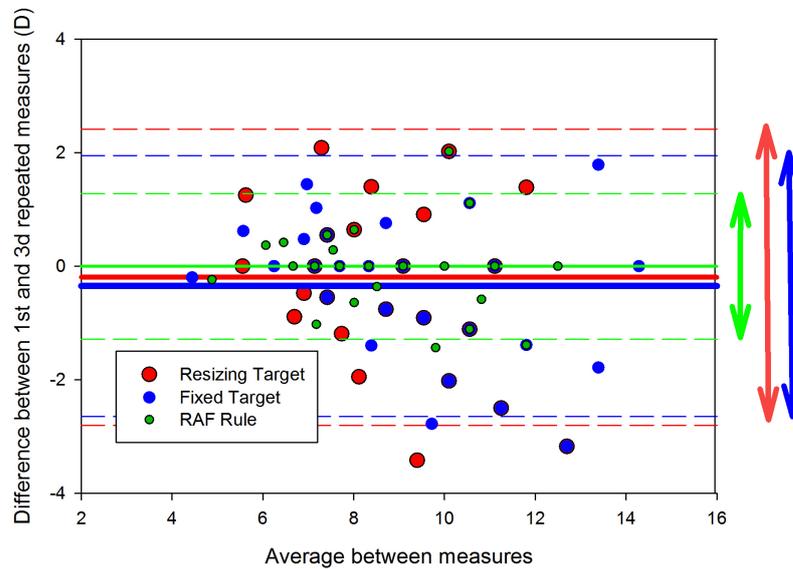


Figure 4.19: Difference plot between the 1st and 3rd repeated measure on each instrument for the push-up technique

Table 4.5: Pooled Standard Deviations for each instrument

Instrument	Push-Up S_p (D)	Pull-Down S_p (D)
RAF rule	0.43	0.50
Resize	0.88	0.61
Fixed	0.82	0.67

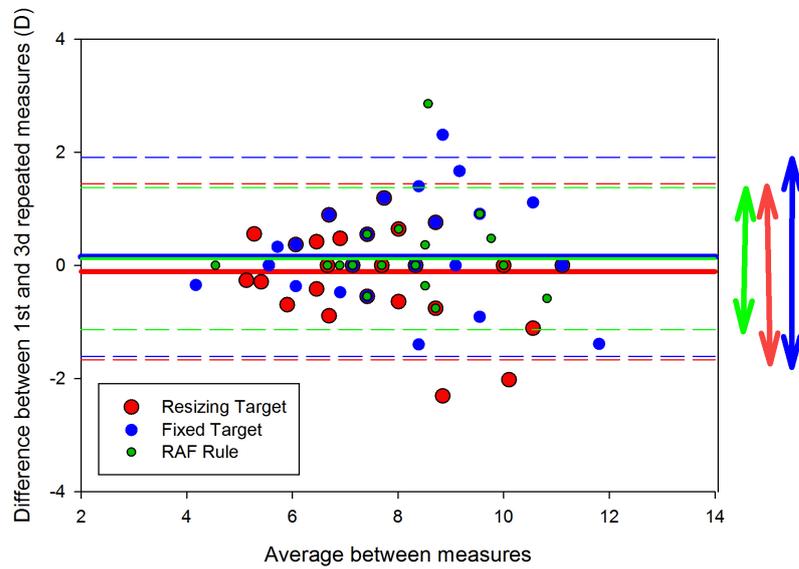


Figure 4.20: Difference plot between the 1st and 3rd repeated measure on each instrument for the pull-down technique

Table 4.6: Bland-Altman repeatability data

Instrument	Push-Up mean (D)	Push-Up 95% CI (D)	Pull-down mean (D)	Pull-down 95% CI (D)
RAF rule	0.00	1.27	0.12	1.24
Resize	-0.19	2.57	-0.11	1.53
Fixed	-0.34	2.26	0.15	1.73

4.4.6 Discussion and Conclusions

Mean push-up and pull-down accommodation measures for all three testing methods (range: 7.5D - 9.1D) for a patient cohort of mean age 23.3 years, were slightly lower than would be predicted by the Hoftstetter equation at 11.5D [Hofstetter, 1950] and lower than binocular measurements taken by Duane [1922] at 10.3D. The reason for this is unknown. push-up values were significantly larger than pull-down values for all three testing conditions which is in agreement with the results of Antona et al. [2009], and Chen and Leary [1998]. However, these previous studies found greater differences (+1.8 and +1.4D respectively) than those in this study (RAF + 0.45D, resize 1.3D and fixed 0.9D) most likely due their younger cohort.

There was no overall systematic difference between the Aston Accommodometer and RAF rule measurements for the push-up technique despite a resizing target being used. The effect of a resizing target on the push-up method was not significant in this study of 34 patients. Maintaining a constant angular subtense with a target may not be important in subjective push-up AOA methods as was once thought. This is in disagreement with Atchison et al. [1994] who found that reducing the size of the target with distance in turn reduced accommodative amplitudes by a mean of 1.4D. However, they used a target that was always at or near the acuity limit, this could not be replicated on the iPhone screen due to resolution limitations.

In contrast to the push-up test results, the pull-down method was significantly affected by the resizing target. In the pull-down method, the magnified relative size of the target means clearance is reported earlier, hence the higher mean AOA in the Aston Accommodometer with the fixed target (8.1D) than when it maintains its angular subtense (7.5D). Furthermore correlation values between push-up and pull-down values were lower with the resizing target ($r=0.662$) than with both of the fixed target conditions ($r=0.84$, $r=0.85$). If patients are consciously aware of the changing target size it may be masking the “clear” condition used with the pull-down method, this interference with subjective pull-down AOA measurements may cause correlation values to drop between push-up and pull-down techniques. It is unknown as to why the resizing target affects the pull-down method more than the the push-up method although it may be related to the differing end conditions between the two tests.

Bland-Altman analyses between the first and third repeated measures found larger confidence intervals with the Aston Accommodometer than the RAF rule. In addition the pooled standard deviations between all three repeated measures were higher

with the Aston Accommodometer. The increase in variability is most likely caused by the ultrasonic beam hitting various elements of the face as it is being brought in or pulled away. This is a limitation to the design of the Aston Accommodometer. Even when attempting to hold the Accommodometer still, small changes in displacement can cause attenuation and acoustic properties to change dramatically for the ultrasonic sensor, leading to higher variability between measurements. This is well known and is described previously in subsection 4.3.1 on page 142. Furthermore, the RAF rule is constrained to move along one dimension and is operated by a trained examiner, meanwhile the Accommodometer is free to move anywhere and is operated by untrained patients, which will also cause higher variability between measures. Within session Bland-Altman repeatability for the pull-down method on the RAF rule, 1.24D, was close to the 1.21D found in a study by León et al. [2012] despite the latter testing AOA monocularly.

Whilst the Aston Accommodometer measures were more variable than those of the RAF rule, it is not a major problem as the standard practice is to take the mean of three measures; as we have already seen in this study, when this is done there are little differences if any between the fixed target Accommodometer and the RAF rule for both the push-up and pull-down method. Given the repeatability of the techniques concerned the observed differences are practically unnoticeable.

In conclusion it appears that the Aston Accommodometer is a reliable and valid method of determining both push-up and pull-down accommodation values when used with a fixed target. Therefore the Aston Accommodometer with a fixed target has the potential to capture subjective AOA measurements within a clinical setting and is interchangeable with the RAF rule. The accommodometer also does not rest invasively on a patient's cheeks, which is an advantage when testing patients who are culturally sensitive to revealing or touching the face. However it should not be used with a resizing target until further studies have demonstrated the effect of resizing targets on patients AOA.

5 A New Model For Simulating Localized Tear Film Breakdown in Video-Keratotomy images

5.1 Introduction

The tear film, as seen in Figure 5.1, is critical to the health and optical clarity of the eye. A healthy tear film can protect the anterior of the cornea from infection, provide a clear smooth surface for optical clarity, and deliver essential nutrients to the corneal epithelium. Diagnosing tear film conditions and maintaining a healthy tear film however provide many challenges as the tear film is difficult to assess using current technologies. Many tests are reliant on trained health professionals, require subjective assessment, and are invasive in nature. Tear film related conditions such as dry eye, which is said to affect 15% of the population [Moss et al., 2000], are thought to impact heavily on vision related quality of life [Miljanović et al., 2007] and ocular health. Improved models of tear film degradation in video-keratotomy images can be used to develop more sensitive algorithms and equipment to measure tear film surface quality. The start of this chapter is going to discuss the various functions of the tear film and how they are reflected in the tear film structure. The causes and effects of tear film breakdown will be explored and their relationship to video-keratotomy images will be discussed. After evaluating current models of Placido Disk based devices and tear film models, a new tear model will be proposed to simulate tear film break down in video-keratotomy images.

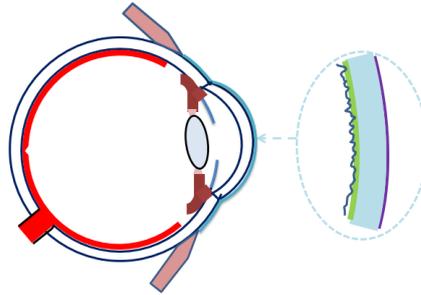


Figure 5.1: Tear film location

5.1.1 Tear Film Functions

5.1.1.1 Optical clarity

The cornea accounts for nearly two thirds of the eyes refractive ability [Gross, 2008]. The pre-corneal tear film is the most anterior optical surface of the cornea. Any changes in thickness or stability of the tear film, can dramatically affect vision. With irregular tear film thickness, higher order wave front aberrations can occur to the detriment of visual clarity [Koh et al., 2002]. When full break up occurs the irregular corneal epithelial surface can be exposed, causing optical scattering [Fine and Yanoff, 1979]. Hence the smooth surface provided by the tear film over the irregular corneal surface greatly aids in a patients vision, and is one of the most important functions. Studies have suggested that an irregular tear film surface, where the thickness is varied or no thickness at all can affect spatial vision and contrast acuity [Tutt et al., 2000]. A diagram illustrating the effect of an irregular tear film on the optical path of light can be seen in Figure 5.2.

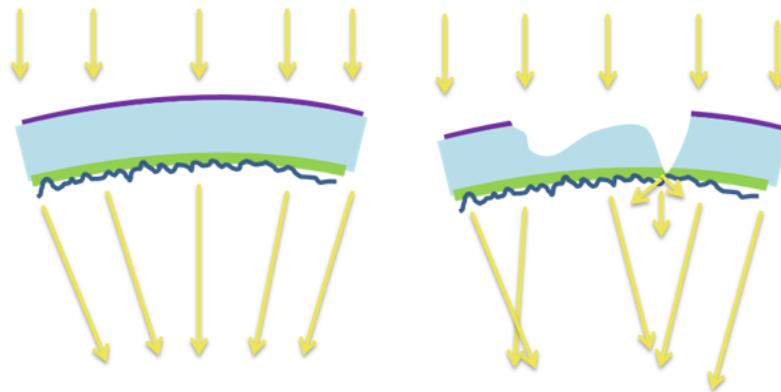


Figure 5.2: The effect of an irregular tear film on the optical path of light

5.1.1.2 Lubrication

A normal blink, where the superior lid margin passes over the anterior eye, would subject the epithelium to high shear forces ($170 \frac{\text{Dynes}}{\text{Cm}^2}$ [Mutharasan and Srinivas, 2002]). However, with the tear film providing a covering lubricating layer, the shear forces on the epithelium are significantly reduced. This limits any damage to the epithelium cells. In cases where the tear film has been disrupted and this lubrication function is absent friction-related damage to epithelium cells is observed [Ren and Wilson, 1997].

5.1.1.3 Protection

The eye has constant exposure to environmental conditions. This includes different temperatures, humidity, air-borne pollutants, particulate matter, wind, UV radiation and foreign bodies. The tear film acts as physical barrier protecting the exposed eye to several of these environmental conditions. Many of the chemicals, for example, which make up the tear film particularly in the mucous, have been known to resist bacterial and viral infections [Holly et al., 1986]. Tear production is stimulated in times when the eye is overcome by environmental conditions, which can wash away particulate matter. Hence the tear film therefore can act not only as a barrier but one that can be shed in order to protect the eye.

5.1.1.4 Nutrition

To maintain transparency the cornea lacks vasculature. Due to the lack of blood vessels, another method for transferring nutritional content to the corneal epithelium takes place. Studies have suggested that the aqueous humor supplies glucose, through diffusion to these cells, but the tear film supplies essential electrolytes [Botelho and Martinez, 1973] and oxygen supply. It has been determined that certain growth factors are supplied in the tear film to enable regeneration of epithelium cells [Rolando and Zierhut, 2001], however there is much debate as to whether the tear film also provides critical amounts of glucose [Chen et al., 1996] to ensure the cells of the corneal epithelium are well nourished.

5.1.2 Tear Film Structure.

The tear film has to provide a stable and maintainable environment for the corneal surface, in spite of huge environmental challenges. In addition it must also be optically optimised. A modern and widely held description of the structure was first described by Wolff [1954]. His depiction of the tear film was of a three layered structure consisting of mucus, aqueous, and lipid layers. Most, if not all, further descriptions of the tear film are additions or corrections of Wolff's initial premise. The three layered tear film is understood to have a thin mucin layer, primarily to anchor the tear film to the corneal epithelium. Directly above the mucin layer, an aqueous layer, to enable the tear film to spread evenly over the eye. On the exterior of the tear film, a lipid layer, to prevent the aqueous from evaporating. This classical model, although still in use, has been superseded by suggestions that the three layers are not so clear cut. It has been proposed that an aqueous-mucin gel is in place rather than three separate layers [Gipson, 2004, Bron, 2004a, Cher, 2007]. The aqueous-mucin gel is thought to have a concentration gradient whereby mucus is more concentrated closer to the corneal epithelium than to the lipid layer. See Figure 5.3 for a diagram depicting the classical trilaminar and the suggested aqueous-mucin gel model of the tear film structure.

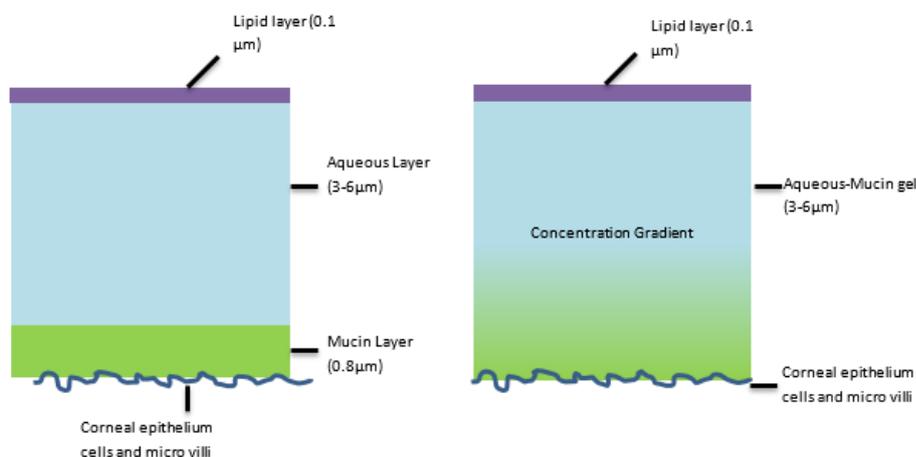


Figure 5.3: The classical trilaminar and the aqueous-mucin gel model of the tear film.

Much of the debate concerning the structure of the tear film is not only between the classical and new structures, but the thickness of the tear film. Several papers have quoted variable thickness of the tear film and its separate layers using several

different assessment methods [Prydal et al., 1992, King-Smith et al., 1999, Benedetto et al., 1975]. Estimates of the thickness of the tear film varies, but with modern interferometric techniques and spectra reflection, thickness is quoted around 3- 6 μ m [Hosaka et al., 2011, King-Smith et al., 2009].

5.1.2.1 Mucin Layer

The mucus phase consists of mucin, immunoglobins, urea, salts, glucose, leukocytes, cellular debris and enzymes [Bron, 2004a]. Mucus is secreted by goblet cells in the conjunctival region of the eye [Gipson et al., 1992] and has several different functions. Firstly it anchors the tear film to the cornea, by attaching to the glycocalyx present on the surface of the epithelium. It is also acts as a physical shield preventing attachment of bacteria and particulates on the ocular surface. Notably it provides a smooth lubricated surface for the eyelid margin to pass without abrasion, and lastly, affords corneal and conjunctiva surfaces with the necessary required hydration for optimum function (the corneal epithelium is hydrophobic [Holly and Lemp, 1971]).

5.1.2.2 Aqueous Layer

The aqueous portion is composed of dissolved oxygen, growth factors, anti-bacterial substances, and electrolytes [Pflugfelder et al., 2004] and is secreted by the main lacrimal gland. The major function of this layer is to provide oxygen and carry waste products away from the corneal epithelium [Agarwal, 2006]. However this is not its only function, anti-bacterial substances have been found to shield the cornea from infection [Van Haeringen, 1981]. The thickness of the aqueous phase of the tear film is debateable, with suggestions ranging from 3-4 [King-Smith et al., 2009] microns through to a thickness of 40 microns [Prydal et al., 1992]. As this layer is relatively thick, it enables tears to be used in the removal of debris from the eye.

5.1.2.3 Lipid Layer

The lipid layer acts as the air tear boundary to which it is exposed directly to environmental conditions. The primary source of the lipids are the meibomian glands and it's composition is mainly that of waxy substances (lipids) cholesterols and some sterols [McCulley and Shine, 2003]. This has led to many reports suggesting its main attribute is one of retarding evaporation of the tear film [Maïssa and Guillon,

2010, Tsubota, 1998] and underlying aqueous layer, a key obstacle in maintaining an equal tear film over the eye. Studies where the lipid layer has been lacking in patients suggest that this layer plays a key role in the spreading, and hence regular surface, of the tear film [Foulks, 2007]. As the outermost layer of the tear film, its job is to provide a smooth optical surface. Its thickness is approximately 100 nm [Bron, 2004a]. As the outermost and arguably the most important layer for tear film stability and visual acuity, it is the most widely studied.

5.1.3 Tear Film Maintenance, Thinning, Rupture and Dry eye.

Maintenance falls to the lacrimal apparatus and its constituent parts for the secretion, distribution and excretion of the majority of the tear film. The distribution of the tear film is complex and can be influenced by several factors. This includes mechanical elements (called hydrodynamic factors) such as lid blinking and lid closure which affect tear spread and tear clearance. The chemical composition and volume of tears, which are comparatively important for an even undisturbed tear film, are secreted and maintained by several glands. A diagram depicting the lacrimal system can be seen in Figure 5.4, whilst a table summarizing the main productive organs of the tear film can be witnessed in Table 5.1. It has been estimated that basal tears are secreted at a rate of $1.05 \pm 0.39 \mu\text{l}/\text{min}$ [Tomlinson et al., 2009].

Drainage of tears within the lacrimal apparatus is the job of the puncta, canaliculi, sac and nasolacrimal duct as well as normal evaporative processes on the ocular surface. A comprehensive review of evaporative tear film papers have concluded that approximately 15% of the tears ($0.14 \pm 0.07 \mu\text{l}/\text{min}$) secreted are lost to evaporation [Tomlinson et al., 2009], however laser interferometry studies have suggested that the evaporation rate may be as high as $1.14 \mu\text{l}/\text{min}$ [Kimball, 2009]. Minor re-absorption of tears through the cornea or Lacrimal canals has also been suggested although it is not thought to constitute the bulk of tear film excretion [Tomlinson and Khanal, 2005].

The remaining tears that are not lost to evaporation and re-absorption are thought to be excreted, through the Puncta along the Lacrimal canaliculi into the Lacrimal sac and out to the Naso-Lacrimal duct. Although this excretion of tears through the nose is hard to measure in vivo, balancing tear secretion flows with those lost to tear evaporation suggests that the bulk of drainage must occur through the lacrimal system. The process of tear drainage can be explained by both active and passive elements. During a blink, lid closure can pump tears through to the nasal cavity

via the lacrimal canaliculi, lacrimal sac and nasolacrimal duct [Rosengren, 1972]. Another theory suggests that a “lacrimal pump” exists, whereby upon a blink the lacrimal sac dilates causing a low pressure within the sac that draws tear fluid through from the canalluci [Jones, 1957]. Passive drainage by gravity may explain the thinning of the tear film during the interblink phase where active blinking affects are absent. Studies where patients were placed both vertically and horizontally (to negate the effects off gravity) show that drainage was significantly affected by gravity [Sahlin and Chen, 1997].

Any deficiencies within this complex system of secretion, distribution and excretion can lead to several negative consequences. An umbrella term for the diseases occurred by the failing of part or parts of this system of secretion and excretion is called dry eye disease.



Figure 5.4: Lacrimal System

5.1.3.1 Tear film dynamics

Three distinct stages are associated with the tear film dynamics during the inter-blink period when the eye is fully exposed to the environment. The first stage is called the tear film build up and occurs directly after blinking. It is then followed by a stable inter-blink stage and finally by a break-up stage as the tear film becomes unstable. It is noted however, that in normal conditions blinking can occur before the break-up stage commences ensuring an even and undisturbed tear film. For

Layer	Production	Functions
Lipid	Meibomian Gland	Smooth Optical Surface,Retards Evaporation,Antibacterial protection,Imparts stability on the tear film
Aqueous	Lacrimal and Wolfring Glands	Provides oxygen to the corneal epithelium, Carries waste products from the corneal epithelium,Provides growth factors, Antibacterial protection
Mucin	Conjunctiva Goblet Cells,Corneal and conjunctiva epthelium	Anchors the tear film to ocular surface, provides wetting, lubrication for eye lid movement.

Table 5.1: Summary of tear film functions

this reason, the three stage model can be observed greater in patients in suppressed blinking conditions (SBC) than those under normal blinking conditions (NBC).

Tear film build up phase Spreading of the lipid and aqueous layers across the surface is most closely associated with this phase. Upon a blink the tear film is pulled down by the upper eyelid. When the lids reopens, marginal reservoirs (meniscus) of lipid and aqueous are spread rapidly upwards over the surface of the eye [Creech et al., 2009]. Surface tension gradients during the upstroke movement of the upper lid is the driver for this movement [Owens and Phillips, 2001, Shiobara and Schnider, 1989]. Two to three seconds post blink has been suggested as the time it takes for the tear film to spread evenly [Benedetto et al., 1984] . Particles that rest on the surface of the tear film whose lateral velocity can be tracked in patients, suggest lipid movement occurs up-to 1 second post blink[Varikooty et al., 2012]. Evidence from laser shearing interferometry and video keratography studies suggest that a stable tear film structure supports movement around 2-3 seconds post blink, however it was noted some patients build up phase was as long as 10 seconds[Szczesna and Iskander, 2010, Németh et al., 2002]. Brown [1969] postulated that the tear build up time can be further split down into two distinct stages, one where tear film is pulled by capillary action from the eye lid movement and a second stage of slow

spreading lipid movement to thicken and stabilize the tear film. Dual stage build up has also been postulated as a result of polar and non polar lipid spreading [Bron, 2004b].

Stable inter-blink phase After the tear film has spread and stabilized a period of relative constancy in thickness and aberrations can occur. In normal blinking conditions it is thought that this phase is immediately followed by a blink [Benedetto et al., 1984] but a latter separate thinning and break-up phase can occur in suppressed blinking conditions. For normal eyes this period of stability can last for a sustained duration, however it may also be completely absent [Szczena and Iskander, 2010].

Thinning and Break up phase Once the tear film has stabilized the penultimate phase is for it to start thinning and eventually break down. The thinning of the tear film has been attributed to both evaporation [King-Smith et al., 2008, Kimball et al., 2010] and tangential flow [McDonald and Brubaker, 1971]. Theories concerning the mechanism of tangential flow point to “Pressure gradient flow” by the concave outer surface of the tear meniscus, Marangoni flow or poor wettability [Nichols et al., 2005]. Both evaporative and tangential flow thinning are heavily reliant on the underlying structure of the tear film for their properties. If the tear film has inadequate aqueous, it may be spread thinly and consequently further thinning either from evaporation or tangential flow will quickly deplete the tear film. Alternatively, if the production of the lipid layer is insufficient increased evaporation and poor spreading properties may cause tear film instability [Fatt, 1991].

The tear film is said to have broken down (sometimes called break-up) when it no longer covers the entirety of the corneal surface. Once the tear film has thinned significantly, dry spots where the tear film has broken down can be seen on the corneal surface. Tear film break-up time (TBUT) is the duration between the last blink and when the first observed dry spots appear, and is a commonly used metric when diagnosing dry eye disease. Normal patients who suppress blinking can also experience tear film breakdown, however the time with which it takes to occur is greatly increased. You would expect that if the tear film was thinning at a constant rate across the cornea that dry spots would appear almost simultaneously (assuming the underlying tear film was of constant thickness) but tear break down tends to be more localized.

It is well known that the corneal surface is hydrophobic and requires the mucoid layer for wetting and adherence of the preceding tear film layers. Many of the purported mechanisms of tear film breakdown explain how de-wetting can occur between the mucoid layer and the corneal surface. This has been attributed to underlying defects and floating debris acting as nucleation sites [Fatt, 1991], lipid distribution [King-Smith et al., 2013], lipid contamination of the mucus layer [Holly, 1973], and dispersion forces [Sharma and Ruckenstein, 1985]. The exact mechanism for how these dry spots form is not known however.

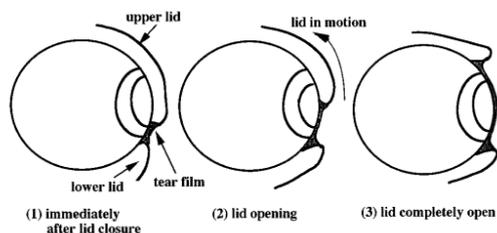


Figure 5.5: Lid movement and the build up phase.

5.1.3.2 Tear film break-up geometry

The geometry of tear film break-ups is invaluable when understanding techniques of tear film analysis that rely on the reflectance of light to determine tear stability. This includes clinical tests such as noninvasive break-up time and tear stability analyses performed by video keratoscopes. Examining the size, shape and geometry of a rupture is challenging owing to the small thickness ($3\text{-}6\mu\text{m}$) of the tear film and the propensity for micro-movement of the eyes. Thus only one paper, from Licznarski, Tomasz J and Kasprzak, Henryk T and Kowalik [1999], has been found that has effectively mapped a 3D tear film break-up in-vivo.



(a) 3D representation

(b) Cross-section

Figure 5.6: Tear film break-up cross sections and 3D shape, taken from [Licznarski, Tomasz J and Kasprzak, Henryk T and Kowalik, 1999]

Licznarski, Tomasz J and Kasprzak, Henryk T and Kowalik [1999] used a Twyman-Green interferometer to observe interference patterns caused by the underlying topography of the tear film on a patients cornea in-vivo. Analysis of the resulting interferograms suggest that the depth of break ups remains constant at approximately $1.5 \mu\text{m}$ however width and length of break up change with time. The shape and cross sections of observed tear film break-ups can be seen in Figure 5.6. Further work from [Kasprzak, 2003, Caneiro, 2010], that was modeling the affects of tear film breakdown on retinal image quality, used the preceding evidence from Twyman Green interferometry to create functions that approximately fitted tear film breakdown. A tear film dimple function using the same naming conventions as [Caneiro, 2010] , is shown in Equation 5.1. Here a is the depth of the dimple, b is a constant that is related to the half width p_{hw} of the dimple (see Equation 5.2) and $c(x, y)$ is a function that describes the radial distances from the dimple center (x_0, y_0) (see Equation 5.3). A 3D representation of the tear film dimple is displayed in Figure 5.7.

$$z(x, y) = a \cdot \left[1 - e^{-b \cdot c(x, y)^2} \right] \quad (5.1)$$

$$b = -\frac{\ln(0.5)}{\left(\frac{p_{hw}}{2}\right)^2} \quad (5.2)$$

$$c(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (5.3)$$

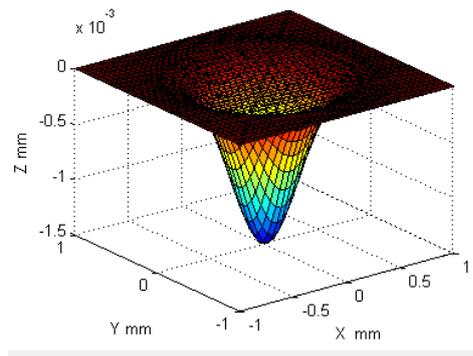


Figure 5.7: 3D Tear film dimple

5.1.4 Placido based devices and modelling

With the onset of low price sensors and the increasing power of computer technology, novel techniques for measuring tear film stability have emerged. The majority are variations and adaptations of commercially available Placido disk based corneal topographers [Kopf et al., 2008, Goto et al., 2003, Gumus et al., 2011, Alonso-Caneiro et al., 2011]. The principles of a Placido disk based corneal topographer can be seen below in Figure 5.8. A projected pattern of concentric rings, called Placido rings, are projected onto the cornea. Reflections of these concentric rings from the surface of the cornea (called Purkinje images) are captured by an image sensor mounted centrally to the cornea behind the projection device.

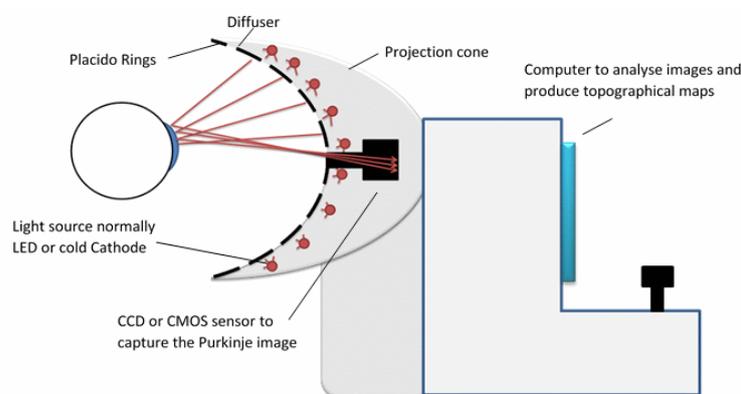


Figure 5.8: Placido based corneal topographer diagram

The relative positions and shape of the reflected rings give an indication of the curvature of the cornea. In a perfect cornea, one in which it acts as a perfect sphere, the rings should be reflected back with no distortion, a flat Purkinje image (see Figure 5.9 (i)). Conversely in an imperfect cornea, such as one that has Keratoconus, the rings become distorted (see Figure 5.9 (ii)).

Once analysed the reflected rings can act like a topographical map of the cornea to aid in a clinician's diagnosis and treatment of a patient. Typically these measure axial curvature, tangential curvature and elevation. Although popularly used to map the shape of the cornea, the image is actually formed by the reflection of the Placido rings on the tear film/air boundary. When the tear film is spread evenly it takes the approximate shape of the cornea however degradation of the tear film causes perturbations on the recorded image (see Figure 5.9 (iii)). These perturbations can be split into two categories; perturbations caused by epithelium exposure, and

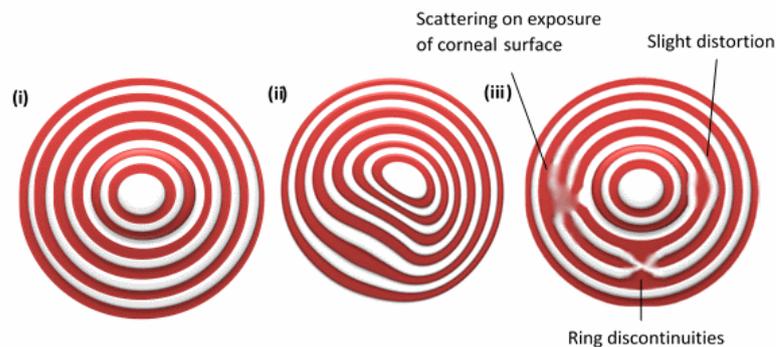


Figure 5.9: Illustration of Placido rings under different corneal conditions

perturbations caused by the tear film thinning. It is well known that the exposed cornea is a highly irregular surface which optically scatters light. This is shown as a blurring of the concentric rings, where the rings may appear wider and with less intensity (see Figure 5.10 (ii)). It can be noted that prior to this optical scattering, the tear film becomes locally irregular and the rings appear to have been deformed. This is caused by the local slope changing as the tear film begins thinning and can be easily explained by Snell's laws of reflection (see Figure 5.10 (i)). The detection of the local slope can also be used to determine the time it takes for the tear film to become uniform post blink [Alonso-Caneiro et al., 2011] as well as an indication of imminent break up [Arnold et al., 2010].

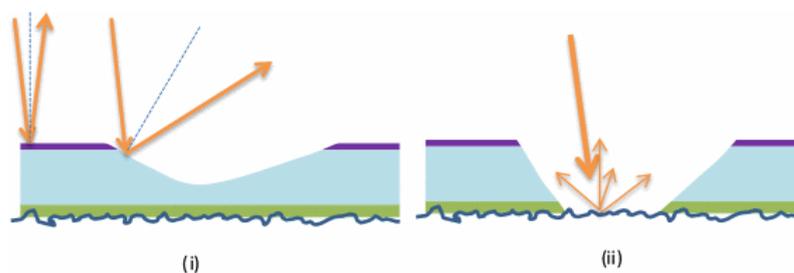


Figure 5.10: Perturbation diagram

5.1.5 Current Models

Placido disc analysis techniques often use geometrical models to deduce the underlying topography of a surface. Models vary but generally consist of a ring projection system, a simplified cornea, a lens or pin-hole aperture and an image plane [Mejia-

Barbosa and Malacara-Hernandez, 2001]. Figure 5.11 displays a basic geometrical model of a Placido cone based topographer. Only light that reaches the image plane is of interest in modeling Placido images, so backwards ray tracing starting from the image plane out through the optical system is used. Whilst ray tracing can involve three rays (principal, central, and focal), if you assume that the image will be in focus on the image plane, only central rays that pass through the equivalent nodal point of the lens system need to be plotted. With this simplification any shape of projection rings can easily be modeled, see Figure 5.12, as a flat Placido disc.

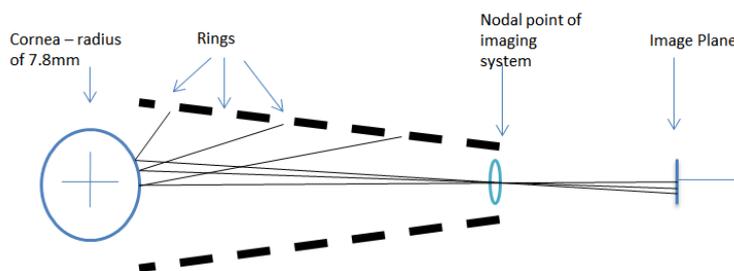


Figure 5.11: Geometrical model of a Placido disc topographer

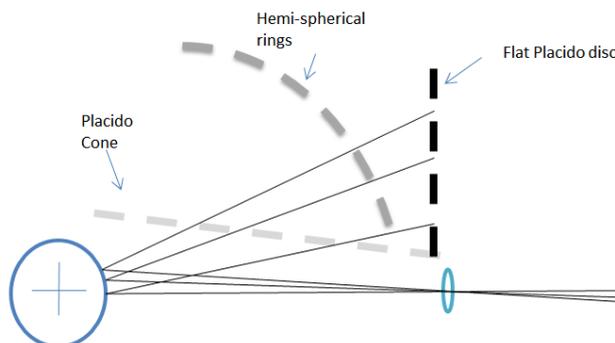


Figure 5.12: Placido Cones and Spheres modeled as an equivalent flat Placido disc

Two dimensional geometrical models are used extensively in Placido disc based corneal topographers who incorporate the arc step algorithm [Doss et al., 1981], and its variation van Sarloos [van Saarloos and Constable, 1991] algorithm, to re-

construct corneal profiles. This type of meridional analysis assumes that the cornea is rotationally symmetric; skew rays caused by toric surfaces cannot be traced back to their original destination. A diagram illustrating skew ray error is shown in Figure 5.13. There is an ambiguity in determining which ring point corresponds to an image point, as the initial ray may not travel in the same meridional plane as its subsequent reflection. For cases where the surface of the cornea may be considered toric, such as during tear film breakdown, a skew incident ray from point Q will hit the surface R and be reflected in a plane different to that of the incident ray, in this instance it happens to have been reflected in the meridian plane passing through P and Z. Any 2D analysis of the meridian plane will determine that the incident ray originated from P although it may have originated from a skew ray such as Q. Therefore to model the toric effects of real world tear film deterioration, three dimensional ray tracing models that can trace skewed rays should be used.

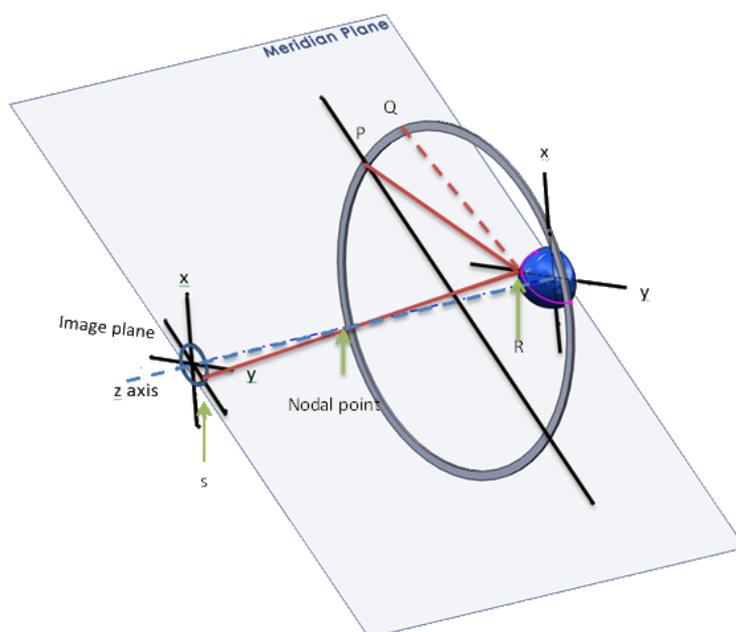


Figure 5.13: Skew ray error

As far as the author is aware, there are two publications [Alonso-Caneiro et al., 2011, Caneiro, 2010] that have modelled the effect of tear film breakdown on video-keratoscope images. The first [Caneiro, 2010] used backward ray tracing, a tear film dimple model (as described in subsection 5.1.3.2), and a flat cornea. A further paper from Alonso-Caneiro et al. [2011] used the tear dimple model but applied it

to a spherical surface. The neat symbolic equation provided for a break-up on a spherical surface is displayed in Equation 5.4, where a is the dimple depth, b the constant related to the half width p_{hw} , (x_0, y_0) the dimple centre and R the radius of a spherical cornea. This equation can be illustrated in Figure 5.14a, where an exaggerated dimple with a $p_{hw} = 0.4mm$ and $a = 1.0mm$ is shown. Both the aforementioned flat and spherical corneas could only model break-ups that happen horizontal to the imaging plane; it can be assumed that they actually occur at the normal to the corneal surface as depicted in Figure 5.14b. Backwards ray tracing used in these two papers took account of specular reflection at the surface of the tear-air interface, but no attempt was made to model the diffuse reflections from an exposed cornea.

$$z(x, y) = a \cdot (1 - e^{-b \cdot p_0(x, y)^2}) + \left| \sqrt{R^2 - p(x, y)^2} \right| - R \quad (5.4)$$

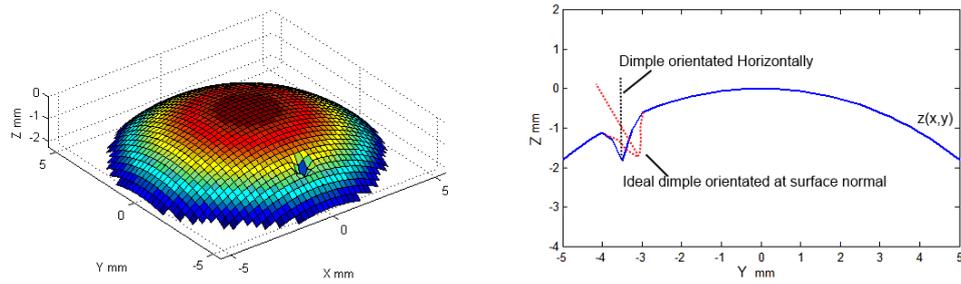
$$b = -\frac{\ln(0.5)}{\left(\frac{p_{hw}}{2}\right)^2} \quad (5.5)$$

$$p_0(x, y) = \left| \sqrt{(x - x_0)^2 + (y - y_0)^2} \right| \quad (5.6)$$

$$p(x, y) = \left| \sqrt{x^2 + y^2} \right| \quad (5.7)$$

5.1.6 Outcome

Two limitations of current Placido disc models have been identified in the previous section. The first that tear film break down is always occurring horizontal to the image plane rather than at the surface normal of the cornea, and secondly, that the backwards ray tracing model does not take into account diffuse reflections that would be associated with exposure of the corneal surface. Therefore, improved mathematical models that can account for these two phenomena would be beneficial for the



(a) 3D representation of equation Equation 5.4 (b) Cross section of the cornea with a tear film breakdown when $x = 0$

Figure 5.14: Spherical break up model

design of tear imaging systems that work with Placido disc based devices. Similarly, by creating more accurate models of tear film deterioration, algorithms and imaging systems could be devised that attempt to reduce the impact of these perturbations for more accurate corneal topography, or conversely track perturbations for objective tear measures.

5.2 Aims and objectives

5.2.1 Aim

The primary aim of this chapter is to create both a new tear film break up model orientated at the surface normal and an improved backwards ray tracing method to simulate diffuse reflection in Placido disc images.

5.2.2 Objectives

1. Identify mathematical techniques that can be used to create a new tear film break-up model on the cornea.
2. Create a ray tracing model of a video-keratoscope that uses both specular and diffuse reflection.
3. Simulate the new cornea and backwards ray tracing models using software.
4. Compare the results of simulations to previous tear models and those of tear film break downs observed in real video-keratoscope images.

5.3 New Combined Cornea and Tear Break-up Models

In this section the author proposes two new models for a tear “dimple” on a corneal surface. The first makes use of novel parametric equations to evaluate tear break up on spherical surfaces; tear dimples are always orientated towards the radial direction. The second uses numerical approximations to create tear dimples that are orientated at the surface normal of any arbitrarily shaped cornea.

5.3.1 Parametric method

Cartesian equations in the form of $z = f(x, y)$ where x, y and z are the Cartesian coordinates along the primary axes, are unsuitable to model tear dimples that are centred at surface normals. Looking again at Figure 5.14b you can see that the Cartesian function described by Alonso-Caneiro et al. [2011] has a unique z value for each (x, y) pair, however if you were to move this dimple to be orientated at the surface normal, as shown by the red dotted line, there could be two solutions for any coordinate pair and hence a function describing this in Cartesian parameters is impossible. Consequently, any new tear film modelled on the cornea should take the form of a parametric function.

Parametric descriptions of the surface of a sphere use both the longitude θ and latitude φ as parameters to define all (x, y, z) coordinates that lie on the spherical surface with a radius of r . A tear dimple centred at the surface normal of a sphere would still have a unique value for each (θ, φ) pair and only the radius r would have to vary depending on the location of the dimple, hence functions adapted from the parametric equations for a sphere can be created. The author proposes a new parametric model of a combined cornea and tear break up where $(x_s(\theta, \varphi), y_s(\theta, \varphi), z_s(\theta, \varphi))$ are the equivalent Cartesian coordinates for the tear surface described by Equation 5.8, Equation 5.9, Equation 5.10. R_c is the radius of the underlying cornea and $r_{tf}(\theta, \varphi)$ is the change in radius associated with the tear dimple centred at (θ_0, φ_0) ; $R_c + r_{tf}(\theta, \varphi)$ is the radii associated with each (θ, φ) pair. A diagram explaining the relations between parameters is shown in Figure 5.15a.

$$x_s(\theta, \varphi) = (R_c + r_{tf}(\theta, \varphi)) \cdot \cos(\theta) \cdot \cos(\varphi) \quad (5.8)$$

$$y_s(\theta, \varphi) = (R_c + r_{tf}(\theta, \varphi)) \cdot \sin(\theta) \cdot \cos(\varphi) \quad (5.9)$$

$$z_s(\theta, \varphi) = (R_c + r_{tf}(\theta, \varphi)) \cdot \sin(\varphi) \quad (5.10)$$

The function $r_{tf}(\theta, \varphi)$ in Equation 5.11, is related to a dimple previously described in Equation 5.1, but with fundamental changes to fit it to a sphere. Let P_{hw} be the half distance along an arc's perimeter with a radius of R_c . Therefore, the angle A_{hw} subtended by half the dimple can be defined as in Equation 5.13. As before a represents the dimple depth, and b is the constant related to A_{hw} . $\Delta\sigma(\theta, \varphi)$ in Equation 5.14 is the central angle between any (θ, φ) pair and (θ_0, φ_0) . The smaller the central angle, the more the dimple will have an effect on the resulting radius $R_c + r_{tf}(\theta, \varphi)$. A diagram of these relationships for a dimple on a spherical surface is presented in Figure 5.15b.

$$r_{tf}(\theta, \varphi) = -a \cdot e^{-b \cdot (\Delta\sigma(\theta, \varphi))^2} \quad (5.11)$$

$$b = -\frac{\ln(0.5)}{\left(\frac{A_{hw}}{2}\right)^2} \quad (5.12)$$

$$A_{hw} = \frac{P_{hw}}{R_c} \quad (5.13)$$

$$\Delta\sigma(\theta, \varphi) = \cos^{-1}(\sin(\varphi_0) \sin(\varphi) + \cos(\varphi_0) \cos(\varphi) \cos(|\theta_0 - \theta|)) \quad (5.14)$$

A three dimensional representation of the previous equations, with a dimple depth of $a = 1.5mm$ and half width $P_{hw} = 1mm$, is displayed in Figure 5.16a. Notice how the preceding cross section in Figure 5.16b demonstrates how the author's new

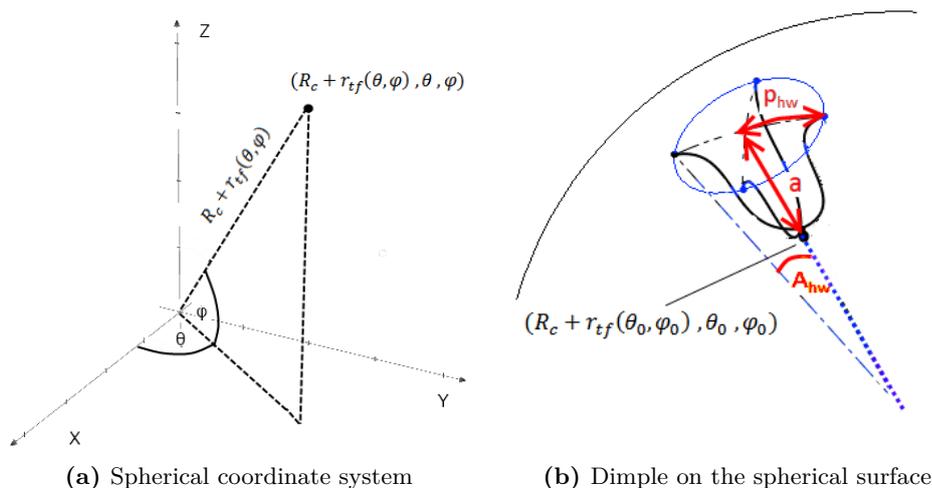
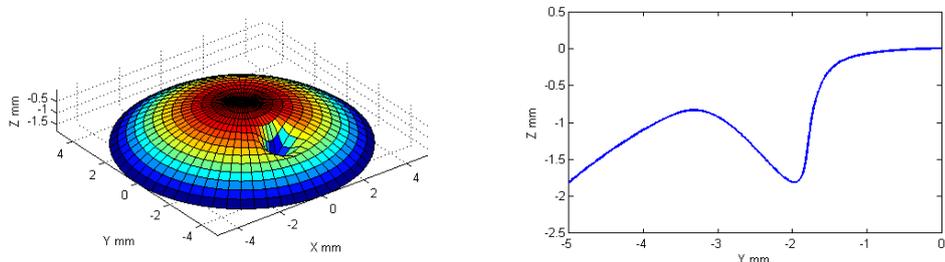


Figure 5.15: Conventions and relationships between parameters

break up model fits a dimple to the surface normal as opposed to the horizontal model of Alonso-Caneiro et al. [2011]. Adding additional r_{tf} functions centred at differing angles, can be used if multiple break-ups are required.

It must be noted that the novel equations derived in this section, always orientate dimples towards the centre of a sphere; if modelled on an oblate spheroid the dimples would no longer orientate at the surface normals. Whilst these neat equations can be used quite readily for a spherical approximations of the cornea, deriving more equations for use on an arbitrarily shaped cornea would be considerable more complex and time consuming. Hence numerical approximation techniques described in the following section, can provide a more efficient method of modeling tear film break-up on aspheric surfaces.



(a) 3D representation of Eqs Equation 5.8, (b) Cross section of the cornea and tear break Equation 5.9, Equation 5.10 with a dimple up at $\theta = \frac{-\pi}{2}$. situated at $\theta_0 = \frac{-\pi}{2}$ and $\varphi_0 = \frac{2\pi}{5}$.

Figure 5.16: Parametric tear film model.

5.3.2 Numerical approximation

The numerical approximation technique outlined in this section will allow a tear film break up to occur at the surface normal of any arbitrarily shaped cornea. Thus this novel technique can be used in a wider range of scenarios than the parametric equations. Additionally, whilst this method requires more calculation, it is conceptually much simpler than deriving formulas for each and every corneal shape.

5.3.2.1 Corneal model and mesh

Step one was to create the initial cornea in the form of a point cloud. Aspherical cornea model's better represent corneal geometry than a sphere [Bonnet et al., 1964], thus a model by Lotmar [1971] was chosen. Equation 5.15 is a Cartesian function where $(x, y, z(x, y))$ represent the coordinates on the corneal surface along the X, Y and Z axes, r_0 is the central radius of curvature and $p(x, y)$ the radial distance.

$$z(x, y) = -\frac{p(x, y)^2}{2 \cdot r_0} \cdot \left[1 + \frac{5}{28} \cdot \left(\frac{p(x, y)}{r_0}\right)^2 - \frac{1}{12} \cdot \left(\frac{p(x, y)}{r_0}\right)^4\right] \quad (5.15)$$

$$p(x, y) = \sqrt{x^2 + y^2} \quad (5.16)$$

It is important that the relations between successive points (called a mesh) on the point cloud is known to avoid surface ambiguities. If each element of $\mathbf{x} = (x_1, x_2 \dots x_n)$ needs to be evaluated with each element of $\mathbf{y} = (y_1, y_2 \dots y_m)$ for the equations above, three m -by- n matrices can be created as follows:

$$xM = \begin{matrix} & x_1 & x_2 & \cdots & x_n \\ x_1 & x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_m & x_1 & x_2 & \cdots & x_n \end{matrix} \quad yM = \begin{matrix} & y_1 & y_1 & \cdots & y_1 \\ y_2 & y_2 & y_2 & \cdots & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_m & y_m & y_m & \cdots & y_m \end{matrix}$$

$$zM = \begin{matrix} & z(x_1, y_1) & z(x_2, y_1) & \cdots & z(x_n, y_1) \\ z(x_1, y_2) & z(x_2, y_2) & \cdots & z(x_n, y_2) \\ \vdots & \vdots & \ddots & \vdots \\ z(x_1, y_m) & z(x_2, y_m) & \cdots & z(x_n, y_m) \end{matrix}$$

For any index $zM_{i,j} = z(xM_{i,j}, yM_{i,j})$. The structure of these matrices can effectively be used as a mesh, if each element is a vertex then triangulation can occur by connecting successive vertices in the pattern shown in Figure 5.17.

It is important to define how these vertices connect with each other at the start. Further down the process these vertex coordinates can change so successive points may not have increased x, y coordinates along the mesh as they had at the start, but their connectivity at the appropriate place will still be assured. The three matrices xM, yM, zM can be converted into a single matrix, V_{coords} , whose row indicates each vertex and whose columns are its (x, y, z) Cartesian coordinates. Furthermore, a connectivity matrix, $C_{triangles}$, whose rows indicate each surface triangle and whose columns define the identity of bounding vertices will fully define the mesh. The structure of these matrices can be seen below and their relationships illustrated in Figure 5.17 :

$$V_{coords} = \begin{matrix} & x_{v1} & y_{v1} & z_{v1} \\ & x_{v2} & y_{v2} & z_{v2} \\ & \vdots & \vdots & \vdots \end{matrix} \qquad C_{triangles} = \begin{matrix} & v_1 & v_{m+1} & v_2 \\ v_2 & v_{m+2} & v_{m+1} \\ \vdots & \vdots & \vdots \end{matrix}$$

With both vertex coordinates and mesh fully defined a surface can be generated, see Figure 5.18.

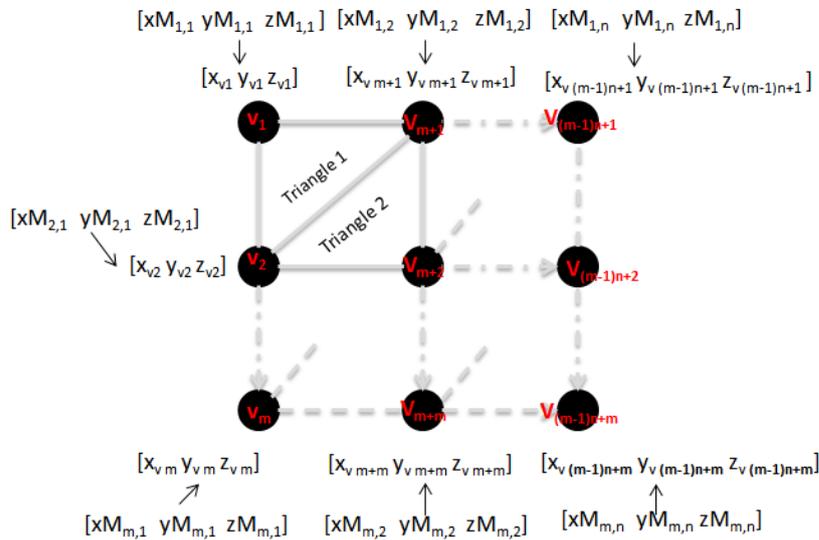


Figure 5.17: Initial Mesh

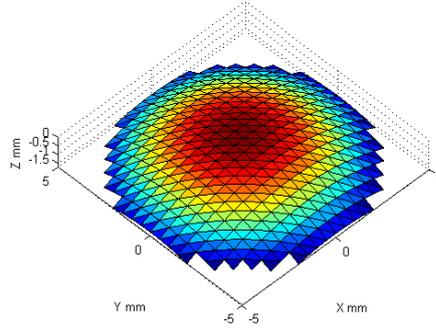


Figure 5.18: Surface mesh

5.3.2.2 Translation and rotation

The second step was to choose a vertex on the surface where the break-up is to occur, translate the cornea model so the vertex rests on the origin, find the surface normal at that point, and rotate the cornea so the selected vertex surface normal matches the Z axis.

Translation involved subtracting the selected vertex coordinates from each and every row in the matrix V_{coords} to produce V_{trans} . Each vertex was connected to a series of triangular faces; the surface normal at each vertex was taken as the average of the surface normal's of each connecting face. Both the connectivity list $C_{triangles}$ and vertex coordinate list V_{trans} are all that are needed to find a vertex's surface normal $S_{normal} = [s_x, s_y, s_z]$.

Having chosen an arbitrary vertex and calculated its surface normal, rotation matrices can be used on all the vertex positions to rotate the cornea in the specified direction. Using the $z - x' - z''$ Euler notation, the rotation matrix R will rotate the apex in the direction of the surface normal, see Equation 5.17. Conversely, to reverse the rotation and make the selected surface normal rotate to meet the Z axis, the transpose of this matrix, R^T , will be combined with the vertices coordinates V_{trans} to output a rotated list of vertex coordinates $V_{trans+rot}$ as described in Equation 5.18.

A visual representation of this method can be viewed in Figure 5.19.

$$R = \begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (5.17)$$

$$\psi = 0 \quad \theta = \text{atan2}(s_z, -\sqrt{s_x^2 + s_y^2}) \quad \phi = \text{atan2}(s_y, s_x) + \frac{\pi}{2}$$

$$V_{trans+rot} = V_{trans}R^T \quad (5.18)$$

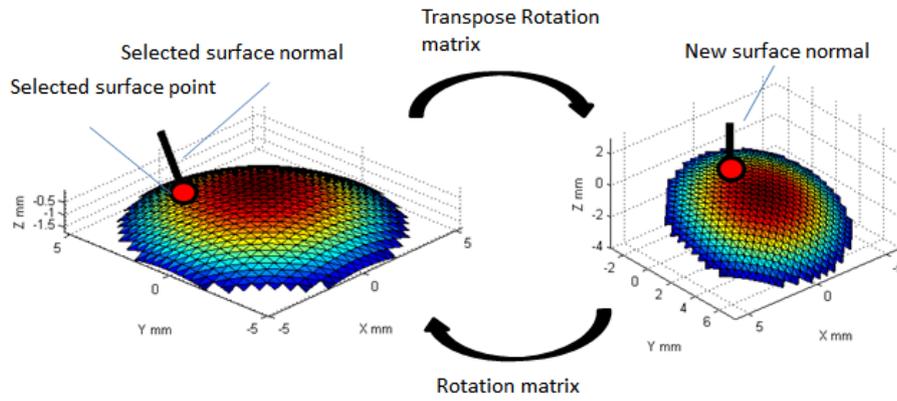


Figure 5.19: Diagram illustrating the rotation method

5.3.2.3 Applying the tear dimple and rotating to original position

As the selected vertex's surface normal now matches that of the Z axis, it becomes trivial to add the tear dimple to this model. For each vertex in the vertices list $V_{trans+rot}$, the x and y coordinates were taken and input into the tear dimple formula previously described in Equation 5.1, Equation 5.2 and Equation 5.3 on page 178. The resulting tear dimple values were then added to the z vertex coordinates to produce a modified vertices list $V_{trans+rot+mod}$. The rotation matrix R can then be applied to rotate the cornea back to its original position. A diagram explaining the process can be viewed in Figure 5.20.

$$V_{final} = V_{trans+rot+mod}R$$

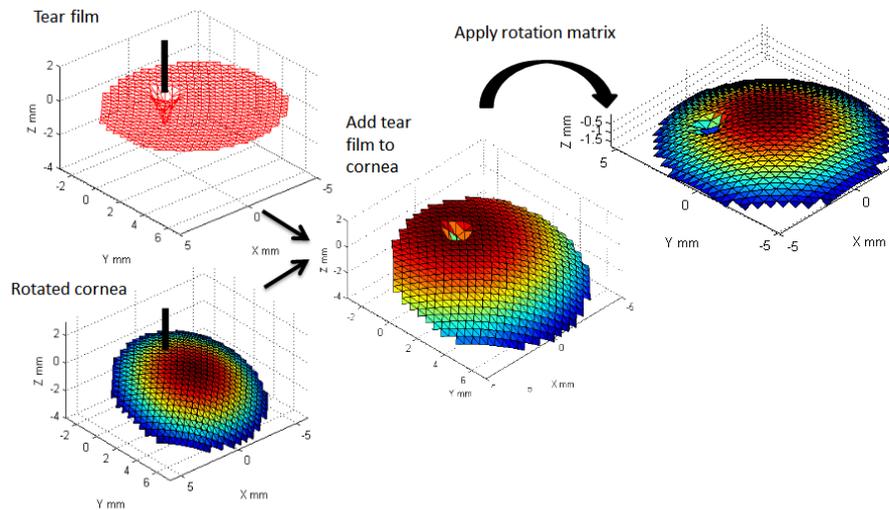


Figure 5.20: Creating and applying the tear dimple

As can be seen in Figure 5.20 tear dimples can now be orientated with respect to the surface normal of the cornea. The author believes this creates a more accurate representation of tear break down than that of Alonso-Caneiro et al. [2011], whose tear break ups always occurred parallel to the viewing axis irrespective of the surface it rested upon, and those for which this thesis derived earlier in subsection 5.3.1 on page 186 whose parametric equations relied on the cornea being modeled as a partial sphere.

5.4 Ray Tracing

Previous research that modelled the effect of tear film deterioration on Placido images [Alonso-Caneiro et al., 2011, Caneiro, 2010], have up to now incorporated only the law of specular reflection in their simulations. The author proposes that an attempt be made to simulate both specular and diffuse reflection, in order to include the effects of corneal scattering on Placido images. In addition, the new and more accurate tear film geometry, derived by the numerical approximation technique in the preceding section, could also be used.

5.4.1 Ray tracer

A basic ray tracing program was created using the MATLAB programming environment. It consisted of four main elements: the image plane, the nodal point of the imaging system, the corneal model, and the Placido projection plane. The system, Figure 5.21a, was set up to approximate the operation of real world Placido based topographers, however simplifications were made. Where many topographers use a lens before the imaging plane, this has been replaced with a small aperture called the nodal point. Rays could therefore be traced from the nodal point out towards the scene. Additionally, rather than trace rays backwards onto the image plane from the nodal point, the image plane can be brought forward between the nodal point and the cornea without any geometric distortion.

Twenty Placido rings that covered the whole of the cornea were used on the models Placido plane, their respective spacings were determined so that the rings traced on the image plane would appear equidistant when no tear film distortion was present. Figure 5.21a identifies the main elements and paths the ray's take through the system; the rays that reach the Placido plane are given their intensities based on whether a bright ring (1) or a dark background (0) is hit. The respective intensities of each ray can then be mapped directly on the image plane. Figure 5.21a is an actual low resolution ray tracing simulation where the paths of each ray are represented graphically.

5.4.2 Diffuse and specular reflection

An ideal diffusely reflecting surface has the same luminance regardless of the observers angle of view. As rays are always backwards traced from the observer to

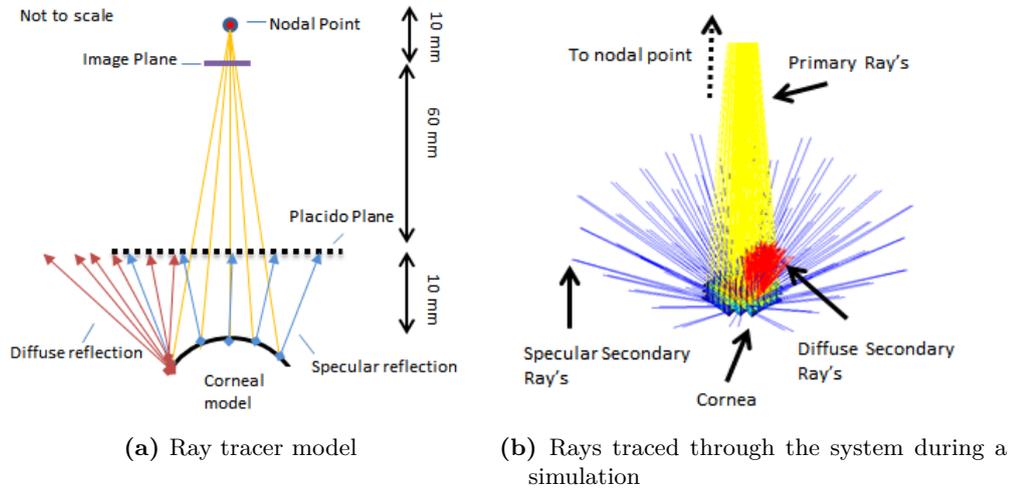


Figure 5.21: Ray tracer

the scene, the primary ray's intensity can be calculated as an average of the many secondary rays, which are spread in a hemisphere above the intersection point. However it is thought that an exposed cornea is not a completely diffuse surface, but one that has both specular and diffuse components as illustrated in Figure 5.22a. To approximate this mixed model in a backwards ray tracer, the majority of the secondary rays were shot out at a wide angle, but a single ray in the specular direction was added, see Figure 5.22b. Heavier weighting was applied to the specular ray than to the sum of the diffuse rays. The formula used to provide the final luminance, L_{final} , for a diffusely reflecting surface is listed in Equation 5.19, where w_{spec} and w_{diff} are the weightings applied to the specular and diffuse components of the model, L_{spec} is the luminance of the specular ray, $L_{i,diff}$ is the luminance of each individual diffuse ray and n_{diff} is the number of diffuse rays spread over an angle of Ω above the intersection point.

$$L_{final} = w_{spec}(L_{spec}) + w_{diff} \left(\frac{\sum_{i=1}^{n_{diff}} (L_{i,diff})}{n_{diff}} \right) \quad (5.19)$$

$$L_{i,diff} = \left(\frac{a_{surface} - (a - 1.5)}{1.5} \right) \quad (5.20)$$

In the tear dimple model, previously described in subsection 5.1.3.2 on page 176, no mention is made of where corneal scattering would begin. It could be assumed that only the deepest point on the tear dimple is where the corneal surface is exposed, however anecdotal evidence from tear film break-up images suggests that “blurring” occurs over a much larger area. Consequently the author has decided that for any tear dimple depth a , the height of the surface, $a_{surface}$, that will use this mixed model must be within $1.5\mu m$ of the deepest point on the dimple. Moreover, the relative luminance value $L_{i,diff}$ given to any diffuse ray that hits a “ring” on the Placido disk will decrease linearly from 1 at the deepest point on the dimple, to 0 at $1.5\mu m$ as described by Equation 5.20. All other rays from surfaces not falling within the respective diffuse zone will be specularly traced.

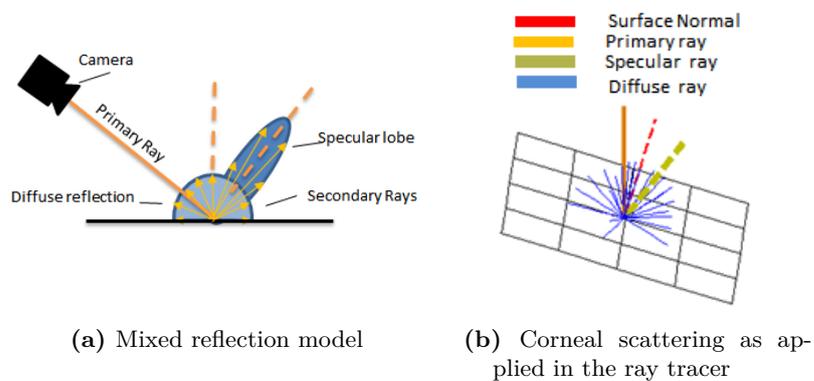


Figure 5.22: Diffuse and specular reflection

5.4.3 Simulation Results and Discussion

Thirty two simulations were run using the parameters listed in Table 5.2, sixteen of which used the ray tracing technique with both specular and diffuse components, and the remainder using specular components only. Both the depth and half width of the tear breakdown were varied between simulations. The cornea model used was identical to the numerical approximation cornea described in subsection 5.3.2 on page 190.

Table 5.2: Simulation Parameters

Parameter(s)	Value(s)	Description
Cornea		
r_0	7.87 mm	Central radius of cornea
x_{min}, y_{min}	-5.5 mm	Minimum x and y coordinates
x_{max}, y_{max}	5.5 mm	Maximum x and y coordinates
res_{model}	0.01 mm	Resolution in the x and y direction
Tear break-up		
a	1.5, 2, 2.5, 3 μm	Break-up depth
p_{hw}	0.1, 0.2, 0.3, 0.4 mm	Half width of the break-up
x_0, y_0	3.8, 2.0 mm	Tear break up centre, x y coordinates.
Ray tracer		
Ω	0.37 sr	Diffuse spread
n_{diff}	16	Number of diffuse rays
n_{rings}	20	Number of Placido rings
w_{diff}	0.9	Weighting given to diffuse rays
w_{spec}	0.1	Weighting given to specular rays

Figure 5.24 and Figure 5.23 contain the simulation results for the mixed and specular ray tracing models. Looking specifically at the shape of the specular only simulations, it appears that bulbous distortions are caused within the rings on the image plane. As would be expected, the bulbous distortion increases as the depth of the tear break up increases, this is primarily caused by the deeper gradients of the surface normals further distorting the incoming rays. For localized tear breaks with small widths ($p_{hw} = 0.1mm$), the gradient distortion is so large that the rings can experience discontinuity's along their perimeter. Increasing the width of tear break-ups reduces this background encroachment on the rings, and increases the overall size of the bulbous distortion up to $p_{hw} = 0.3mm$, beyond this the larger

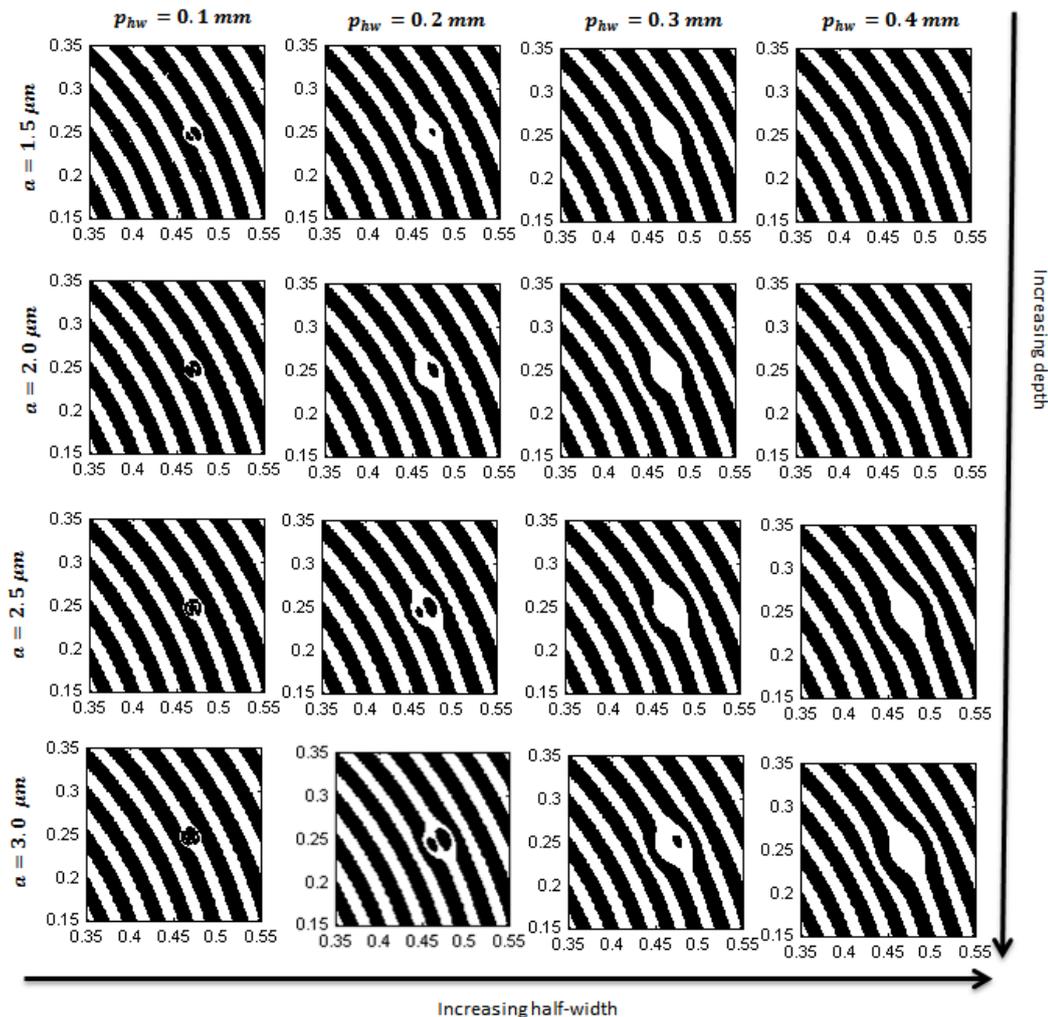


Figure 5.23: Specular only simulation results

widths cause smaller ring distortions albeit over a much larger area.

Adding a diffuse component to the simulations as seen in Figure 5.24 causes small low intensity islands to appear within the bulbous regions of the specularly reflected rays. It is not until the half-width of the break up increases to around 0.3-0.4 mm do the diffuse components of the image start to distort a wider area than the specular components. Thus it could be said that during a break up, where depths resulting from tear thinning will either decrease or remain constant, and the width of the break up will increase, the major distortions within the Placido images will change from a specular ring distortion to that of a diffuse blurring.

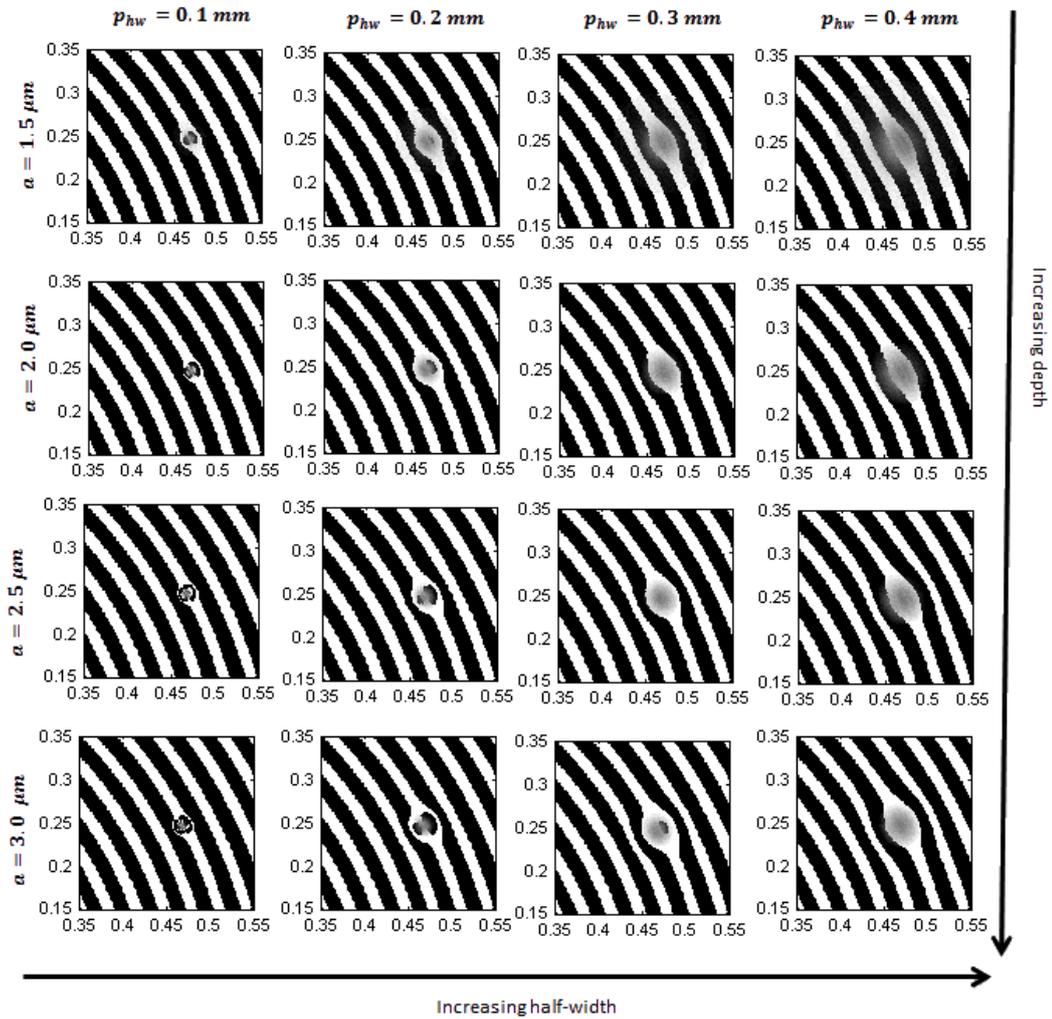


Figure 5.24: Diffuse and Specular simulation results

The specular results are somewhat in agreement with those of Alonso-Caneiro et al. [2011], who ray traced eleven rings within the area of the tear film break up, Figure 5.25, where a bulbous region can be seen between the second and sixth ring. Furthermore, distortion to the outer ring spacings was more severe than those to the inner ring spacings, which was replicated again in this study but not to the same extent, most likely a result of the new tear film model used. However none of the ring spacings appear to collide, as was the case in our study, and the choice of such a high density of rings in what is a small area of the cornea is unusual and may be exaggerating the effect of tear breakdown on Placido images.



Figure 5.25: Results of Alonso-Caneiro et al. [2011], who used a half-width of $p_{hw} = 1.8mm$ and depth of $a = 1\mu m$ taken from [Alonso-Caneiro et al., 2011]

The simulations predict bulbous like distortions appearing due to the small changes in surface profile associated with tear film breakdown. These resemble actual tear film break-down images produced using Placido disc devices, as seen in Figure 5.26. Although limited to predicting rotationally symmetric tear break downs, further work can be done to adapt the numerical approximation technique to those of randomly distorted tear break ups for more life-like simulations. We are also aware that while the diffuse scattering in our model is not based on the underlying physics behind light scatter, it does give a representation of the effect that real break-ups will have on Placido images.

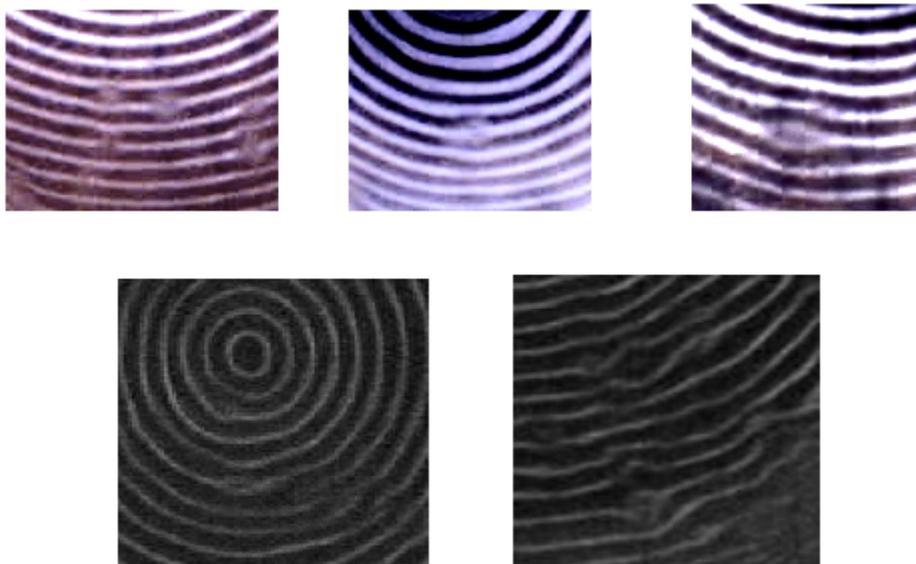


Figure 5.26: Bulbous like tear distortions in real Placido topographer images

For automated tear film measurements using a Placido disc topographer, these simulation results may mean that tear metrics that look at the position and shape of the rings for their values, as would be appropriate if only specular reflection occurred, may lose sensitivity when the break up increases in area and ring distortion decreases. Conversely, if tear film metrics were to concentrate on spotting the low intensity diffuse areas in images, they may miss the initial ring distortion that precedes full tear film break down. One of the challenges for automated tear analysis will be to combine these two conflicting distortions into one tear metric, although it must be said that quantifying ring distortions using image processing will be certainly easier than identifying diffuse scattering .

5.5 Conclusion

It can be seen that the new numerical approximation technique set out in this chapter can successfully orientate tear break-ups parallel with respect to the surface normal of the cornea. Moreover, the approach used can be flexible as it doesn't rely on underlying assumptions for the shape of the cornea, or the shape of the tear film breakdown; different cornea models and or break down profiles can be easily applied. Therefore, the author believes that not only does the aforementioned numerical approximation technique offer a more accurate representation of the geometry of tear film break up on the surface of the cornea, but also greater flexibility to future researchers that may want to adapt this technique to both new cornea and tear film models. For example, instrument designers may wish to use different tear break-up models to assess the impact of a deteriorating tear film on corneal topography images, and try and mitigate its impact.

Secondly, by including both specular and diffuse reflection within a custom made ray tracing system that simulates a Placido disc device, a more accurate representation of the effect that tear break-down has on Placido images was achieved. The simulations represented in this chapter therefore can be used to further develop custom tear metrics that are sensitive to ring distortion and will be applied to the analysis of in-vivo tear film placido disc reflection video images as part of chapter 6

6 Automated Assessment of Tear Film Stability

6.1 Introduction

Although many cases of dry eye are diagnosed and treated based on symptoms alone [Foulks and Jester, 2007], a set of diagnostic tests have emerged to evaluate the quality of the tear film. The most recent report from the international dry eye workshop [Bron et al., 2007] recommended a series (Table 6.1) for clinical evaluation . This is not an exhaustive list of tear film tests, but it does give a suggestion of the most popular and widely used tear film tests in clinical practice. It has been asserted that no single test is considered adequate enough to diagnose disorders of the tear film alone. This is confirmed by the accepted definition of dry eye which includes symptoms, tear film properties, and ocular surface effects. Each test generally examines a specific attribute of dry eye disorders with the results of several tests normally influencing a clinicians diagnosis [Pflugfelder et al., 1998].

The most widely practiced tear film tests, besides symptom questionnaires, are based primarily on the tear film break up [Korb, 2000, Smith et al., 2008]. This is a test through which the tear film is imaged either digitally or subjectively by a clinician, and the time taken from when the patient first blinks to when the tear film breaks up (or when the patient re-blinks) is recorded. The recorded time then gives an indication of the stability of the tear film. The shorter the time the less stable the tear film. Currently there are no standardizations of this test and it relies on the subjective judgment of the examiner, although a general consensus is widely followed by clinicians.

Recently objective testing methods have been developed to attempt to mitigate subjective errors in stability measurement, however as to yet no objective assessment of tear film stability has become widespread in practice [Smith et al., 2008]. This chapter will therefore review several invasive and non-invasive methods of determ-

ining tear film stability, evaluate different equipment and analysis methods that can be used for tear stability assessment, and propose a new objective assessment of tear stability using video tear reflection image analysis. It is hoped that by further researching and developing objective tear film methods, they may one day provide a quick and reliable alternative to subjective methods of assessment.

Table 6.1: Dry eye testing methods

Attribute	Tests
Tear secretion	Schirmer's test Thread methods
Tear clearance	Fluorescein clearance Tear Function Index Fluoro-photometry
Ocular Surface Damage	Fluorescein Rose Bengal. Lissamine green staining
Tear film stability	Fluorescein tear break up time (FBUT) Non-invasive break up (NIBUT) Video Keratometry (VK) Interferometry Wavefront aberrometry
Tear volume	Tear meniscus measurement
Lipid layer	Meibometry Meiobogrpahy
Tear Evaporation	Evaporimeter Closed chamber Ventilated chamber
Chemical properties	Osmolarity Freezing point Depression Conductivity Mucin analysis

6.1.1 Fluorescein tear film break up time

The most widely applied tear film stability assessment method is the fluorescein break-up time [Korb, 2000, Smith et al., 2008]. Sodium fluorescein is an orange dye which under certain blue light conditions fluoresces green due to the excitation of its molecules. Strips of paper infused with the dye can be applied to the eyes of patients, where the sodium fluorescein molecules will readily diffuse throughout the tear film but not into the cornea, making the tear film visible under a blue exciter

light. The use of a yellow filter, ideally with a cutoff wavelength of 500nm [Peterson et al., 2006] placed in front of the observation system, can exclude the reflected blue exciter light and enhance tear film observation. The fluorescein tear film break up time (FBUT) can then be observed by measuring the time after a blink that the first dry spots (seen as dark spots within the green tear film) are detected. FBUT has been shown to be reliable in cases of extreme dry eye as well as being cost and time efficient [Holly et al., 1986]; its popularity owes much to its cost effectiveness, simplicity and familiarity amongst clinicians.

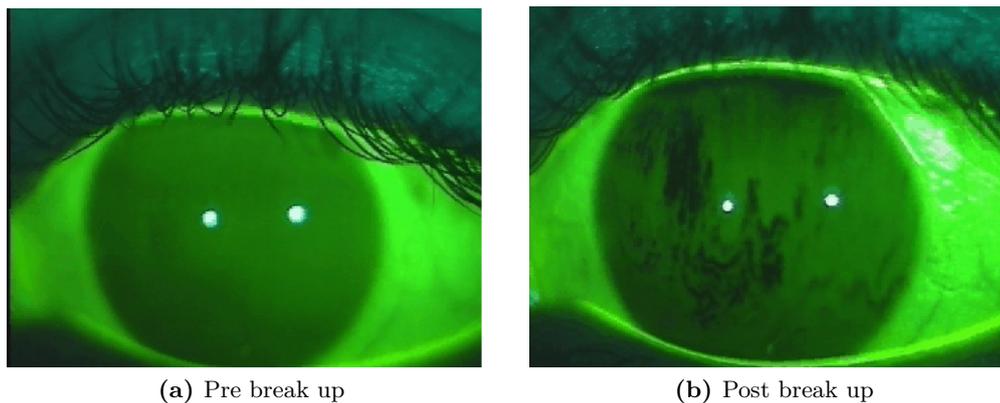


Figure 6.1: Fluorescein staining pre and post break up

There are several noticeable drawbacks concerning the use of fluorescein. Primarily being an invasive procedure, adding a diluted dye can modify the tear film volume and stability. It has been reported that this may fundamentally change some of the physical properties of the tear film altering the stability and skewing break up test results [Mengher et al., 1985]. As with any examination of an invasive nature, patient discomfort may be greater in comparison to non-invasive methods, and the strips used to implement the dye may promote reflex tearing. In addition, the application, volume, acidity and make up of any staining solution can vary between tests and subjects, although this can be mitigated by using sealed strips and careful application of saline.

Recently efforts have been made to automate the fluorescein FBUT and make it an objective test [Yedidya et al., 2007b, 2008, Cebreiro et al., 2011]. As with most eye imaging algorithms, the iris diameter and centre need to be found. However during fluorescein imaging the pupil is partly masked by the fluorescein and the iris-sclera boundary is blurred significantly. Yedidya et al. [2007b] invented an algorithm to

detect the eye lids, first by using the Sobel operator at a high threshold and then circle fitting to find the upper and lower part respectively. Once any pixels outside the eyelids were masked, RANSAC [Hartley, 2003] (an iterative fitting method) could be used to detect the iris centre and dimensions in the first image. As the device used to take fluorescein images was hand held, image translation was used to steady images that otherwise would have been blurred. Further work by the team [Yedidya et al., 2008] took the average intensity for all of the pixels inside the segmented area in the first frame, and then compared pixel intensities in further frames to this reference. That way a drop off in intensity for each individual pixel could be found as an indication of break up in the image. Once a pre-determined number of the pixels had broken up (decreased to a threshold level) an estimation of the FBUT could be found (see Figure 6.2 for an illustration). Similarly Cebreiro et al. [2011] used this method and found that the FBUT error interval of 2.5 secs was within the same range as experts themselves.

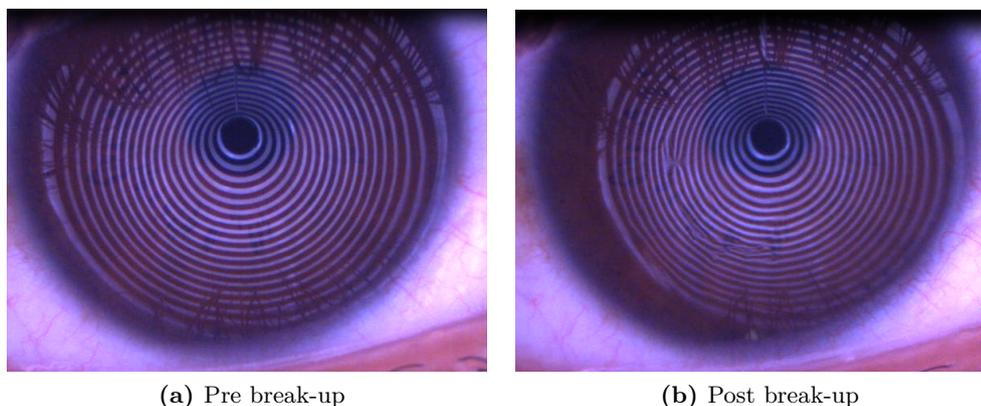


Figure 6.2: Objective fluorescein imaging taken from [Yedidya et al., 2007a]

6.1.2 Non-invasive break up time

To overcome the potential limitations of using fluorescein dye for tear film assessment, non invasive break up time (NIBUT) assessment was developed. One of the first instruments designed to test NIBUT was the Tearscope Plus [Guillon, 1998], a rather bulky slit lamp attachment that could provide a 180 degree diffuse light source for illuminating the cornea. For NIBUT assessment, a ringed insert could be placed in front of the light source to be specularly reflected off the cornea. Once imaged through a slit-lamp, the ringed pattern could be seen clearly (Figure 6.3). Any perturbations of the reflected grid could be taken as a break within the tear film,

the duration between when a patient blinks to the appearance of perturbations in the reflected grid, as recorded on the integrated manual timer, was taken as the non-invasive break up time. As video topographers also use the same principle to image Placido rings, see subsection 5.1.4 on page 179 for an explanation, they can be used in lieu of the Tearscope for NIBUT assessment (Figure 6.3).



(a) Pre break-up (b) Post break-up

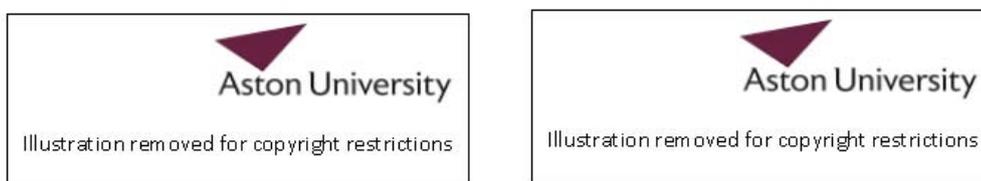
Figure 6.3: Tearscope mires pre and post break-up

NIBUT was found to be repeatable [Elliott et al., 1998] and to correlate highly with existing FBUT measures [Pult and Riede-Pult, 2012] but tends to give a longer break up time [Sweeney et al., 2013], which suggests fluorescein instillation does skew break up results. As such NIBUT has been used widely in research as a non-invasive replacement for FBUT but its use in a clinical environment however has been minimal [Smith et al., 2008] due to the ease of viewing the tearfilm over the whole of the cornea with fluorescein testing. Whilst it may be a step forward to be using a non-invasive measurement NIBUT still relies on subjective assessment by the examiner. Researchers have thus looked towards creating objective non-invasive measures of tear stability assessment, most notably using interferometry and video keratometry that rely on the reflectance spectra of the tear film.

6.1.3 Interferometry

Interferometry is a technique that uses the principle of constructive and destructive interference of light to characterize the underlying properties of surfaces. When imaging the tear film with instruments that use a wide band light source such as the Tearscope Plus or the KOWA DR1 tear camera, interference fringes can be clearly seen coming from the tear film layer (Figure 6.4). The fringes are a product of light

interacting from the front and rear surface of the lipid layer, this is often called thin film interference and occurs when the difference in refractive index between layers is large.



(a) Tearscope Plus taken from [Remeseiro et al., 2013] (b) DR1 tear interference camera taken from [Goto, 2004]

Figure 6.4: Lipid layer interference fringes as observed on two instruments

Guillon [1998] proposed a grading system for the interference fringes observed with the tearscope images, that implied that the thickness of the lipid and its stability are directly related to the interference fringes observed. Again this relied heavily on the subjective judgment of examiners. With the creation of the later DR1 interference camera that successfully eliminated reflectance spectra from anywhere but the lipid layer [Khamene et al., 2000], researchers have gone on to objectively analyze the interference fringes using computer methods. Goto [2003] first quantified the colors of the interference fringes and related them to specific lipid thicknesses, which he later used to objectively analyze tear stability and lipid spread time [Goto and Tseng, 2003a,b]. The lipid spread time in particular could be used to differentiate between different tear disorders. Despite this the high cost, limited field of view ($<8mm^2$) and the complicated analysis needed for tear interference images has hampered its use outside of the research laboratory.

Narrow band shearing interferometry techniques that use a coherent light source of only one wavelength, have also been used to image the tear film. Objective analysis techniques have been developed to analyse the resulting patterns (Figure 6.5) from the interferometer, which normally consist of looking at the disturbance of the horizontal lines as an indication of the stability of the tear film. It has been claimed that lateral shearing interferometry can fully characterize tear stability from tear

build up stage all the way to the tear break up [Szczena and Iskander, 2010]. In addition, in a study where multiple objective tear stability techniques were evaluated, lateral shearing interferometry stability measures were the best indication of dry eye disease above those of wave-front aberrometry and video keratography [Szczena et al., 2011]. The practicality of narrow band interferometry techniques is in question however as it is extremely sensitive to head and eye movements and as such requires head restraints and bite bars which limits its clinical use [Dubra et al., 2004].



(a) During the stable inter-blink phase

(b) During the break-up stage

Figure 6.5: Lateral shearing interferometry images taken from [Dubra et al., 2004]

6.1.4 Video Keratography tear stability measures

6.1.4.1 Topographic metrics

It had been noticed by researchers that some of the errors in topographical mapping could be attributed to tear film deterioration [Cronje-Dunn and Harris, 1996]. Experimenters soon realized that by using the metrics provided by corneal topography over a series of images, a tear film break up time could be deduced reliably. Research teams soon modified existing PD based corneal topographers, and took advantage of the corneal topographic maps provided to estimate new metrics to measure tear film stability. One such early system was named the TSAS (Tear stability analysis system)[Goto et al., 2003], which recorded an image every second for 10 seconds. A change in the corneal topography map of 0.5 diopters was taken as a break up in the tear film. Figure 6.6 shows the corneal topography map at $t=1$ sec and at $t=5$ secs,

the change in the map can be attributed to a break down in the tear film. They found that this method correlated strongly ($p>0.81$) with FBUT.



(a) Tear stability map taken after a blink (b) Topography map taken after a blink

Figure 6.6: Tear stability maps using topographical data taken from [Goto et al., 2003]

This approach, first of producing topographical maps and then analyzing the data was used several times to find metrics that could actively measure tear stability. Analysis of topographical aberrations [Montés-Micó et al., 2005], RMS wavefront error [Montes-Mico, 2004], topographic indices such as SRI and SAI [Németh et al., 2002, Kojima et al., 2004] have all followed. SRI and SAI in particular showed promise in differentiating between types of dry eye [Kojima et al., 2004].

Topographic metrics however have their own flaws. When the Placido rings becomes locally irregular inaccurate topographical maps can be constructed, hence software designed primarily for corneal topography will try and minimize any interference by removing the data from analysis. This is generally to remove erroneous data from eyelash shadows and lid movement but could easily exclude tear film break up data. In addition topographic measures are extremely sensitive to the micro-movement of eyes [Buehren et al., 2002] which can cause false tear film break ups to be reported.

6.1.4.2 Custom tear analysis

To overcome some of the limitations of using topographic metrics for tear film stability measures, custom image processing algorithms were developed that worked

on the raw images produced by video keratographs. Iskander et al. [2005] first developed such a system that took the original image data straight from a corneal topographer, and used an algorithm that related the limbus position to that of the instruments axis, in order to reduce error from micro movements. The TFBUT could be estimated as the time in which the quality of the ring pattern in the acquired PD topographer degrades. A set of 256 intensity profiles centred on the axis of the keratograph and spread out radially in equal angular steps were sampled. These intensity profiles could then be used to reconstruct an image into polar coordinate representation as seen in Figure 6.7. The Marr Hildreth edge detection routine was applied to produce a binary image ready for analysis. The variance in the number of rings detected in each intensity profile was used as an indicator of the quality of tear film. Once a noticeable rise in variance existed, this was to be taken as the BUT. Unfortunately this algorithm was described but not tested on patients.



Figure 6.7: Image processing steps undertaken by [Iskander et al., 2005]

The advantage of using statistical metrics on the initial ring image was that the effect of ocular micro-fluctuations could be reduced. For rings covering the pupil where the contrast of the rings was fairly stable, the algorithm worked successfully. But for the rings overlapping the iris, where the contrast had lowered, edge detection proved difficult. A subsequent paper [Kopf et al., 2008] from the same team used background data subtraction to greatly help in the edge detection process. The team fitted polynomials to each of the intensity profiles generated and then subtracted from the original data prior to edge detection (see Figure 6.8). The data after subtraction became independent of pupil size. Using the ring number variance as before, they found the average variance could be used to differentiate patients tear quality before and after the use of contact lenses but it correlated poorly with FBUT

($r=0.22$). It appears that the idea of an automated break up time was abandoned most likely as image noise would mask a rise in the variance or cause false positives to occur.

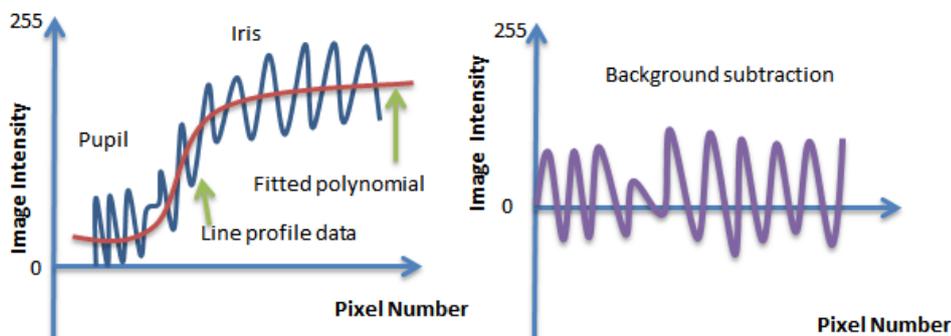


Figure 6.8: Background data subtraction to increase efficiency of edge detection algorithms

Gumus et al. [2011] furthered this technique in the tear stability analysis system. The team noticed that peripheral rings were lower in intensity due to the reflection and color of the iris so a weighted system was used to normalize the periphery rings. Recognition of the break up during the line profile, was provided by assessing the difference in the area underneath each curve. Once the difference reached a designated threshold the point was said to have broken up. The total number of points identified was called the ring break-up (RBU) sum. If the RBU sum exceeded a certain threshold this was taken as the BUT. Break up times were found to be significantly lower than those found with fluorescein staining, sensitivity (80%) and specificity (60%) to dry eye were poor and the noise of the RBU sum was high as a result of lid movement affecting results. An increase in noise from erroneous data may be tripping the break up condition which would explain the lower break up time values, however correlation and agreement between FBUT and the objective BUT was not stated.

All of the different methods mentioned previously have used a fixed region of interest centered at the apex of the cornea. To ensure that the majority of lids or eyelash artefacts are excluded in all patients, the fixed analysis area must remain small and is limited to a few central rings; tear breakups occurring outside the centre will be missed. This may cause a problem as the inferior cornea is the place where break-ups are more likely to occur[Fenner and Tong, 2013]. Alonso-Caneiro [2008] first created a method to separate interference from eyelashes and tear degradation

to maximize the analysis area for corneal topography in each video frame. The use of frequency normalization, gabor filters and statistical decision making were novel in their application to provide dynamic ROI's (region of interest) to Placido based devices. A Subsequent paper [Alonso-Caneiro, 2009] focused on using these processing algorithms to classify the different interferences (tear or eyelash) observed and provide new metrics for tear film evaluation. They found their new tear surface quality metric could discriminate between patients with bare eyes and those that had worn contact lenses.

It appears that tear film surface metrics based on Placido disk devices can be used to discriminate between different tear conditions successfully, however there are contradictory results about how these metrics can be used for stability measurement. Goto et al. [2003] found their automated BUT assessment to correlate highly with FBUT, meanwhile Best et al. [2012] in an independent study on a different device found automated BUT to have poor agreement with NIBUT despite the descriptions of both systems used being broadly similar. What is unclear is how much of the automated BUT relies on the inter-blink period for its agreement with FBUT and NIBUT not the assessment of the tear film surface itself, for example if the threshold that triggers the BUT is set low, erroneous break-ups may be observed leading to poor correlation with NIBUT; however if the threshold is set too high any break ups will not be detected and the BUT will be defined by the inter-blink period. In the latter case the image analysis may be holding little sway over the resulting automated BUT.

6.1.5 Outcome

Assessing tear film stability with the subjective fluorescein FBUT is the most commonly used procedure. FBUT appears to be an appropriate measure of dry eye as many of the symptoms occur due to premature exposure of the corneal surface. The popularity of the fluorescein test owes much to its cost effectiveness, simplicity and familiarity with clinicians. Automation of the invasive fluorescein FBUT however requires additional capital cost, in the form of computation, and an extra layer of complexity, without negating any of the invasive effects of fluorescein instillation.

A more favourable approach would be to use non-invasive methods for TFBUT and stability measures, to create an objective automated tear film measure. Standard narrow band interferometry whilst giving the most accurate readings for stability measurements [Szczena et al., 2011] is impractical for use in a clinical setting; it

requires restraints to limit head and eye movement and can only assess a limited area of the corneal surface (around $4mm^2$). Diffuse lipid layer imaging (wide band interferometry) using such systems as the DR1 interference camera can also be used successfully to assess tear film stability in patients. Semi automation has already taken place [Goto, 2003], but limited analysis areas ($8mm^2$) expensive equipment and complexity may limit its clinical application.

Video keratotomy can provide a balance between precision tear film stability measurements and clinical application. Many Placido based devices have already been brought to market for corneal topography and clinicians are well accustomed to them. Systems that have used the changes in topographical maps, rely on estimators that assume that the reflectance of the rings occurs from a high quality tear film [Mejia-Barbosa and Malacara-Hernandez, 2001] and hence are not tuned for tear film analysis. They are also very sensitive to micro-fluctuations of the eye. Analysis of raw Placido Disk based images may provide more effective and sensitive metrics to the quality of the tear film surface [Szczena-Iskander and Iskander, 2012, Alonso-Caneiro et al., 2011] than those based on topographical maps. Dynamic analysis areas have also improved to cover much of the cornea that traditional topographers would miss. Already in chapter 5, the effect that tear film break down can have on raw Placido disk images and this insight can be used to build upon the work of current tear analysis systems to help improve the reliability of objective tear film measures.

6.2 Aims and Objectives

6.2.1 Aim

To create a new and novel image processing algorithm that can accurately measure tear film stability and surface quality objectively.

6.2.2 Objectives

1. Selection of a suitable Placido disc device.
2. Design and implement a full image processing routine to be used with the Placido Disc device for automated assessment of the tear film surface.
3. Test the new tear film surface metrics validity and repeatability in comparison with current non invasive measures.

6.3 Hardware Development

Selection of a suitable device was crucial to the design of the author's tear analysis system. First to be explored was the idea of creating a custom device specifically suited to tear analysis. It was reasoned that current tear analysis systems did not make use of all the data available to them, for example most topographers recorded in black and white missing any chromaticity differences that could be used to separate tear interferences from those of eyelash shadows. In addition, by creating a custom device certain characteristics such as number of rings, ring luminance and frame rate could all be controlled. The following sections describe the creation of such a device and some of the manufacturing difficulties that resulted in this approach not being pursued.

6.3.1 Design and manufacture

The concept was to build a small Placido cone type device that would sit into the eye socket. This design allowed the device to remain portable and compact whilst still providing an adequate Placido target for reflection off the cornea. Secondly, by fitting ergonomically into the eye any obstructions due to facial features on the target could be reduced. Several other designs were considered that sat away from the eye but the size of the Placido grid needed was deemed too large for such a compact device.

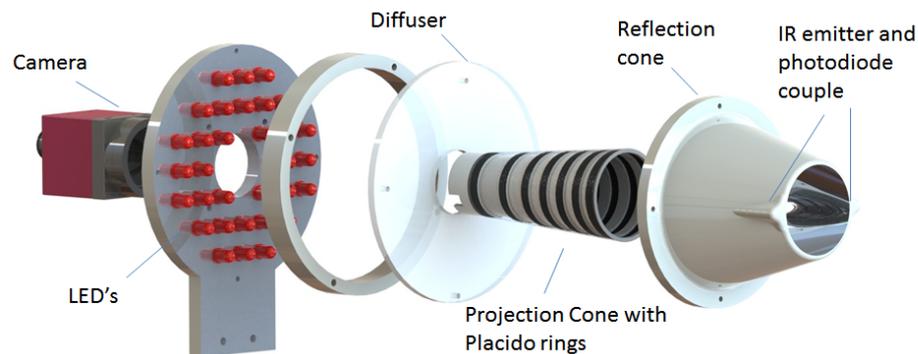


Figure 6.9: Placido Disc Concept

The principles of this design consisted of an illumination system that would disperse light evenly throughout the Placido cone, reflect off the tear film surface, and the resulting Purkinje image to fall on an image sensor placed on the instrument axis. In

addition an infra-red emitter and photo-diode pair placed at the opening of the cone could be used to position a patients cornea at the required distance. A summary and specification for the major components of the device can be viewed in Table 6.2.

Table 6.2: Device summary

Component	Specification	Value
Camera	Sensor:	1/3" color CMOS
	Resolution:	1280x1024 pixels
	Max FPS:	25 fps
Lens	Focal Length:	16mm
	F-Number:	f/2.8
	Mount:	S-mount
Illumination	LED's:	6 x OSRAM Golden dragon
	Color Temp:	4000K
	Luminance:	20,500 mcd
	Viewing Angle:	120°
	LED Driver:	THORLABS current limiting (1.2A) driver
Casings/Mounts	Diffuser:	Laser cut acrylic "Pearl" diffusing sheet
	Inner cone:	Sandblasted extruded acrylic, turned on lathe
	Outer Cone:	SLA prototyped. Coated in Barium sulphate
	LED Mount:	Laser cut Aluminum
Placido Grid	No of Rings:	15
	Manufacture:	Turned on lathe
Cornea Detector	LED:	875nm Peak, 1800 mW/sr radiant intensity
	Photo Diode:	820nm Peak
	Amplifier:	LTC1050 chopper - transimpedance configuration

6.3.1.1 Illumination system

LED's were used to illuminate the Placido disc target due to their small dimensions, low cost and high flexibility in comparison to other light sources. However as a point type light source an appropriate diffuser would have to be used to avoid "hot" spots appearing in images. It was decided that an acrylic diffusing filter set away from the top of the flat LED would be sufficient to avoid hot spots and achieve homogeneous illumination of the cone. The outer cone was shaped to fit into the eye socket whilst still being able to project light onto the outer rings. The choice of using white illumination was taken so the device could be used to see "colour" fringes in the tear film (thin film interferometry) although bright white illumination is known to cause reflex tearing.

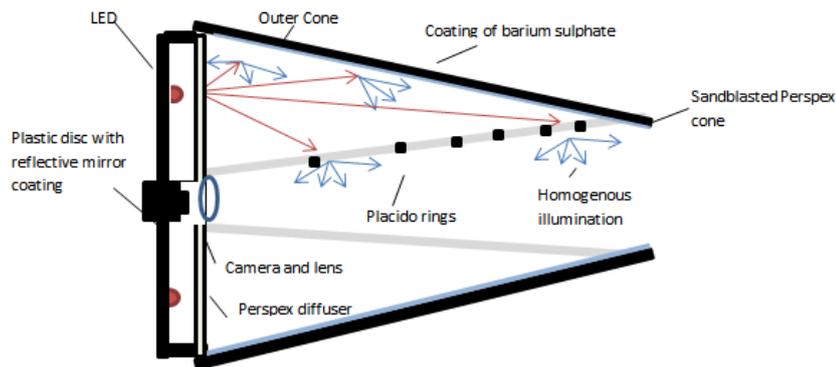
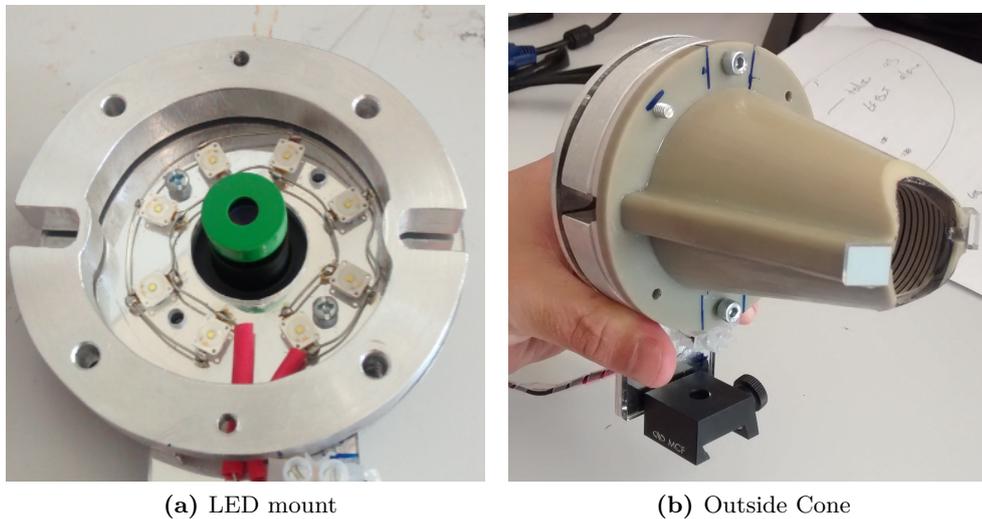


Figure 6.10: Illumination System

A ring of eight bright white LED's were mounted onto a metal disc that was placed at the rear of the device and controlled by a current limited LED driver (Figure 6.11a). The large viewing angle and high intensity reduced the quantity of LED's needed. Above the LED's rested a acrylic light diffuser to scatter light evenly across the Placido disc. The outer cone (Figure 6.11b) was rapidly prototyped using a stereolithography machine and coated in a layer of paint that contained large amounts of barium sulphate. Barium sulphate scatters incident light that falls on the outer cone to achieve more homogenous illumination. In the center of the device a sandblasted Perspex cone was used to provide further dispersion.



(a) LED mount

(b) Outside Cone

Figure 6.11: Illumination System

6.3.1.2 Corneal displacement measurement

To ensure the correct working distance was used on the device a simple beam break detector was incorporated at the nose of the illumination cone. Two hollow tubes (Figure 6.12) that ran parallel to the sides of the illumination cone housed an LED and a photo diode. The light from the LED ran through the tube was reflected at the mirrors and cut across the Placido disc opening, it was then reflected down the opposite hollow tube and into the photo diode. If the opening of the Placido cone was completely unobstructed, a large response would be registered on the detector and a message sent to the examiner that the device should be brought closer to the eye, conversely if there was little or no response on the photo diode it could be assumed that the cornea was completely blocking the path and a warning that the device had to be brought further away. The ideal response was when only half of the beam was being intruded upon.

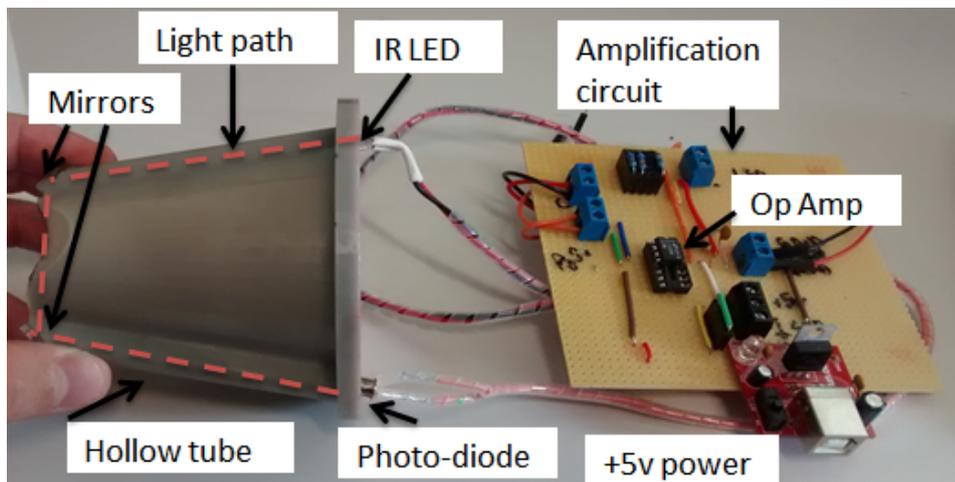


Figure 6.12: Beam detector setup

Figure 6.12 is the final beam breaker design which consisted of an infra-red LED and photo diode pair coupled with an trans-impedance amplifier. Photo-diode current was directly proportional to the amount of photons absorbed by its sensor, the trans-impedance amplifier circuit was used to convert the current into an amplified voltage which could be fed into a digital acquisition device.

6.3.1.3 Placido Target and Imaging system

Two optical designs could be used on the placido device, a telecentric (Figure 6.13a) and single lens (Figure 6.13b) design. Many corneal topographers make use of a telecentric lens for easier topography analysis as only light that runs parallel to the instrument axis is imaged, the drawback being that the lens must at least be as large as the object it is imaging. For a compact Placido device this would not allow rings to be positioned on the central cornea as the reflection of the lens would take a disproportionate amount of space. Thus a standard single lens design was chosen for the device to give greater ring coverage.



Figure 6.13: Imaging system designs taken from [Mejia-Barbosa and Malacara-Hernandez, 2001]

A lens with an F number of 2.8 and focal length of 16 mm was coupled with a 1/3" compact board level CMOS camera that could record images at 25 Hz. A simple ray tracing program in MATLAB was set up to determine the position of the rings on the inside of the Placido cone so they would appear appear equally spaced on the imaging sensor (Figure 6.14). As the rings targeted for the outer edges of the cornea corresponded to small ring spacings on the Placido cone, this would later become a significant problem when attempting to manufacture such a cone.

Extruded acrylic block was used as the base material for the Placido cone due to its transparency and cost. The block was placed on a lathe and turned to the required profile inside and out, due to the brittle nature of acrylic and the complexity of the shape each cone took approximately 3 days to manufacture. Once the cone had

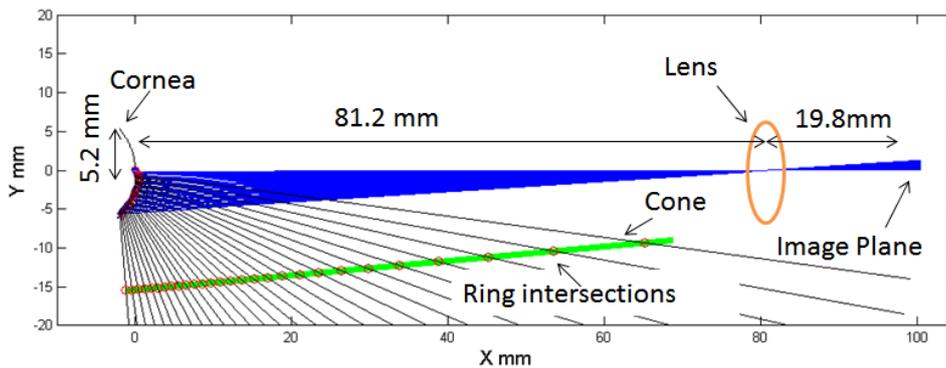


Figure 6.14: Ring spacing determination

been turned the inside was primed with plastic undercoat and sprayed black. The idea was to use the lathe to scribe out the black paint to leave the transparent cone underneath, hence producing the Placido rings. Painting however proved difficult as the cone could not be taken from the lathe; if the cone was taken out from the lathe and then reset it would never run truly concentric again.

Several attempts at creating the Placido rings took place however the small ring spacings required could not be manufactured accurately using a manual lathe despite the use of a fine cutting tool. In addition repeated turning caused fracture cracks to appear and the acrylic cone would often break up. Other manufacturing ideas such as casting the cone in less brittle material were tried but creating the rings still proved difficult. A picture of one of the erroneous cones manufactured is illustrated in Figure 6.15b, meanwhile a cone being turned on the lathe is shown in Figure 6.15a. To produce a cone with accurate ring spacings it would require the use of CNC (computer numerical control) lathe.

After several failed attempts it was decided to use a commercially available corneal topographer as the basis of the tear analysis system. The Medmont E300 corneal topographer was selected as it had a substantial number of rings (30), a large analysis area as it could be positioned within the eye socket and could record at a fast frame rate (25 Hz), but the idea of using colour data to supplement tear analysis had to be abandoned.



(a) Placido cone being turned on the lathe



(b) Erroneous Placido cone

Figure 6.15: Placido cone manufacture

6.4 Development

The concept behind the new tear analysis system developed is illustrated in Figure 6.16. Firstly, blinks in the video would have to be identified and any frames outside the inter-blink period be discarded. Following on from blink detection the inner ring of the Placido disc would have to be tracked as micro movement from eyes shifts the Pukinje image of the rings from the centre. Once the Placido disc was properly centered in each frame, two ROI algorithms would determine the ROI's to be used over the whole video, the first to discard data artifacts resulting from the eye lids and the second to remove eyelash shadows and spurious interference. Finally data could be ready for the tear film metric to be applied.

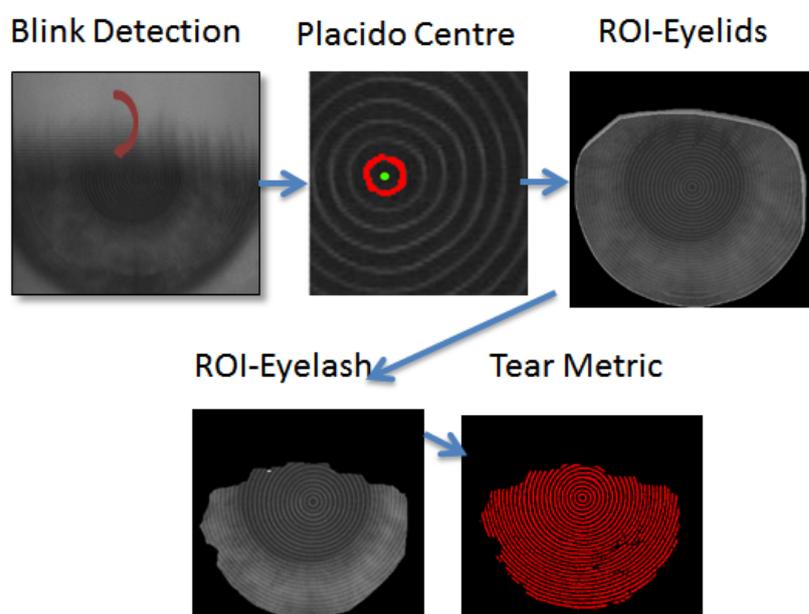


Figure 6.16: Software concept for custom tear analysis

6.4.1 Blink detection

Blink detection was achieved by taking the raw image frames produced by the topographer and assessing their statistical properties. It was found that during a blink the standard deviation of pixel intensities increased, when plotting this metric against time peaks can be clearly seen in the resulting graph. Using a peak detection algorithm provided by LABVIEW, the blinks could be found and image frames residing outside the inter-blink time excluded from analysis.

The threshold for the size of the peak would have to individually match each patients video, as the standard deviation of the pixels could not be known in advance. To start with the peak threshold was set so low that hundreds of peaks would be found in the graph; when the threshold was slowly increased there would be a certain point where the majority of erroneous blinks would disappear abruptly, leaving only true blinks. Once a threshold value had found less than twelve blink affected frames and the largest duration between the two blink frames was at least 40 frames long the inter-blink time could be determined. Figure 6.17a and Figure 6.17b show the blink detection on two separate videos.

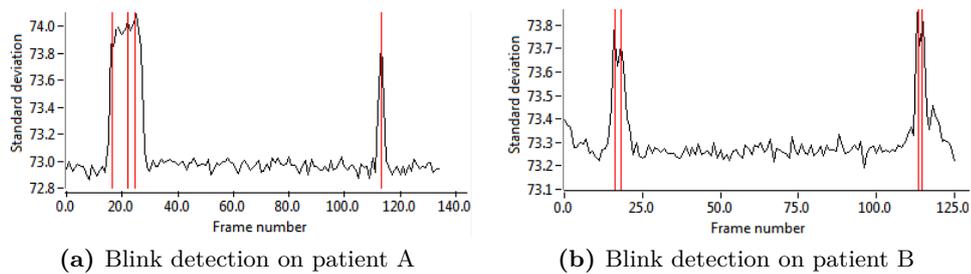


Figure 6.17: Pixel standard deviation plotted against frame number for two subjects.

6.4.2 Centre detection

One of the effects of tear film break up on images, as simulated in chapter 5, is that the spacings between rings can become erratic. If a tear film metric required the distances between rings to be known, the location of the Placido centre would be beneficial. However, raw images from the corneal topographer are not completely centred with respect to the reflected Placido rings; micro-movements of the eye and changes in gaze during the inter-blink period require the image to be continually readjusted.

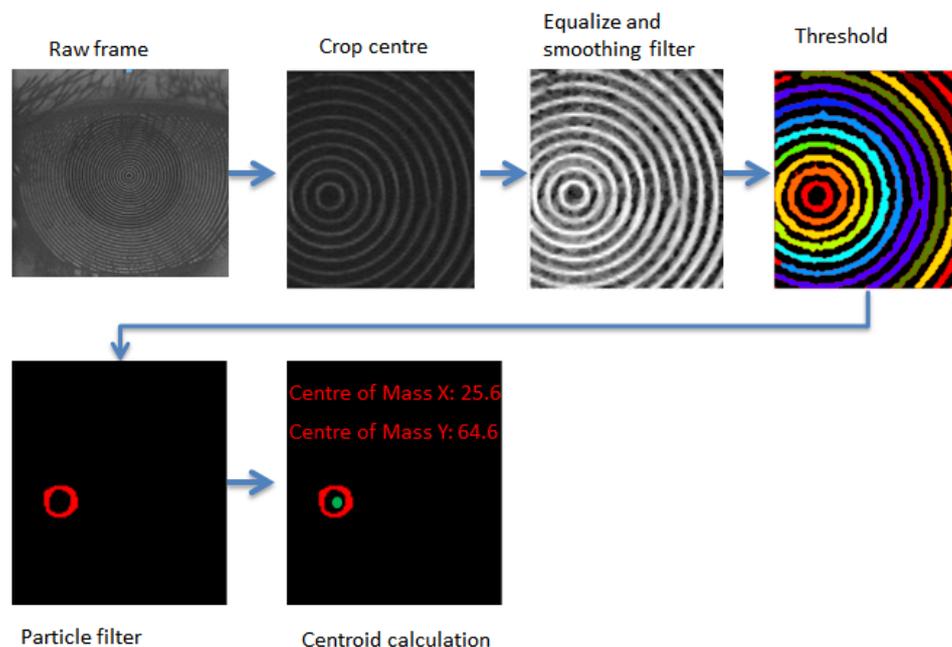


Figure 6.18: Image processing steps in center detection

The major image processing steps used to determine the innermost ring are illustrated in Figure 6.18. Raw images were cropped, histogram equalized and then a threshold was applied to isolate the rings from the background data. Each separate particle in the image (denoted by a separate colour) was classified using metrics such as particle elongation, number of holes and holes area. A filter could then be applied to discard particles that did not have the characteristics that would be expected of the innermost Placido ring. If no ring was found in the image, the analysis would be repeated with a change in threshold and smoothing filter parameters; see Table 6.3 for a summary of the values used to filter out erroneous particles and arrive at the innermost ring. After three analysis attempts if no ring was found the image was

discarded and would no longer be used, conversely if a ring had been found the geometric centroid would be taken as the Placido centre and the image would be reset within its frame.

Table 6.3: Image processing values used in the center detection

Parameter/Characteristics	Values	Repeat 1	Repeat 2
Raw Image Size (px):	645x576	-	-
Cropped Image Size (px):	114x131	-	-
Smoothing filter			
Type:	Median	Mean	Mean
Kernel size:	3x3	3x3	3x3
Threshold			
Type:	Niblack localized [Niblack, 1985]	Global	Global
Threshold value (8-bit value):	-	179	202
Kernel size:	32x32	-	-
Particle filter			
Bounding rectangle diagonal (px):	<23	-	-
Elongation factor:	0.85-1.15	-	-
No of holes:	1	-	-
Holes area (px^2):	100-260	-	-

6.4.3 ROI - Stage 1, Eye lid removal

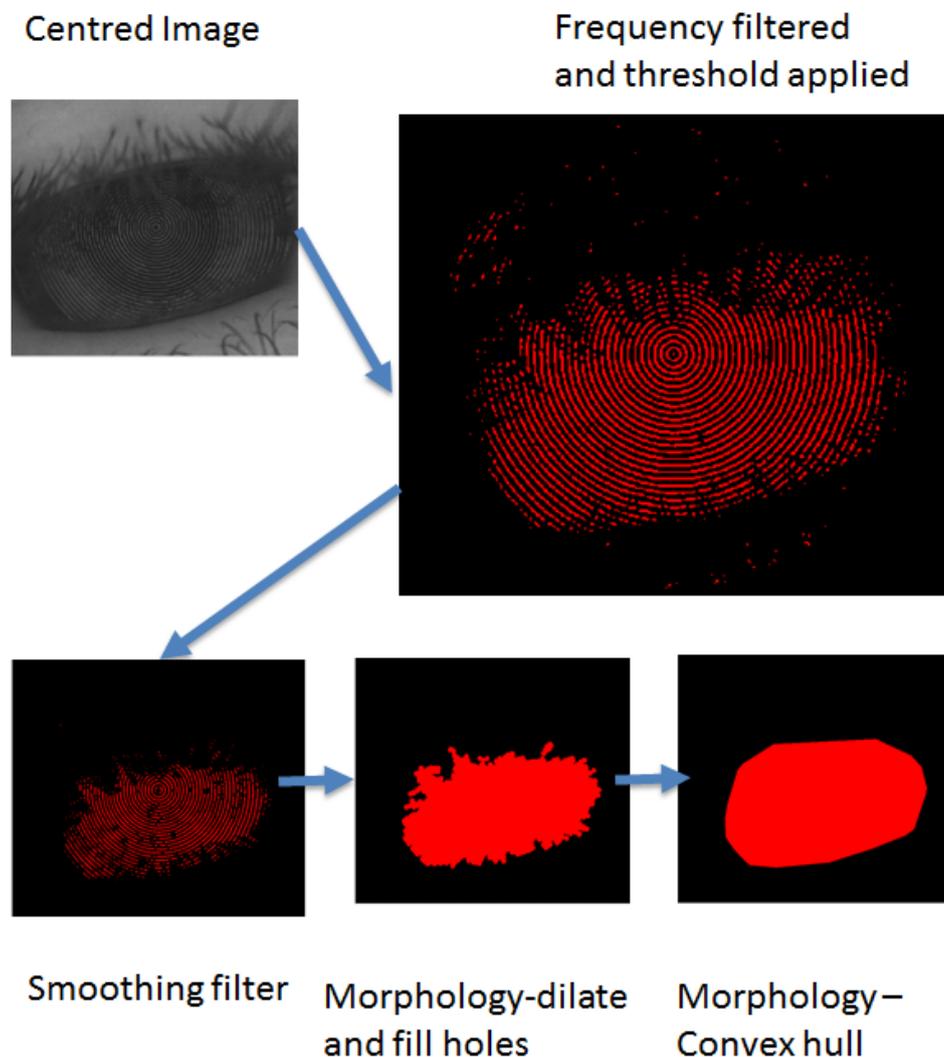


Figure 6.19: Image processing steps to exclude eyelids

Of the whole tear analysis system the method in which erroneous data from eyelashes and eyelid movement were separated from the Placido ring data could be considered novel. Previous methods had taken each individual frame in a video and analyzed it separately for creating dynamic regions of interest (ROI) that exclude eye lash data [Alonso-Caneiro, 2009, 2008]. Whilst this may give you the maximum ROI for any given frame, the differences in size and shape of the ROI across the frames may increase noise in the final metric. Furthermore significant lid movement during

the inter-blink period could be unveiling areas of the cornea that in previous frames were not accounted for. The author proposed using ROI's that were tailored to each individual video rather than each frame as it could be reasoned that the movement of lids and eyelash shadows therefore would not impact on the final tear metric.

The image processing technique described in Figure 6.19 was created to isolate lid movement from the Placido disc images and was the first of a two stage process to eliminate non-tear interference. One of the main characteristics of Placido disc images is the repeating pattern of rings of a set frequency. These rings can be isolated from the background data by use of a band pass filter to remove image elements that do not correspond with the frequency of the Placido disk. Once a threshold was applied, smooth filtering and morphology could transform the rings into one continuous particle. See Table 6.4 for a summary of the techniques and parameters used in the analysis.

With a region of interest collected for each frame, the ROI's were summed together into one image and normalized. The intensity of each pixel would correlate with the areas of the video that had least been affected by eye lid movement. Only the areas that had remained free of eye lid movement in 95% of video frames were kept, the rest were excluded from any tear analysis. See Figure 6.20 for the ROI images after 5 and 50 frames on a single video; intensity is shown on a color map with red being the least affected by eye lids and blue being the most.

Table 6.4: Image processing values used in the eye lid removal

Parameter	Value	Parameter	Value
Frequency Filter		Morphology-Dilate	
Low pass cut-off (Hz)	0.25	Size	11x11
High pass cut-off (Hz)	0.08	Iterations	1
Threshold		Structuring Element	Square
Type	Otsu's method [Otsu, 1975]	Morphology-Fill Holes	
Smooth Filter		Hole size (px^2)	1-400
Type	Rounded local average	Morphology-Convex Hull	
Kernel size	5x5	Type	Graham scan [Graham, 1972]

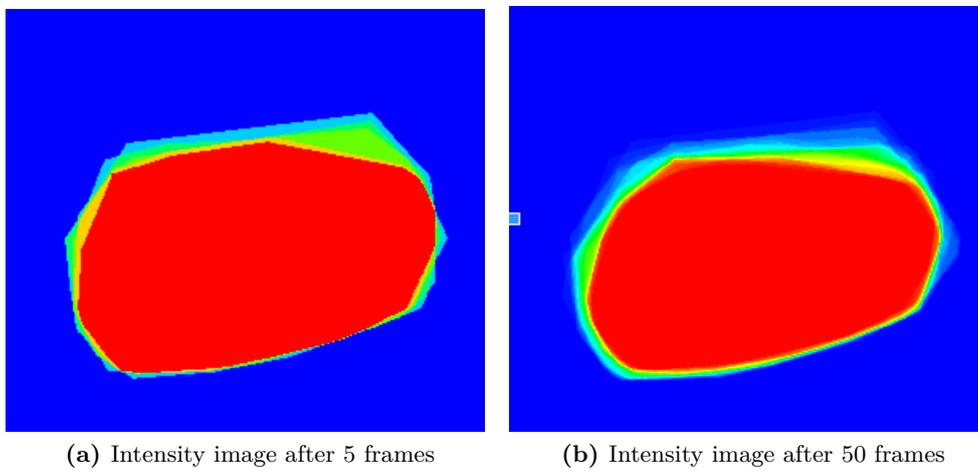


Figure 6.20: ROI intensity images

6.4.4 ROI - Stage 2, Eye lash removal

Following on from the eye lid removal, another adaptive ROI algorithm was developed to further isolate the tear film from eyelash shadows. One of the difficulties in separating these two types of interference is that they can have a similar effect on Placido rings. Even though geometric distortion of the rings does not occur with eyelash shadows, breaks within rings can mimic the effects of such distortion so as to be difficult for computers to isolate from real tear film breakdown. In addition, corneal exposure can cause the rings to drop in intensity so much that they cannot be detected, hence breaks in the rings that could appear to be from eyelash shadows. There is, however, one characteristic that is shared by most eyelash shadows, they always occur at the top and bottom of the Placido disc where conveniently they are connected to the outside of the ROI described in the previous section.

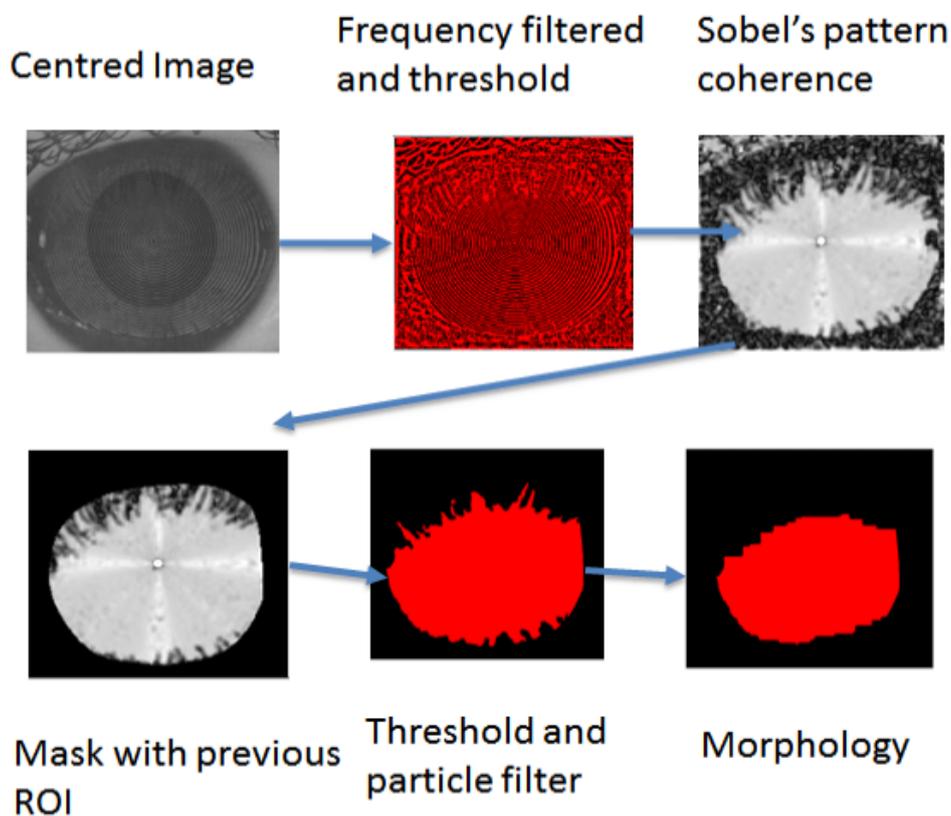


Figure 6.21: Eyelash removal image processing steps

Figure 6.21 is an illustration of the image processing algorithm. While similar in parts to the previous ROI that also used frequency filters and particle filtering, the

major difference is in the use of orientation filters for effective interference detection. The Placido disc in an unaltered state has rings that have strong local orientation (pattern coherence), for example the rings at the bottom right of the cornea will be orientated at approximately a 45° angle from the horizontal. Disturbance from both eye lashes and tear breakdown will disturb the local orientation pattern and impact on the resulting image. While it is the case that concentric rings will have different orientations at different places they should have a strong orientation in at least one direction where as interference will generally have much weaker orientation. The exact pattern coherence operation used Sobel's gradient detection operator [Sobel, 1990] and was borrowed from and fully described in [Alonso-Caneiro, 2008].

Once the image had been converted to a coherence map, a mask made from the ROI that removed eye lids was applied. The resulting image was then subject to a threshold to differentiate the signal from the noise, the eyelash interference became connected to the outside mask while tear interference did not. A particle filter could therefore be used to remove the much smaller particles leaving only the Placido disc without eyelashes. Finally a morphological "erode" function was applied to leave a single ROI. As before, all of the ROI's from each frame were summed and only the areas free of eyelashes in 80% of frames were kept. Table 6.5 is a summary of the parameters and operations used in this section.

Table 6.5: Image processing parameters used in the removal of eye lashes

Parameter	Value	Parameter	Value
Frequency Filter		Particle filter	
Low pass cut-off (Hz)	0.25	Area (px^2)	>10,000
High pass cut-off (Hz)	0.08	Morphology Erode	
Pattern coherence		Size	3x7
Edge Detector Type	Sobel [Sobel, 1990]	Iterations	8
Coherence map	Local Orientation [Alonso-Caneiro, 2008]	Structuring element	Rectangle
Threshold			
Type	Global		
Threshold Value	80		

6.4.5 Tear metric

The final stage of the image processing was to implement the tear film metric used to quantify the surface quality of the tear film. As seen in chapter 5, geometric distortion of the rings occurs when the tear film breaks down. Before breakup the rings would be equally spaced, whilst after breakup the rings would either become too close or too far spread. If the ring spacings could be automatically determined, tear breakdown could be tracked by assessing the standard deviation of all the spacings. This was the assumption on which the tear metric relied. While chapter 5 did recommend a tear metric that is sensitive to both the shape and intensity of rings, it was hard to include intensity measurements as there was significant noise produced from the iris.

Figure 6.22 gives an overview of the image processing steps used to arrive at the tear metric. Again, the centred image would be frequency filtered and the mask from the previous image processing step applied. Having already found the centre, the rings could be unwrapped to produce a polar image, whose radial pixel values lie along the Y axis and the angle at which it was sampled along the X axis. Each resulting image contained 556 radial samples, ring spacings could then be easily determined for each column of the image. The tear metric, as described in Eq. Equation 6.1, summed the standard deviation σ_i of each radial sample's ring spacings together. Table 6.6 is a summary of the parameters and operations used in this section.

$$TSFQ_A = \sum_{i=1}^{556} \sigma_i \quad (6.1)$$

Table 6.6: Image processing parameters used in determining the tear metric

Parameter	Value	Parameter	Value
Frequency Filter		Cartesian to polar	
Low pass cut-off (Hz)	0.25	Radial samples	556
High pass cut-off (Hz)	0.08	Sample length (px)	250
Threshold			
Type	Global		
Threshold Value (8-bit)	80		

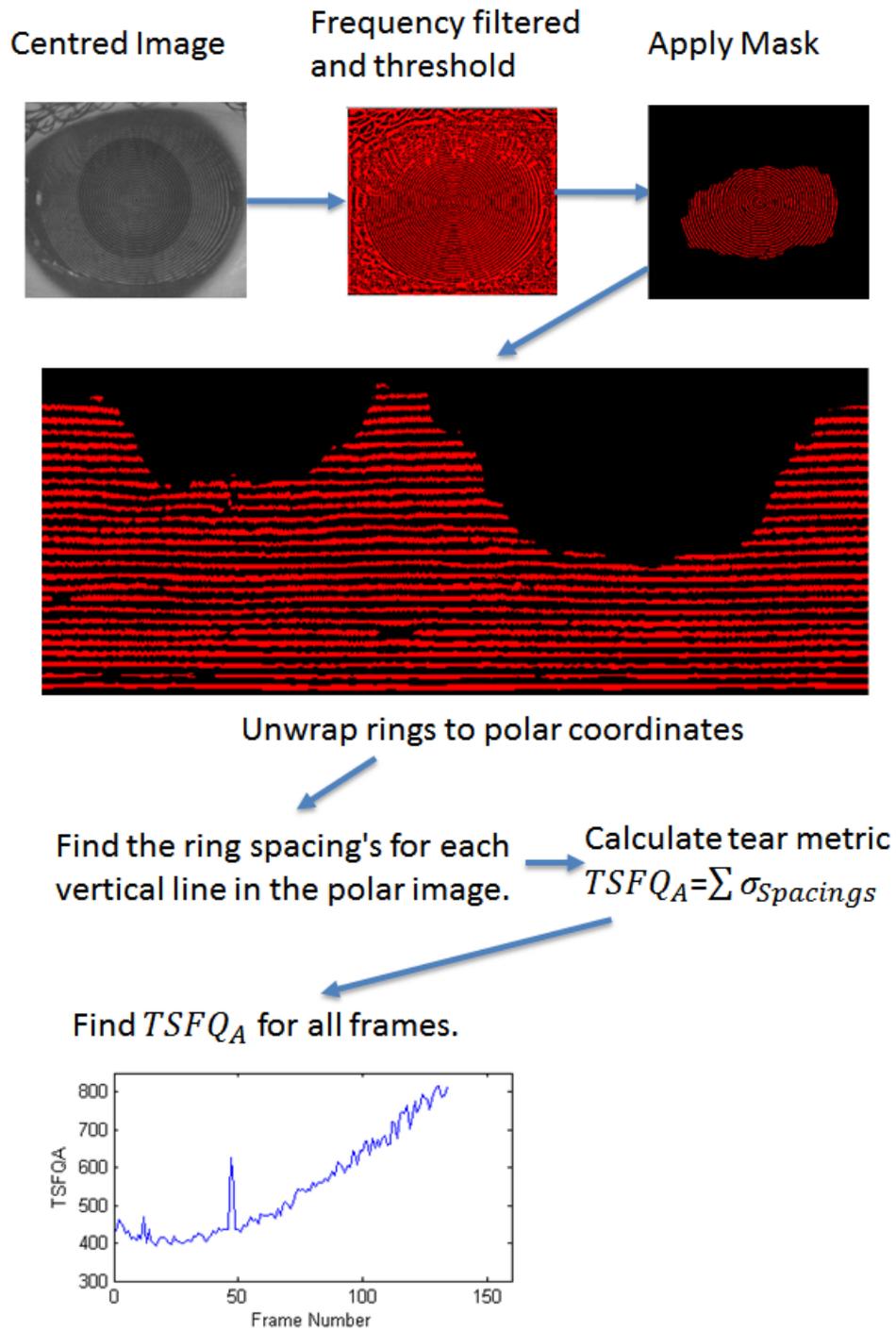


Figure 6.22: Tear metric image processing steps

6.5 Study

6.5.1 Aim

The main aim of this study was to evaluate the new image processing algorithm summarized in section 6.4 against both existing tear film algorithms and traditional tear stability measures.

6.5.2 Subjects

Twenty nine young subjects aged 18-25 years (average 21.4 ± 1.9 years) were recruited from Aston Universities student body and were asked to attend two testing sessions for the evaluation of the new image processing algorithm. Of the twenty nine subjects who attended the first session only twenty returned for the second. Informed consent was obtained from all participants and the study was conducted in accordance with the declarations of Helsinki. Ethical approval for the study was granted by the Aston University Ethics Committee.

6.5.3 Protocol

Five different tear stability measures and a symptomatic questionnaire metric were evaluated: symptom questionnaire (OSDI), tearscope break up time (TSBUT), video break up time (VBUT), inter-blink time (IBT), a tear film surface quality metric as produced by [MEDMONT, Australia] (TSFQ-M) and a tear film surface quality metric developed by the author (TSFQ-A).

After informed consent had been taken, patients filled in the OSDI questionnaire. Once complete, they were asked to sit for tear film breakdown measures on two instruments, a slit-lamp with a tearscope attachment and the MEDMONT E-300 video keratographer. All the raw video data needed for the tear film surface quality metrics, inter-blink times and video break up times could be recorded by the MEDMONT E-300 video keratographer at 15 Hz, while the tearscope break up time could be evaluated with the slit-lamp. The order in which patients sat for each instrument were randomized. At the end of the first testing session, patients were asked to return for a second, no longer than 4 days after the first and at approximately the same time of day.

On the slit-lamp, patients were asked to blink and hold their eyes open for as long as possible, meanwhile the examiner operated a stopwatch and recorded the time at which either the first perturbation in the reflected grid pattern was seen or a blink was recognized. This was repeated three times, with a time duration of larger than two minutes between each test. The mean of the three measures was taken as the TSBUT for the patient. Similarly on the video keratographer patients underwent three repeated tests under suppressed blinking conditions. As both the MEDMONT and the authors tear metric were calculated for each frame in the video, and as it is the change in tear film surface quality from the start of the video to the end that is important for tear stability, the resulting plots were fitted with a least squares first order polynomial. The mean gradients for all three fits were taken as the TSFQ-A and TSFQ-M for the Medmont and the authors tear metric respectively. Post-test the recorded videos were analyzed manually to find the interblink and video break up time. The researchers were masked as to the results of the other tear film metrics to prevent bias.

6.5.4 Statistics

Data was checked for normality (Shapiro-Wilk, $p > 0.05$) before testing could begin. All of the stability metrics returned non normal distributions so non-parametric tests were used. As most of tear film testing metrics were not directly comparable, Spearman's correlation coefficients were calculated to evaluate their agreement between each other in both the first and second testing sessions. With both the TSBUT and VBUT relying on the first perturbation of the reflected Placido rings for their measurement, a Wilcoxon signed rank tests was undertaken to establish if there was a significant difference between them. Test-retest repeatability and reliability between sessions was evaluated by both Bland-Altman analyses and Spearman's correlation coefficients. As the tear metrics repeatabilities are not directly comparable, a ratio of the 95% CI range over the range of values in session 1 for each metric was computed. In addition, Wilcoxon signed rank tests were conducted to determine if there were significant differences for the tear film metrics between sessions.

6.5.5 Results

6.5.5.1 Session 1

Of the twenty-nine subjects that completed the OSDI (Ocular surface disease index) questionnaire, eleven of the subjects scored high enough ($OSDI > 30$) to be rated as having severe symptomatic dry eye. A break down of the OSDI scores can be viewed in Table 6.7. Mean results for each metric can be seen in Table 6.8, no significant difference was found between the break up time evaluated with the tearscope to those calculated with the video keratographer ($Z = -0.4$, $p = 0.69$).

A correlation matrix consisting of the Spearman's rank correlation coefficient ρ between each of the six tear film measures can be viewed in Table 6.9. Weak correlations between all of the tear film measures and the OSDI were found in the first session of testing. Strong correlations existed between the TSBUT and VBUT ($\rho = 0.71$), and the VBUT and IBT ($\rho = 0.76$). The authors tear metric correlated better with the TSBUT ($\rho = -0.61$) than those of the Medmont ($\rho = -0.4$), but both correlated equally as well with the VBUT ($\rho = -0.67, \rho = -0.68$ respectively).

Table 6.7: Symptomatic dry eye as evaluated by the OSDI Questionnaire, $n=29$

	None OSDI < 15	Mild 15 > OSDI > 30	Severe OSDI > 30
No. of Subjects	14	4	11

Table 6.8: Means and standard deviations for the metrics across Session 1, $n=29$

Symptomatic Group	OSDI	TSBUT (s)	IBT (s)	VBUT (s)	TSFQ_M	TSFQ_A
All	22.1±19.6	14.2±10.3	20.6±11.4	15.4±12.1	3.5±5.9	1.0±1.3
None	5.8±4.3	18.6±12.2	25.2±11.6	19.8±13.3	2.2±4.3	1.0±1.3
Mild	19.8±4.8	9.9±5.0	20.0±10.9	14.4±14.0	8.9±7.8	2.1±1.4
Severe	43.5±12.0	10.2±5.1	15.1±7.6	10.1±4.7	3.2±5.7	0.6±0.9

Table 6.9: Spearman's rank correlation coefficients ρ between each metric in the first session, n=29

	OSDI	TSBUT	IBT	VBUT	TSFQ_M	TSFQ_A
OSDI	-	-0.22	-0.26	-0.17	0.03	-0.07
TSBUT	-	-	0.42	0.71	-0.41	-0.61
IBT	-	-	-	0.76	-0.24	-0.40
VBUT	-	-	-	-	-0.67	-0.68
TSFQ_M	-	-	-	-	-	0.66
TSFQ_A	-	-	-	-	-	-

Weak correlations, $\|\rho\| < 0.4$, highlighted in blue

Mild correlations, $0.7 > \|\rho\| > 0.4$, highlighted in green

Strong correlations, $\|\rho\| > 0.7$, highlighted in red

6.5.5.2 Session 2

Of the twenty subjects that returned for a second testing session, seven had severe symptomatic dry eye as evaluated by the OSDI questionnaire. A break down of the OSDI scores for the second session can be viewed in Table 6.10. Mean results for each metric are summarized in Table 6.11, there were no significant differences found between TSBUT and VBUT ($Z=-1.1$, $p=0.28$).

A matrix consisting of Spearman's rank correlation coefficients between each of the metrics collected in session 2 can be viewed in Table 6.12. Contrary to the results seen in session 1, OSDI scores correlated more highly with all tear film metrics, although only significant mild correlations were found in comparisons with the VBUT ($\rho = -0.4$) and TSBUT ($\rho = -0.65$). Correlations using the authors tear metric TSFQ_A with VBUT ($\rho = -0.70$) and IBUT ($\rho = -0.51$) increased between sessions, although there was small decrease in the correlation seen with the TSBUT ($\rho = -0.50$). Medmonts TSFQ_M metric only demonstrated mild correlations with VBUT ($\rho = -0.45$) and TSFQ_A ($\rho = -0.64$).

Table 6.10: Symptomatic dry eye as evaluated by the OSDI Questionnaire, n=20

	None OSDI<15	Mild 15>OSDI>30	Severe OSDI>30
No. of Subjects	11	2	7

Table 6.11: Means and standard deviations for the tear film metrics across session 2, n=20

Symptomatic Group	OSDI	TSBUT (s)	IBT (s)	VBUT (s)	TSFQ_M	TSFQ_A
All	21.7±19.6	15.5±10.2	20.4±11.4	14.7±10.1	2.91±5.5	1.1±1.4
None	7.3±4.2	19.4±10.3	24.6±10.8	17.2±9.7	3.5±6.2	1.1±1.4
Mild	22.5±2.5	21.6±7.7	29.8±6.1	27.1±3.4	1.9±1.6	0.1±0.1
Severe	46.4±13.5	7.6±1.2	10.4±4.6	7.9±4.1	2.2±4.4	1.4±1.3

Table 6.12: Spearman's rank correlation coefficients ρ between each tear film measure in the second session $n=20$

	OSDI	TSBUT	IBT	VBUT	TSFQ_M	TSFQ_A
OSDI	-	-0.65	-0.67	-0.48	-0.193	0.197
TSBUT	-	-	0.78	0.83	-0.12	-0.50
IBT	-	-	-	0.82	-0.09	-0.51
VBUT	-	-	-	-	-0.45	-0.70
TSFQ_M	-	-	-	-	-	0.64
TSFQ_A	-	-	-	-	-	-

Weak correlations, $\rho < 0.4$, highlighted in blue

Mild correlations, $0.7 > \rho > 0.4$, highlighted in green

Strong correlations, $0.7 > \rho > 0.4$, highlighted in red

6.5.5.3 Test-Retest Repeatability

Table 6.13 is a summary of the Spearman's correlations, Wilcoxon signed rank tests and the Bland Altman analyses. There were no significant differences between the means of any of the tear film metrics between the first and second session. TSBUT, IBT and VBUT displayed stronger correlations across sessions than those of the automated metrics TSFQ_M and TSFQ_A. Bland Altman plots for all metrics can be seen in Figure 6.23 and Figure 6.24. Across all metrics besides the OSDI, as the mean between the sessions increased so did the magnitude of the difference.

Table 6.13: Wilcoxon signed rank tests, Spearman's correlation coefficients and Bland Altman repeatability data

Statistic	OSDI	TSBUT	IBT	VBUT	TSFQ_M	TSFQ_A
Correlation						
Spearman's ρ	0.95	0.68	0.73	0.72	0.50	0.52
Wilcoxon's test						
Z-Statistic	-1.42	-0.34	-0.04	-0.05	-0.11	-0.37
p value	0.16	0.74	0.97	0.97	0.91	0.71
Bland Altman						
Mean Difference	1.90	0.67s	0.39s	1.14s	0.26	-0.18
COR	11.16	15.56s	17.44s	18.50s	12.22	2.53
(95% CI range)/(session 1 range)	0.36	0.84	1.07	1.02	1.08	1.09

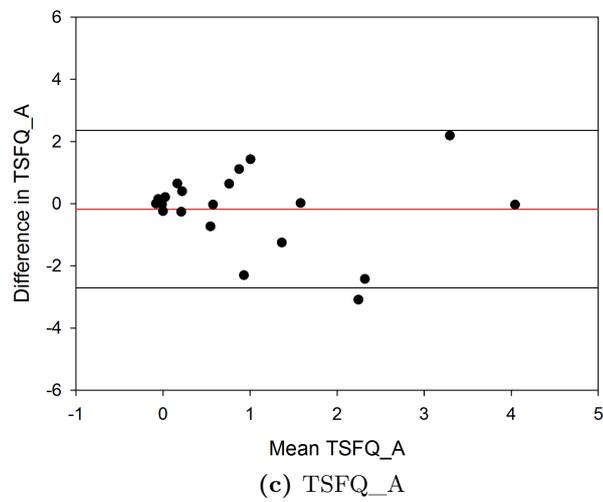
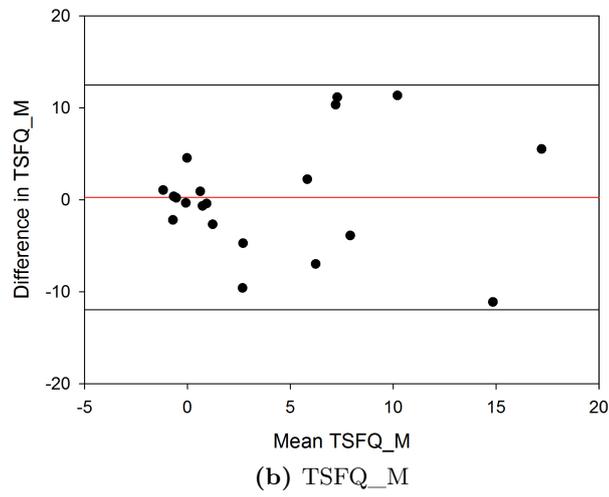
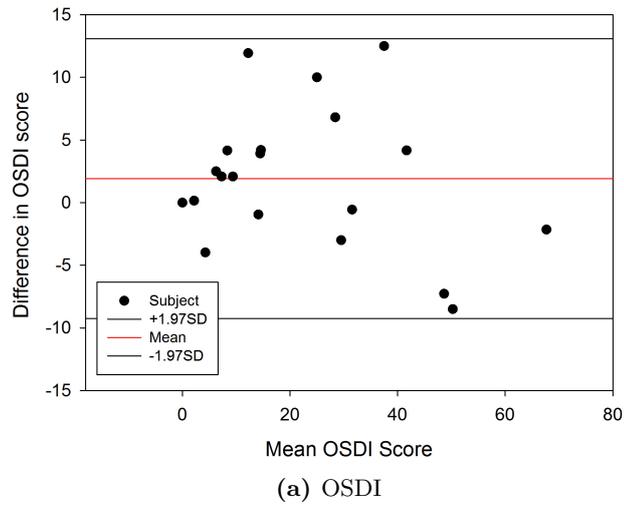


Figure 6.23: OSDI, TSFQ_M and TSFQ_A Bland-Altman plots for repeatability

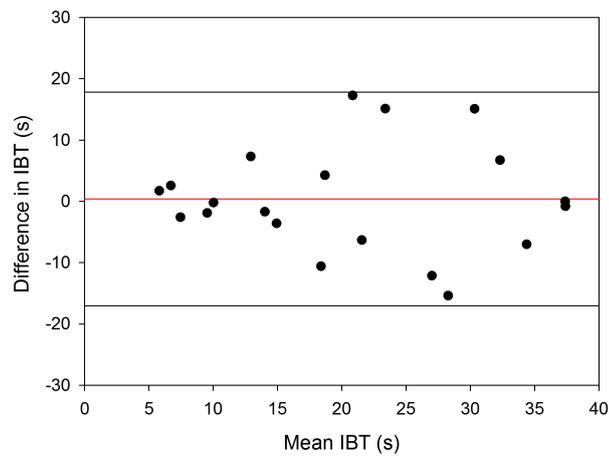
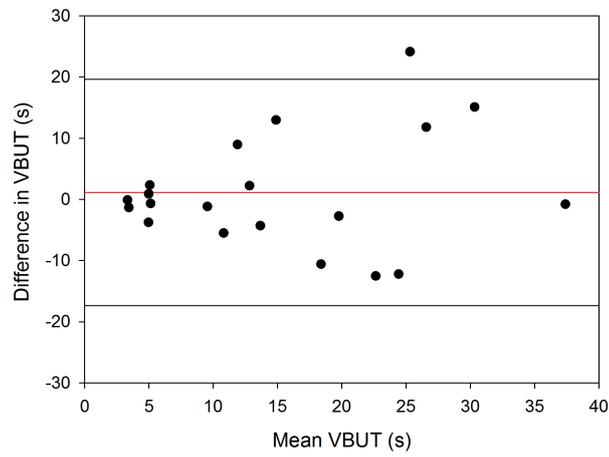
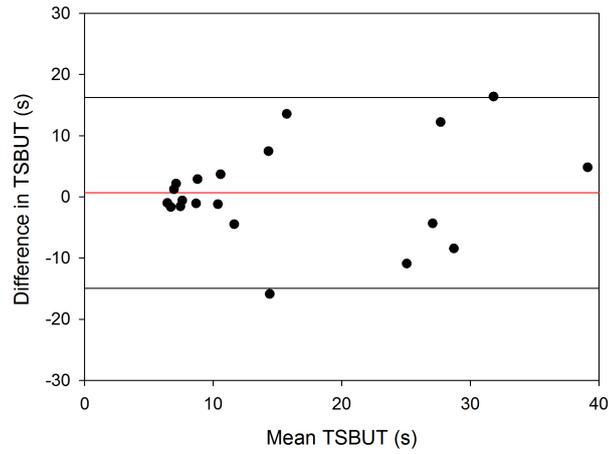


Figure 6.24: TSBUT, VBUT and IBT Bland-Altman plots for repeatability 242

6.5.6 Discussion and conclusions

The automated tear metrics had greater correlation with the break up time evaluated subjectively with the topographer videos than those measured using the Tearscope in both sessions. As both the automated metrics and the VBUT derived their values from the same underlying video data this is unsurprising. With the VBUT being the gold standard, the higher correlation values seen with the authors tear metric, TSFQ_A, (Table 6.9, Table 6.12) than those of Medmont's tear metric, TSFQ_M, in both sessions suggests that TSFQ_A is a better indicator of tear stability. However, with correlation values of ($\rho = -0.68$, $\rho = -0.70$), lower than between the TSBUT and VBUT ($\rho = 0.71$, $\rho = 0.83$) it still compares unfavourably to subjective methods. The results for this study are in agreement with the results of Goto et al. [2003] who found correlation values of ($r = 0.72$) between their automated break up time and those subjectively assessed on raw Placido images.

Correlations between the TSBUT and the automated metrics were either weak or mild in both testing sessions. Best et al. [2012] found little agreement between their automated break up time and the TSBUT, although no correlation values were given. Kopf et al. [2008] found equally poor correlations between their automated metric and fluorescein break up time ($r = 0.22$), as did Hong et al. [2013] ($\rho = 0.55$) and in their subsequent paper [Hong et al., 2014] ($\rho = 0.34$).

It should be noted that there are significant differences between the assessment method of break up time and the automated tear film metrics used in this study; break up time is the time at which either the first perturbation or blink is observed while the TSFQ_A and TSFQ_M measure the rise in tear instability over time. Therefore a low value of VBUT or TSBUT could arise through a blink without any changes in tear stability, which may be one reason why both the automated metrics do not correlate more heavily with the VBUT or TSBUT. In the first session 52% of the VBUT values were equal to the IBT, while in the second session 45% were. This explains the high correlations between IBT and VBUT.

In session 1 there was little correlation between any of the tear measures and the OSDI score, however, in session 2 stronger correlations were found particularly with the TSBUT ($\rho = 0.65$) and IBT ($\rho = 0.67$). The results from session 2 are in agreement with Pult et al. [2011], who found a similar correlation ($r = 0.59$) between TSBUT and OSDI. The reason for this disparity in results between testing sessions is unknown, especially as no significant difference was found between any of the tear metrics across the two testing sessions.

Both the TSFQ_A and the TSFQ_M suffered from poor reliability and repeatability in comparison to the TSBUT and VBUT, with low correlation values between the two testing sessions and large 95% confidence interval ratios (Table 6.13). Hence using these automated measures to differentiate between patient conditions could be problematic. Meanwhile the OSDI symptom questionnaire showed the highest repeatability and reliability.

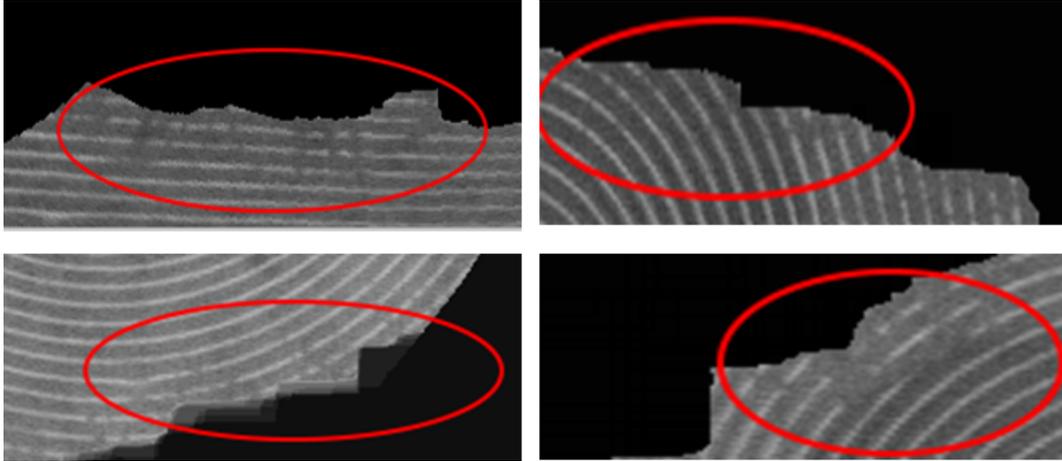


Figure 6.25: Segmentation faults in the TSFQ_A metric

The low correlation values between the automated and existing tear film measures and questionnaire metric, and the lack of repeatability brings the validity of the current image processing techniques into question. While it can be argued that objective measures may eliminate observer biases, they can be strongly affected by interference from lid, eyelash and eye movement that many of the image processing methods find difficult to separate. In the case of the TSFQ_A metric for which it is possible to see the segmentation results, Figure 6.25, the image processing technique outlined in this chapter could not fully mask tear shadow interference for many of the videos. As a result eyelash movement could dramatically effect the tear film metric, leading to the poor repeatability observed. It can be assumed that this problem is common to many automated tear stability techniques; Best et al. [2012] argued that the smaller break up time results found with his automated instrument was a result of interference in the image capture process being interpreted as a tear film break up.

For the time being subjective methods of tear stability assessment offer greater clinical repeatability, reliability and utility, however objective automated analysis of tear film stability is still in its infancy. Further improvements in segmenting eyelash

shadows from tear analysis videos should be researched.

7 Conclusion

7.1 Introduction

This thesis set out to demonstrate the potential of both dynamic image processing and mobile computing technology to enhance clinical measurement in optometric practice by investigating these two questions:

1. What are the advantages of clinical testing on mobile devices in comparison to existing measures?
2. How can a new computer vision system improve the objective measurement of tear stability?

The first question was explored by the creation and validation of three novel mobile apps for reading performance (Chapter 2), contrast sensitivity (Chapter 3), and subjective amplitude of accommodation (Chapter 4) assessment with several technological enhancements. These included, touch screen interaction with patients, colour bit stealing, automated distance tracking, voice detection, computerized scoring and hardware interfacing of sensors.

The second question was explored by modelling the affect of tear film instability on video ketratoscopy images using novel geometrical models of tear film breakdown combined with an advanced ray tracing technique (Chapter 5). In addition, a new automated computer vision technique was created and validated against existing subjective and objective tear stability measures (Chapter 6).

7.2 Discussions and conclusions

Chapter 2 outlined the development of a novel reading performance test on a tablet computer, the first of its kind. The “Aston Read” chart had the ability to

control illumination, present sentences effectively and reduce testing time. Furthermore automatic scoring and displaying of patients data lessened the burden on the examiner and could reduce subjective bias.

The results from the study demonstrated that while reading metrics between the app and the paper based charts are not interchangeable, they correlate highly and thus the app based charts have the potential to capture functional visual ability in clinical practice. With repeatability that was equal to those of traditional paper based charts for optimal reading speed and reading acuity, and greater for critical print size, clinical testing of reading performance can be improved by the use of mobile devices. Furthermore, the Aston Read chart could reduce examination time. Following on from the study, development of the Aston Read app to reduce patient and examiner biases further, led to the creation of objective measurements of reading speed using voice recording and an image processing method for determining patient distance. While currently untested these two innovations have the potential of improving reading performance assessment further.

Chapter 3 presented a new method of testing contrast sensitivity at both near and far distance using a tablet computer and an external monitor. Previous PC and tablet based contrast sensitivity testing relied on using the 2-AFC testing paradigm, which while accurate, required hundreds of contrast targets to be presented to subjects. Presenting the targets and collecting responses takes too long to be considered practical for clinical testing. Paper based grating charts meanwhile suffer from poor repeatability and lack precision, with only a handful of spatial frequencies and contrasts to be tested. The new near and far chart used a swept frequency curve with a colour bit stealing routine to present nearly all target frequencies and contrasts to the patient in one presentation. Using the touch-screen technology embedded into mobile devices, patients could draw out their contrast sensitivity curve (using the method of adjustment) in one go, reducing testing time.

The study conducted found that the repeatability was better using the new near and far distance apps than a paper based gratings chart (CSV-1000) but lower than a letter based contrast chart (Pelli-Robson). As contrast letter charts are restricted to testing only one spatial frequency they cannot be used to plot a full contrast sensitivity function. Thus the new apps provide a more repeatable method of plotting a patients contrast sensitivity function than clinical alternatives such as the paper based gratings chart. However, repeatability was performed with Bangertter foils

which may have affected results (artificially increased repeatability).

Poor correlations between all of the test charts, even between the near and far contrast sensitivity apps, leaves unanswered questions about the new apps validity in differentiating between patients, although this may be a result of the limited age range of the patients and hence the limited range of true contrast sensitivities studied. No gold standard from which to compare all of the contrast tests were included in the study. Further studies should compare the contrast sensitivity measures of the new apps to those of a laboratory CRT set up to test their validity in a wider range of patients.

Chapter 4 introduced a new smart-phone attachment to measure subjective amplitude of accommodation. The attachment in the form of a smart-phone case, interfaced an ultrasonic distance sensor to that of a mobile device; using the feedback from the distance sensor to resize a target to maintain a constant angular sub-tense. It was thought that maintaining angular sub-tense of targets could reduce pseudo accommodation and over estimation of AOA measures. It was theorised that patients can more easily spot blur in small targets rather than large targets, so when a fixed target increases in size as it gets close to the patient, the patients tolerance to blur decreases skewing AOA readings. Additionally, the new accommodometer offered a free space method of measuring AOA without having to resort to resting a near point rule invasively on patients cheeks.

The study that was ran found no difference between AOA measures obtained with the new Aston Accommodometer and the traditional near point rule when a fixed target was used. This validated the smart-phone attachment and free space method of assessing AOA. While there was increased intra-session variability with the Aston Accommodometer, once three measures had been taken, as is normal practice in AOA measurement, AOA measures were indistinguishable from those of the near rule. All of the studies in this thesis were undertaken without reference to a priori power calculations, so some elements of the study may not have had sufficient power.

Conversely, the use of a resizing target did affect AOA measures taken using the pull down but not the push up technique. The pull down method had significantly lower AOA values with a resizing target than the corresponding fixed target on the Aston Accommodometer. While this would seem to suggest that a resizing target reduces pseudo accommodation, it is strange that only the pull down method was affected. Furthermore, correlations between push up and pull down measures dropped using

the resizing target. It is unknown why this occurred, but it may be as result of the resizing target introducing errors unrelated to actual accommodation. Further studies should be undertaken to investigate this anomaly.

In regards to the first question, the advantages of using mobile devices for clinical testing are numerous. Firstly, this thesis showed that repeatability and reliability can be increased with the novel application of mobile devices; both the new reading performance test and contrast sensitivity app were more reliable and repeatable than their existing paper counterparts. This is a result of reducing examiner biases in testing methods, automating test procedures and controlling factors such as illumination efficiently. They can also reduce testing time, the evaluation of reading speed for example does not need the examiner to manually unveil sentences, record patient times and plot the results to be able to find reading performance metrics. Equally for contrast sensitivity testing, time could be saved by allowing the patients to trace out their whole contrast sensitivity function using the touch-screen in one go. Contrast sensitivity tests are no longer constrained to paper based gratings charts whose spatial frequencies and contrasts are limited, or to traditional computer tests that present hundreds of targets. Furthermore, this thesis showed that mobile devices can be used to digitize near-vision testing effectively, both the new accommodometer and the reading performance app proved to be valid methods of assessment. Previous PC based tests had neither the resolution nor the portability of mobile devices to test at near distances. New approaches to near testing within the clinic can be implemented using mobile devices; the resizing target on the Aston Accommodometer and the patient interaction with the contrast sensitivity app could not of been achieved using paper based charts which are inflexible or PC based tests that are too large for clinical use.

Chapter 5 documented the creation of a new geometrical model and an advanced ray tracer for use in simulating the effects of tear film breakdown on video-keratotomy images. The geometrical model was novel as it could orientate tear film break ups with respect to the surface normal of the corneal, a more natural portrayal of tear film geometry than past research that modelled all tear film break up geometry in the same direction. The approach used, termed the “numerical approximation” technique, was uniquely flexible as it does not rely on the underlying shape of the cornea; any corneal shape or tear film break up profiles can easily be applied using the same technique. Additionally, the custom ray tracer took account of both specular and diffuse reflection (albeit an approximation) hence any simulations that were

produced gave a more complete picture of tear film breakdown. The simulations that were ran found that initially spatial distortions are more noticeable, as the break-up widens with time, however, intensity distortions become ever more prominent. Hence if automated tear metrics were to be created they should be sensitive to both spatial and intensity based ring distortions.

Chapter 6 documented the practical challenges of creating a custom Placido disc based tear device and a new computer vision system of assessing tear film stability. Ultimately the creation of the Placido disc based tear device was unsuccessful as manufacturing difficulties could not be overcome. On the new tear analysis system, the author proposed using regions of interests that separated eyelid and eyelash data from Placido images, that were tailored to each individual video rather than to each separate frame as it could be reasoned that the movement of lids and eyelash shadows therefore would not impact on the final tear metric.

A study was conducted to test the new objective tear analysis system against existing subjective and objective measures of tear stability. While it was found that the authors computer vision system correlated better with NIBUT and VBUT than a rival objective tear stability system, it still suffered from poor repeatability in comparison to subjective measures. The custom tear analysis system outlined in the chapter struggled to mitigate interference from eyelash movements despite efforts to segment eyelashes. Often the segmentation would remove a large part of the inferior cornea, the area where break-ups happens first, from the tear analysis reducing the effectiveness of the image processing system.

In regards to second question, “how can a new computer vision system improve the objective measurement of tear stability?”, Chapter 5 showed that computer vision systems can be improved by using a metric that can measure both intensity and spatial distortions. However, Chapter 6, which used a more practical rather than theoretical approach, found that objective measurements can be improved by better segmentation of interference in tear images. While a new segmentation technique was used in a custom automated computer vision system, and proved to be in better agreement with existing subjective tear measures than a competing computer vision system, it still could not match the repeatability of subjective tests. It appears that subjective tear stability measures offer greater validity and repeatability than current objective systems, hence the utility of computer vision systems to measure tear stability is still under question. Pattern detection and segmentation is a simple

task for human observers but is challenging for computer vision systems.

Bibliography

- M. Abelson, G. Ousler, and L. Nally. *Lacrimal gland, tear film, and dry eye syndromes 3: basic science and clinical relevance*. Springer, New York, 2002.
- P. M. Adler, M. Cregg, A. Viollier, and J. Margaret Woodhouse. Influence of target type and RAF rule on the measurement of near point of convergence. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 27(1):22–30, 2007.
- A. Agarwal. *Dry Eye: A Practical Guide to Ocular Surface Disorders*. Slack, Thorofare, 2006.
- S. Ahn and G. Legge. Psychophysics of reading XIII. Predictors of magnifier-aided reading speed in low vision. *Vision Research*, 35(13):1931–1938, 1995.
- S. Ahn, G. Legge, and A. Luebker. Printed cards for measuring low-vision reading speed. *Vision Research*, 35(13):1939–1944, 1995.
- D. Albrecht, R. D. Valois, and L. Thorell. Visual cortical neurons: are bars or gratings the optimal stimuli? *Science*, 207(1):88–90, 1980.
- J. L. Alió, W. Radner, A. B. Plaza-Puche, D. Ortiz, M. C. Neipp, M. J. Quiles, and J. Rodríguez-Marín. Design of short Spanish sentences for measuring reading performance: Radner-Vissum test. *Journal of cataract and refractive surgery*, 34(4):638–42, 2008.
- D. Alonso-Caneiro. Computationally efficient interference detection in videokeratometry images. In *Multimedia Signal Processing, 2008 IEEE 10th Workshop*, volume 1, pages 486–490, 2008.
- D. Alonso-Caneiro. Assessment of tear film surface quality using dynamic-area high-speed videokeratometry. *IEEE Transactions On Biomedical Engineering*, 56(5):1473–1481, 2009.
- D. Alonso-Caneiro, J. Turuwhenua, D. R. Iskander, and M. J. Collins. Diagnosing dry eye with dynamic-area high-speed videokeratometry. *Journal of biomedical optics*, 16(7):076012, 2011.

- H. Anderson, G. Hentz, and A. Glasser. Minus Lens Stimulated Accommodative Amplitude Decreases Sigmoidally with Age: A Study of Objectively Measured Accommodative Amplitudes from Age 3. *Investigative ophthalmology & visual science*, 49(7):2919–2926, 2008.
- B. Antona, F. Barra, A. Barrio, E. Gonzalez, and I. Sanchez. Repeatability intraexaminer and agreement in amplitude of accommodation measurements. *Graefe's archive for clinical and experimental ophthalmology*, 247(1):121–7, 2009.
- S. Arnold, A. Walter, T. Eppig, H. Bruenner, and A. Langenbacher. Simultaneous examination of tear film break-up and the lipid layer of the human eye: a novel sensor design (Part 1). *Zeitschrift für medizinische Physik*, 20(4):309–15, 2010.
- K. Arundale. An investigation into the variation of human contrast sensitivity with age and ocular pathology. *The British journal of ophthalmology*, 62(4):213–5, 1978.
- D. Atchison, E. Capper, and K. McCabe. Critical subjective measurement of amplitude of accommodation. *Optometry & Vision Science*, 71(11):699–706, 1994.
- D. Atchison, W. Charman, and R. Woods. Subjective depth of focus of the eye. *Optometry & Vision Science*, 74(7):511–520, 1997.
- G. Avidan and M. Harel. Contrast sensitivity in human visual areas and its relationship to object recognition. *Journal of Neurophysiology*, 87(6):3102–3116, 2002.
- I. Bailey and J. Lovie. The design and use of a new near-vision chart. *American journal of optometry and physiological optics*, 57(6):378–387, 1980.
- L. J. Balcer, M. L. Baier, V. S. Pelak, R. J. Fox, S. Shuwairi, S. L. Galetta, G. R. Cutter, and M. G. Maguire. New low-contrast vision charts: reliability and test characteristics in patients with multiple sclerosis. *Multiple Sclerosis*, 6(3):163–171, 2000.
- O. M. Bartuccio and M. Taub. Accommodative Insufficiency: A Literature and Record Review. *Optometry & Vision Development*, 13(2):35–40, 2008.
- A. Bastawrous, R. C. Cheeseman, and A. Kumar. iPhones for eye surgeons. *Eye*, 26(3):343–54, 2012a.
- A. Bastawrous, C. Leak, F. Howard, and V. Kumar. Validation of Near Eye Tool for Refractive Assessment (NETRA) Pilot Study. *Journal of Mobile Technology in Medicine*, 1(3):6–16, 2012b.

- M. Ben Abdallah, J. Malek, R. Tourki, J. E. Monreal, and K. Krissian. Automatic estimation of the noise model in fundus images. In *10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13)*, pages 1–5. IEEE, 2013.
- D. A. Benedetto, D. O. Shah, and H. E. Kaufman. The instilled fluid dynamics and surface chemistry of polymers in the precocular tear film. *Investigative ophthalmology*, 14(12):887–902, 1975.
- D. A. Benedetto, T. E. Clinch, and P. R. Laibson. In Vivo Observation of Tear Dynamics Using Fluorophotometry. *Archives of Ophthalmology*, 102(3):410–412, 1984.
- N. Best, L. Drury, and J. S. Wolffsohn. Clinical evaluation of the Oculus Keratograph. *Contact lens & anterior eye : the journal of the British Contact Lens Association*, 35(4):171–4, 2012.
- J. M. Black, R. J. Jacobs, G. Phillips, L. Chen, E. Tan, a. Tran, and B. Thompson. An assessment of the iPad as a testing platform for distance visual acuity in adults. *BMJ open*, 3(6):5–7, 2013.
- H. R. Blackwell. Studies of psychophysical methods for measuring visual thresholds. *Journal of the Optical Society of America*, 42(9):606–16, 1952.
- C. Blakemore and F. Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of physiology*, 203(1):237–260, 1969.
- C. Blakemore, J. Nachmias, and P. Sutton. The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain. *The Journal of Physiology*, 210(3):727–750, 1970.
- R. Bonnet, Y. L. Grand, and C. Rاپilly. *La topographie cornéenne et l'adaptation des lentilles de cornée*. N. Desroches, Paris, 1964.
- Y. Botelho and V. Martinez. Electrolytes at various in lacrimal gland fluid and in tears flow rates in the rabbit. *American Journal Of Physiology*, 225(3):2–5, 1973.
- A. Bradley, E. Switkes, and K. De Valois. Orientation and spatial frequency selectivity of adaptation to color and luminance gratings. *Vision research*, 28(7):841–56, 1988.
- C. I. Braun, W. E. Benson, N. A. Remaley, E. Y. Chew, and F. L. Ferris. Accommodative amplitudes in the Early Treatment Diabetic Retinopathy Study. *Retina*, 15(4):275–81, 1995.

- a. Bron. Functional aspects of the tear film lipid layer. *Experimental Eye Research*, 78(3):347–360, 2004a.
- A. Bron. Functional aspects of the tear film lipid layer. *Experimental Eye Research*, 78(3):347–360, 2004b.
- A. J. Bron, J. A. Smith, and M. Calonge. Methodologies to diagnose and monitor dry eye disease: report of the Diagnostic Methodology Subcommittee of the International Dry Eye WorkShop (2007). *The ocular surface*, 5(2):108–52, 2007.
- S. I. Brown. Hydrodynamics of Blinking. *Archives of Ophthalmology*, 82(4):541, 1969.
- T. Buehren, M. Collins, and D. Iskander. The stability of corneal topography in the post-blink interval. *Cornea*, 20(8):826–833, 2001.
- T. Buehren, B. Lee, M. Collins, and D. Iskander. Ocular microfluctuations and videokeratometry. *Cornea*, 21(4):346, 2002.
- F. Campbell and D. Green. Optical and retinal factors affecting visual resolution. *The Journal of Physiology*, 181(3):576–593, 1965.
- F. Campbell and J. Robson. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3):551–566, 1968.
- D. A. Caneiro. *Non-invasive Assessment of Tear Film Surface Quality*. PhD thesis, Queensland University of Technology, 2010.
- T. Carter, M. Rodrigues, A. Robertson, and R. Brady. Smartphone and medical applications use by contemporary surgical trainees: A national questionnaire study. *Journal of Mobile Technology in Medicine*, 3(2):2–10, 2014.
- E. Cebreiro, L. Ramos, and A. Mosquera. Automation of the tear film break-up time test. In *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, number 1, page 123, 2011.
- T. H. Cedarstaff and A. Tomlinson. Human Tear Volume, Quality And Evaporation: A Comparison Of Schirmer, Tear Break-Up Time And Resistance Hygrometry Techniques. *Ophthalmic and Physiological Optics*, 3(3):239–245, 1983.
- R. Chakrabarti. Application of mobile technology in ophthalmology to meet the demands of low-resource settings. *Journal of Mobile Technology in Medicine*, 1(4):1–4, 2012.

- R. Chakrabarti and C. Perera. Smartphone medical imaging: applications and future considerations. *Journal of Mobile Technology in Medicine*, 3(1):1–1, 2014.
- J. B. Chan, H. C. Ho, N. F. Ngah, and E. Hussein. DIY - Smartphone Slit-Lamp adaptor. *Journal of Mobile Technology in Medicine*, 3(1):16–22, 2014.
- W. N. Charman. The eye in focus: accommodation and presbyopia. *Clinical & experimental optometry : journal of the Australian Optometrical Association*, 91(3):207–25, 2008.
- A. Chen and D. J. O. Leary. Validity and repeatability of the modified push-up method for measuring the amplitude of accommodation. *Clinical & experimental optometry : journal of the Australian Optometrical Association*, 81(2):63–71, 1998.
- M. Chen. Accommodation in pseudophakic eyes. *Taiwan Journal of Ophthalmology*, 2(4):117–121, 2012.
- R. Chen, Z. Jin, and L. A. Colón. Analysis of tear fluid by CE/LIF: a noninvasive approach for glucose monitoring. *Journal of capillary electrophoresis*, 3(5):243–8, 1996.
- I. Cher. Another Way to Think of Tears: Blood, Sweat, and Dacruon. *The ocular surface*, 5(3):4, 2007.
- S.-H. Cheung, C. S. Kallie, G. E. Legge, and A. M. Y. Cheong. Nonlinear mixed-effects modeling of MNREAD data. *Investigative ophthalmology & visual science*, 49(2):828–35, 2008.
- S. T. Chung, J. S. Mansfield, and G. E. Legge. Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision research*, 38(19):2949–62, 1998.
- E. Colombo, L. Issolio, J. Santillan, and R. Aguirre. What characteristics a clinical CSF system has to have. *Optica Applicata*, 39(2):415–428, 2009.
- T. Costa and R. Nogueira. Differential effects of aging on spatial contrast sensitivity to linear and polar sine-wave gratings. *Brazilian Journal of Medical and Biological Research*, 46(10):855–860, 2013.
- J. Creech, L. T. Do, I. Fatt, and C. Radke. In vivo tear-film thickness determination and implications for tear-film stability. *Current eye research*, 17(11):1058–1066, 2009.

- S. Cronje-Dunn and W. F. Harris. Keratometric variation: the influence of a fluid layer. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 16(3):234–6, 1996.
- E. Davis, J. Hovanesian, and J. Katz. Professional life and the smartphone. *Cataract & Refractive Surgery Today*, 20(7), 2010.
- R. L. De Valois, H. Morgan, and D. M. Snodderly. Psychophysical studies of monkey Vision-III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research*, 14(1):75–81, 1974.
- L. Dell’Osso and R. Daroff. Nystagmus and saccadic intrusions and oscillations. In *Neuro-ophthalmology*, pages 369–401. Lippincott Williams & Wilkins, Philadelphia, 3rd edition, 1999.
- G. Derefeldt, G. Lennerstrand, and B. Lundh. Age Variations In Normal Human Contrast Sensitivity. *Acta Ophthalmologica*, 57(4):679–690, 2009.
- R. DeValois and K. DeValois. *Spatial vision*. Oxford University Press, New York, 1988a.
- R. DeValois and K. DeValois. *Spatial vision*. Oxford University Press, New York, 1988b.
- A. K. Dextl, H. Schlögel, M. Wolfbauer, and G. Grabner. Device for improving quantification of reading acuity and reading speed. *Journal of refractive surgery*, 26(9):682–8, 2010.
- M. A. Di Leo, S. Caputo, B. Falsini, V. Porciatti, A. Minnella, a. V. Greco, and G. Ghirlanda. Nonselective loss of contrast sensitivity in visual system testing in early type I diabetes. *Diabetes Care*, 15(5):620–625, 1992.
- N. Dodgson. Variation and extrema of human interpupillary distance. In *Proceedings Of The Society Of Photo-Optical Instrumentation Engineers (Spie)*, volume 12, pages 36–46, 2004.
- M. Dogru, R. Honda, and M. Omoto. Early visual results with the 1CU accommodating intraocular lens. *Journal of Cataract & Refractive Surgery*, 31(5):895–902, 2005.
- F. Donders and W. Moore. *On the anomalies of accommodation and refraction of the eye with a preliminary essay on physiological dioptrics*. The New Sydenham Society, London, 1864.

- M. Dorr, L. A. Lesmes, Z.-L. Lu, and P. J. Bex. Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Investigative ophthalmology & visual science*, 54(12):7266–73, 2013.
- J. D. Doss, R. L. Hutson, J. J. Rowsey, and D. R. Brown. Method for calculation of corneal profile and power distribution. *Archives of ophthalmology*, 99(7):1261–5, 1981.
- A. Duane. Studies in monocular and binocular accommodation, with their clinical application. *Transactions of the American Ophthalmological Society*, 20:132–157, 1922.
- A. Dubra, C. Paterson, and C. Dainty. Study of the tear topography dynamics using a lateral shearing interferometer. *Optics express*, 12(25):6278–6288, 2004.
- M. Elder, C. Murphy, and G. Sanderson. Apparent accommodation and depth of field in pseudophakia. *Journal of Cataract & Refractive Surgery*, 22(5):615–619, 1996.
- D. Elliott. *Clinical procedures in primary eye care, 4th Edition*. Elsevier Health Sciences, New York, 4th edition, 2013a.
- D. Elliott, M. Trukolo-Ilic, and J. Strong. Demographic characteristics of the vision-disabled elderly. *Investigative ophthalmology & visual science*, 38(12):2566–2575, 1997.
- D. B. Elliott. *Clinical Procedures in Primary Eye Care, 4th Edition*. Elsevier Health Sciences, New York, 4th edition, 2013b.
- D. B. Elliott, B. Patel, and D. Whitaker. Development of a reading speed test for potential-vision measurements. *Investigative ophthalmology & visual science*, 42(8):1945–9, 2001.
- M. Elliott, H. Fandrich, T. Simpson, and D. Fonn. Analysis of the repeatability of tear break-up time measurement techniques on asymptomatic subjects before, during and after contact lens wear. *Contact Lens and Anterior Eye*, 21(4):98–103, 1998.
- I. Fatt. Observations of tear film break up on model eyes. *The CLAO journal : official publication of the Contact Lens Association of Ophthalmologists*, 17(4): 267–81, 1991.
- G. Fechner. *Elemente der psychophysik*. Breitkopf und Hartel, Leipzig, 1860.

- B. Fenner and L. Tong. Corneal staining characteristics in limited zones compared to whole cornea documentation for the detection of dry eye subtypes. *Investigative ophthalmology & visual science*, 13(1), 2013.
- B. S. Fine and M. Yanoff. *Ocular histology : a text and atlas*. Harper & Row, Hagerstown, 1st edition, 1979.
- E. M. Fine and E. Peli. Scrolled and rapid serial visual presentation texts are read at similar rates by the visually impaired. *Journal of the Optical Society of America*, 12(10):2286–92, 1995.
- R. C. Fitch. Procedural effects on the manifest human amplitude of accommodation. *American journal of optometry and archives of American Academy of Optometry*, 48(11):918–26, 1971.
- M. Flom, G. Heath, and E. Takahashi. Contour interaction and visual resolution: Contralateral effects. *Science*, 142(3594):979–980, 1963.
- G. N. Foulks. The correlation between the tear film lipid layer and dry eye disease. *Survey of ophthalmology*, 52(4):369–74, 2007.
- G. N. Foulks and J. V. Jester. 2007 Report of the International Dry Eye WorkShop (DEWS). *Ocular Surface*, 5(2):65–206, 2007.
- S. Franco, A. C. Silva, A. S. Carvalho, A. S. Macedo, and M. Lira. Comparison of the VCTS-6500 and the CSV-1000 tests for visual contrast sensitivity testing. *Neurotoxicology*, 31(6):758–61, 2010.
- M. Fukuyama, T. Oshika, S. Amano, and F. Yoshitomi. Relationship between apparent accommodation and corneal multifocality in pseudophakic eyes. *Ophthalmology*, 106(6):1178–1181, 1999.
- M. A. García-Pérez and R. Alcalá-Quintana. Interval bias in 2AFC detection tasks: sorting out the artifacts. *Attention, perception & psychophysics*, 73(7):2332–52, 2011.
- M. A. García-Pérez, R. Alcalá-Quintana, R. L. Woods, and E. Peli. Psychometric functions for detection and discrimination with and without flankers. *Attention, perception & psychophysics*, 73(3):829–53, 2011.
- E. Gibson. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1):1–76, 1998.

- E. Gibson. The dependency locality theory: A distance-based theory of linguistic complexity. In *Image, language, brain*, pages 95–126. MIT press, Cambridge, 1st edition, 2000.
- B. Gilmartin. The Marton Lecture: Ocular Manifestations Of Systemic Medication. *Ophthalmic and Physiological Optics*, 7(4):449–459, 1987.
- B. Gilmartin. The aetiology of presbyopia: a summary of the role of lenticular and extralenticular structures. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 15(5):431–7, 1995.
- A. P. Ginsburg and M. W. Cannon. Comparison of three methods for rapid determination of threshold contrast sensitivity. *Investigative ophthalmology & visual science*, 24(6):798–802, 1983.
- I. Gipson, M. Yankauckas, S. Spurr-Michaud, A. Tisdale, and W. Rinehart. Characteristics of a glycoprotein in the ocular surface glycocalyx. *Investigative ophthalmology & visual science*, 33(1):218–227, 1992.
- I. K. Gipson. Distribution of mucins at the ocular surface. *Experimental Eye Research*, 78(3):379–388, 2004.
- E. Goto. Computer-Synthesis of an Interference Color Chart of Human Tear Lipid Layer, by a Colorimetric Approach. *Investigative Ophthalmology & Visual Science*, 44(11):4693–4697, 2003.
- E. Goto. Quantification of tear interference image: tear fluid surface nanotechnology. *Cornea*, 23(8):520–524, 2004.
- E. Goto and S. Tseng. Differentiation of lipid tear deficiency dry eye by kinetic analysis of tear interference images. *Archives of ophthalmology*, 121(2):173, 2003a.
- E. Goto and S. Tseng. Kinetic analysis of tear interference images in aqueous tear deficiency dry eye before and after punctal occlusion. *Investigative ophthalmology & visual science*, 44(5):1897, 2003b.
- T. Goto, X. Zheng, S. D. Klyce, H. Kataoka, T. Uno, M. Karon, Y. Tatematsu, T. Bessyo, K. Tsubota, and Y. Ohashi. A new method for tear film stability analysis using videokeratography. *American Journal of Ophthalmology*, 135(5):607–612, 2003.
- P. A. Gounder, E. Cole, S. Colley, and D. M. Hille. Validation of a Portable Electronic Visual Acuity System. *Journal of Mobile Technology in Medicine*, 3(2):35–39, 2014.

- R. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Information processing letters*, 1(4):132–133, 1972.
- M. Grasso, M. Yen, and M. Mintz. Survey of handheld computing among medical students. *Computer Methods and Programs in Biomedicine*, 82(3):196–202, 2006.
- D. Green and J. Swets. *Signal detection theory and psychophysics*. Wiley, New York, 1st edition, 1966.
- H. Gross. Human eye. In *Handbook of Optical Systems: Vol. 4 Survey of Optical Instruments*, volume 4, pages 3–58. 2008.
- J. Guillon. Non-invasive Tearscope Plus routine for contact lens fitting. *Contact Lens and Anterior Eye*, 21(1):531–540, 1998.
- K. Gumus, C. H. Crockett, K. Rao, E. Yeu, M. P. Weikert, M. Shirayama, S. Hada, and S. C. Pflugfelder. Noninvasive assessment of tear stability with the tear stability analysis system in tear dysfunction patients. *Investigative ophthalmology & visual science*, 52(1):456–61, 2011.
- L. J. Haddock, D. Y. Kim, and S. Mukai. Simple, inexpensive technique for high-quality smartphone fundus photography in human and animal eyes. *Journal of ophthalmology*, 2013(1), 2013.
- G. A. Hahn, D. Penka, C. Gehrlich, A. Messias, M. Weismann, L. Hyvärinen, M. Leinonen, M. Feely, G. Rubin, C. Dauxerre, F. Vital-Durand, S. Featherston, K. Dietz, and S. Trauzettel-Klosinski. New standardised texts for assessing reading performance in four European languages. *The British journal of ophthalmology*, 90(4):480–4, 2006.
- D. Hamasaki, J. Ong, and E. Marg. The amplitude of accommodation in presbyopia. *Am J Optom Arch Am Acad Optom*, 33(1):3–14, 1956.
- R. Hartley. *Multiple view geometry in computer vision*. Cambridge university press, Cambridge, 2003.
- C. Hazel, K. Petre, and R. Armstrong. Visual function and subjective quality of life compared in subjects with acquired macular disease. *Investigative ophthalmology & visual science*, 41(6):1309–1315, 2000.
- H. Helmholtz. *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg, Braunschweig, 1865.
- P. Herse and H. Bedell. Contrast sensitivity for letter and grating targets under various stimulus conditions. *Optometry & Vision Science*, 66(11):774–781, 1989.

- K. Higgins, M. Jaffe, R. Caruso, and F. Demonasterio. Spatial contrast sensitivity: effects of age, test-retest, and psychophysical method. *Journal of the Optical Society of America.*, 1988.
- H. Hofstetter. Useful age-amplitude formula. *Optom World*, 38(12):42–45, 1950.
- F. J. Holly. Formation and rupture of the tear film. *Experimental Eye Research*, 15(5):515–525, 1973.
- F. J. Holly and M. A. Lemp. Wettability and wetting of corneal epithelium. *Experimental Eye Research*, 11(2):239–250, 1971.
- F. J. Holly, D. W. Lamberts, and D. L. MacKeen. *The Preocular tear film in health, disease, and contact lens wear*. Dry Eye Institute, Lubbock, 1986.
- J. Hong, X. Sun, A. Wei, X. Cui, Y. Li, and T. Qian. Assessment of tear film stability in dry eye with a newly developed keratograph. *Cornea*, 32(5):716–721, 2013.
- J. Hong, Z. Liu, J. Hua, A. Wei, F. Xue, Y. Yang, X. Sun, and J. Xu. Evaluation of age-related changes in noninvasive tear breakup time. *Optometry and vision science : official publication of the American Academy of Optometry*, 91(2):150–5, 2014.
- E. Hosaka, T. Kawamorita, Y. Ogasawara, N. Nakayama, H. Uozato, K. Shimizu, M. Dogru, K. Tsubota, and E. Goto. Interferometry in the Evaluation of Pre-corneal Tear Film Thickness in Dry Eye. *American Journal of Ophthalmology*, 151(1):18 – 23, 2011.
- F. Hou, C.-B. Huang, L. Lesmes, L.-X. Feng, L. Tao, Y.-F. Zhou, and Z.-L. Lu. qCSF in clinical application: efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative ophthalmology & visual science*, 51(10):5365–77, 2010.
- D. Hubel and T. Wiesel. Receptive fields of single neurones in the cat’s striate cortex. *The Journal of physiology*, 148(3):574–591, 1959.
- T. Ide, K. Negishi, T. Yamaguchi, S. Hara, I. Toda, and K. Tsubota. New Compact Accommodometer to Measure Accommodation Amplitude as a Biomarker. *Asia-Pacific Journal of Ophthalmology*, 1(1):24–27, 2012.
- D. R. Iskander, M. J. Collins, and B. Davis. Evaluating tear film stability in the human eye with high-speed videokeratoscopy. *IEEE transactions on bio-medical engineering*, 52(11):1939–49, 2005.

- F. Jäkel and F. A. Wichmann. Spatial four-alternative forced-choice method is the preferred psychophysical method for naïve observers. *Journal of vision*, 6(11):1307–22, 2006.
- J. Javitt, M. Brenner, and B. Curbow. Outcomes of cataract surgery: improvement in visual acuity and subjective visual function after surgery in the first, second, and both eyes. *Archives of ophthalmology*, 111(5):686–691, 1993.
- K. Jayaraman. Factors influencing the success of platform centric ecosystem strategies: A case study of Google Android. In *The International Conference on E-Technologies and Business on the Web*, pages 212–217. The Society of Digital Information and Wireless Communication, 2013.
- J. P. Jones and L. a. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of neurophysiology*, 58(6):1233–58, 1987.
- L. Jones. Epiphora. II. Its relation to the anatomic structures and surgery of the medial canthal region. *American journal of ophthalmology*, 43(2):203–212, 1957.
- P. K. Kaiser, Y.-Z. Wang, Y.-G. He, A. Weisberger, S. Wolf, and C. H. Smith. Feasibility of a novel remote daily monitoring system for age-related macular degeneration using mobile handheld devices: results of a pilot study. *Retina*, 33(9):1863–70, 2013.
- K. Kamiya, K. Shimizu, D. Aizawa, and H. Ishikawa. Time course of accommodation after implantable collamer lens implantation. *American journal of ophthalmology*, 146(5):674–8, 2008.
- H. Kasprzak. Modeling of the influence of the tear film deterioration on the cornea on the refractive properties of the eye. *Mopane*, 1(1):197–200, 2003.
- P. A. Keane, P. J. Patel, Y. Ouyang, F. K. Chen, F. Ikeji, A. C. Walsh, A. Tufail, and S. R. Sadda. Effects of retinal morphology on contrast sensitivity and reading ability in neovascular age-related macular degeneration. *Investigative ophthalmology & visual science*, 51(11):5431–7, 2010.
- S. A. Kelly, Y. Pang, and S. Klemencic. Reliability of the CSV-1000 in adults and children. *Optometry and vision science : official publication of the American Academy of Optometry*, 89(8):1172–81, 2012.
- A. Khamene, S. Negahdaripour, and S. C. Tseng. A spectral-discrimination method for tear-film lipid-layer thickness estimation from fringe pattern images. *IEEE transactions on bio-medical engineering*, 47(2):249–58, 2000.

- S. Kimball. *Evaporation is the Primary Mechanism of Tear Film Thinning*. PhD thesis, Ohio State University, 2009.
- S. Kimball, P. King-Smith, and J. Nichols. Evidence for the major contribution of evaporation to tear film thinning between blinks. *Investigative ophthalmology & visual science*, 51(12):6294–6297, 2010.
- P. E. King-Smith, B. A. Fink, and N. Fogt. Three interferometric methods for measuring the thickness of layers of the tear film. *Optometry and vision science : official publication of the American Academy of Optometry*, 76(1):19–32, 1999.
- P. E. King-Smith, J. J. Nichols, K. K. Nichols, B. a. Fink, and R. J. Braun. Contributions of evaporation and other mechanisms to tear film thinning and break-up. *Optometry and vision science : official publication of the American Academy of Optometry*, 85(8):623–30, 2008.
- P. E. King-Smith, B. A. Fink, R. M. Hill, K. W. Koelling, and J. M. Tiffany. The thickness of the tear film. *Current eye research*, 29(4-5):357–68, 2009.
- P. E. King-Smith, K. S. Reuter, R. J. Braun, J. J. Nichols, and K. K. Nichols. Tear film breakup and structure studied by simultaneous video recording of fluorescence and tear film lipid layer images. *Investigative ophthalmology & visual science*, 54(7):4900–9, 2013.
- C. G. Kiss, T. Barisani-Asenbauer, S. Maca, S. Richter-Mueksch, and W. Radner. Reading performance of patients with uveitis-associated cystoid macular edema. *American journal of ophthalmology*, 142(4):620–4, 2006.
- S. Kleen and R. Levoy. Low vision care: correlation of patient age, visual goals, and aids prescribed. *Journal of optometry and physiological optics*, 58(3):200, 1981.
- M. Knorz and D. Claessens. Evaluation of contrast acuity and defocus curve in bifocal and monofocal intraocular lenses. *Journal of Cataract & Refractive Surgery*, 19(4):513–523, 1993.
- S. Koh, N. Maeda, T. Kuroda, Y. Hori, H. Watanabe, T. Fujikado, Y. Tano, Y. Hirohara, and T. Mihashi. Effect of tear film break-up on higher-order aberrations measured with wavefront sensor. *American journal of ophthalmology*, 134(1):115–7, 2002.
- T. Kojima, R. Ishida, M. Dogru, E. Goto, and Y. Takano. A New Noninvasive Tear Stability Analysis System for the Assessment of Dry Eyes. *Investigative Ophthalmology*, 45(5):1369–1374, 2004.

- P. Kollbaum and M. Jansen. Validation of an iPad Test of Letter Contrast Sensitivity. *Optometry & Vision Science*, 91(3):291–296, 2014.
- M. Kopf, F. Yi, D. Robert Iskander, M. J. Collins, A. J. Shaw, and B. Straker. Tear Film Surface Quality with Soft Contact Lenses Using Dynamic Videokeratometry. *Journal of Optometry*, 1(1):14–21, 2008.
- D. R. Korb. Survey of preferred tests for diagnosis of the tear film and dry eye. *Cornea*, 19(4):483–6, 2000.
- J. Koretz, P. Kaufman, M. Neider, and P. Goeckner. Accommodation and presbyopia in the human eye aging of the anterior segment. *Vision research*, 29(12):1685–1692, 1989.
- R. Krueger, M. Mrochen, M. Kaemmerer, and T. Seiler. Understanding refraction and accommodation through 'retinal imaging' aberrometry: A case report. *Ophthalmology*, 108(4):674–678, 2001.
- Y. Kuo, S. Verma, T. Schmid, and P. Dutta. Hijacking power and bandwidth from the mobile phone's audio interface. In *Proceedings of the First ACM Symposium on Computing for Development*, page 24. ACM, 2010.
- D. Laming. Contrast discrimination by the methods of adjustment and two-alternative forced choice. *Attention, perception & psychophysics*, 75(8):1774–82, 2013.
- S. J. Leat and G. C. Woo. The validity of current clinical tests of contrast sensitivity and their ability to predict reading speed in low vision. *Eye*, 11(6):893–9, 1997.
- B. Lee, G. Lee, Y. Shim, and A. Choi. Let developers run into the app store by lowering the barrier-to-entry. *International Journal of Electronic Finance*, 4(3):201–220, 2010.
- G. Legge, T. Klitz, and B. Tjan. Mr. Chips: an ideal-observer model of reading. *Psychological review*, 104(3):524, 1997.
- G. E. Legge and J. M. Foley. Contrast masking in human vision. *Journal of the Optical Society of America*, 70(12):1458, 1980.
- G. E. Legge, D. G. Pelli, G. S. Rubin, and M. M. Schleske. Psychophysics of reading I. Normal vision. *Vision research*, 25(2):239–52, 1985a.
- G. E. Legge, G. S. Rubin, D. G. Pelli, and M. M. Schleske. Psychophysics of reading II Low vision. *Vision Research*, 25(2):253–265, 1985b.

- G. E. Legge, J. A. Ross, and K. T. Maxwell. Psychophysics of reading. VII. Comprehension in normal and low vision. *Clinical Vision Sciences*, 4(1):51–60, 1989.
- A. A. León, S. M. Medrano, and M. Rosenfield. A comparison of the reliability of dynamic retinoscopy and subjective measurements of amplitude of accommodation. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 32(2):133–41, 2012.
- L. A. Lesmes, Z.-l. Lu, and T. D. Albright. Bayesian adaptive estimation of the contrast sensitivity function : The quick CSF method. *Journal of Vision*, 10(3):17, 2010.
- W. Licznarski, Tomasz J and Kasprzak, Henryk T and Kowalik. Application of Twyman Green interferometer for evaluation of in vivo breakup characteristic of the human tear film. *Journal of biomedical optics*, 4(1):176–182, 1999.
- S.-J. Lin, C.-M. Yang, P.-T. Yeh, and T.-C. Ho. Smartphone funduscopy for retinopathy of prematurity. *Taiwan Journal of Ophthalmology*, 4(2):82–85, 2014.
- W. Lotmar. Theoretical Eye Model with Aspherics. *Journal of the Optical Society of America*, 61(11):1522, 1971.
- A. Luebker and J. Lamay. Psychophysics of reading. VIII. The Minnesota low vision reading test. *Optometry & Vision Science*, 37(14):843–853, 1989.
- K. Maaijwee, P. Mulder, and W. Radner. Reliability testing of the Dutch version of the Radner Reading Charts. *Optometry & Vision Science*, 85(5):353–358, 2008.
- R. N. Maamari, J. D. Keenan, D. a. Fletcher, and T. P. Margolis. A mobile phone-based retinal camera for portable wide field imaging. *The British journal of ophthalmology*, 98(4):438–41, 2014.
- D. MacLeod, J. Bergsma, and E. Vul. Functional Adaptive Sequential Testing. *Seeing and Perceiving*, 23(5):483–515, 2010.
- N. Macmillan and C. Creelman. *Detection theory: A user's guide*. LEA, Mahwah NJ, 2nd edition, 2004.
- C. Maïssa and M. Guillon. Tear film dynamics and lipid layer characteristics—effect of age and gender. *Contact lens & anterior eye : the journal of the British Contact Lens Association*, 33(4):176–82, 2010.
- E. A. Mallen, J. S. Wolffsohn, B. Gilmartin, and S. Tsujimura. Clinical evaluation of the Shin-Nippon SRW-5000 autorefractor in adults. *Ophthalmic & physiological*

- optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 21(2):101–7, 2001.
- C. Mangione and R. Phillips. Improved visual function and attenuation of declines in health-related quality of life after cataract extraction. *Archives of ophthalmology*, 112(11):1419–1425, 1994.
- J. S. Mansfield, S. J. Ahn, G. E. Legge, and A. Luebker. A new reading-acuity chart for normal and low vision. *Ophthalmic and Visual Optics/Noninvasive Assessment of the Visual System Technical Digest*, 3:232–235, 1993.
- B. Martin, A. Gilpin, D. Jabs, and A. Wu. Reliability, validity, and responsiveness of general and disease-specific quality of life measures in a clinical trial for cytomegalovirus retinitis. *Journal of clinical epidemiology*, 54(4):376–386, 2001.
- P. Maudgal, R. Stout, and A. V. Balen. VCTS chart evaluation as a screening test. *Documenta ophthalmologica*, 64(9):399–405, 1988.
- J. McAnany and K. Alexander. Contrast sensitivity for letter optotypes vs. gratings under conditions biased toward parvocellular and magnocellular pathways. *Vision research*, 46(10):1574–1584, 2006.
- J. P. McCulley and W. E. Shine. Meibomian gland function and the tear lipid layer. *The ocular surface*, 1(3):97–106, 2003.
- J. E. McDonald and S. Brubaker. Meniscus-induced thinning of tear films. *American journal of ophthalmology*, 72(1):139–46, 1971.
- M. Meeker, R. Du, and P. Bacchetti. Pupil examination: validity and clinical utility of an automated pupillometer. *Journal of Neuroscience Nursing*, 37(1):35–40, 2005.
- Y. Mejia-Barbosa and D. Malacara-Hernandez. A review of methods for measuring corneal topography. *Optometry & Vision Science*, 78(4):240–253, 2001.
- L. S. Mengher, A. J. Bron, S. R. Tonge, and D. J. Gilbert. Effect of fluorescein instillation on the pre-corneal tear film stability. *Current eye research*, 4(1):9–12, 1985.
- B. Miljanović, R. Dana, D. A. Sullivan, and D. A. Schaumberg. Impact of dry eye syndrome on vision-related quality of life. *American journal of ophthalmology*, 143(3):409–15, 2007.

- R. Montes-Mico. Temporal Changes in Optical Quality of Air-Tear Film Interface at Anterior Cornea after Blink. *Investigative Ophthalmology & Visual Science*, 45(6):1752–1757, 2004.
- R. Montés-Micó and J. L. Alió. Distance and near contrast sensitivity function after multifocal intraocular lens implantation. *Journal of Cataract & Refractive Surgery*, 29(4):703–711, 2003.
- R. Montés-Micó, J. L. Alió, and W. N. Charman. Dynamic changes in the tear film in dry eyes. *Investigative ophthalmology & visual science*, 46(5):1615–9, 2005.
- Y. Morad, E. Werker, and P. Nemet. Visual acuity tests using chart, line, and single optotype in healthy and amblyopic children. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 3(2):94–97, 1999.
- E. Morales and K. Rocha. Comparison of optical aberrations and contrast sensitivity between aspheric and spherical intraocular lenses. *Journal of refractive surgery*, 27(10):723–728, 2011.
- F. Moret, C. M. Poloschek, W. a. Lagrèze, and M. Bach. Visualization of fundus vessel pulsation using principal component analysis. *Investigative ophthalmology & visual science*, 52(8):5457–64, 2011.
- S. E. Moss, R. Klein, and B. E. Klein. Prevalence of and risk factors for dry eye syndrome. *Archives of ophthalmology*, 118(9):1264–8, 2000.
- J. Movshon and W. Newsome. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *The Journal of neuroscience*, 16(23):7733–7741, 1996.
- R. Mutharasan and S. P. Srinivas. Topographic Distribution Of Blink Induced Shear Stress (BLISS) On The Corneal Surface. *Investigative Ophthalmology and Visual Science*, 43(12):974, 2002.
- D. Myung, A. Jais, L. He, M. S. Blumenkranz, and R. T. Chang. 3D Printed Smartphone Indirect Lens Adapter for Rapid, High Quality Retinal Imaging. *Journal of Mobile Technology in Medicine*, 3(1):9–15, 2014a.
- D. Myung, A. Jais, L. He, and R. T. Chang. Simple, Low-Cost Smartphone Adapter for Rapid, High Quality Ocular Anterior Segment Imaging: A Photo Diary. *Journal of Mobile Technology in Medicine*, 3(1):2–8, 2014b.
- M. Nakazawa and K. Ohtsuki. Apparent accommodation in pseudophakic eyes after

- implantation of posterior chamber intraocular lenses: optical analysis. *Invest. Ophthalmol. Vis. Sci.*, 25(12):1458–1460, 1984.
- National Academy of Sciences. Recommended standard procedures for the clinical measurement and specification of visual acuity. *Report of working group 39. Committee on vision. Assembly of Behavioral and Social Sciences, National Research Council, National Academy of Sciences, Washington, D.C.*, 1(1):103–148, 1980.
- G. Nemeth, A. Tsorbatzoglou, P. Vamosi, Z. Sohajda, and A. Berta. A comparison of accommodation amplitudes in pseudophakic eyes measured with three different methods. *Eye*, 22(1):65–9, 2008.
- J. Németh, B. Erdélyi, B. Csákány, P. Gáspár, A. Soumelidis, F. Kahlesz, and Z. Lang. High-speed videotopographic measurement of tear film build-up time. *Investigative ophthalmology & visual science*, 43(6):1783–90, 2002.
- W. Niblack. *An introduction to digital image processing*. Strandberg Publishing Company, Birkerød, 1st edition, 1985.
- J. J. Nichols, G. L. Mitchell, and P. E. King-Smith. Thinning rate of the precorneal and prelens tear films. *Investigative ophthalmology & visual science*, 46(7):2353–61, 2005.
- K. Nichols. The repeatability of clinical measurements of dry eye. *Cornea*, 23(3):272–285, 2004.
- OFCOM. Communications Market Report 2014. Technical report, London (UK), 2014.
- M. O’Mahony. Understanding Discrimination Tests: A User-Friendly Treatment Of Response Bias, Rating And Ranking R-Index Tests And Their Relationship To Signal Detection. *Journal of Sensory Studies*, 7(1):1–47, 1992.
- G. Öquist and M. Goldstein. Towards an improved readability on mobile devices: evaluating adaptive rapid serial visual presentation. *Interacting with Computers*, 15(4):539–558, 2003.
- L. A. Ostrin and A. Glasser. Accommodation measurements in a prepresbyopic and presbyopic population. *Journal of cataract and refractive surgery*, 30(7):1435–44, 2004.
- N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 3(11):23–27, 1975.

- H. Owens and J. Phillips. Spreading of the tears after a blink: velocity and stabilization time in healthy eyes. *Cornea*, 20(5):484–487, 2001.
- C. Owsley. Contrast sensitivity. *Ophthalmology Clinics of North America*, 16(2):171–177, 2003.
- C. Owsley and M. E. Sloane. Contrast sensitivity, acuity, and the perception of ‘real-world’ targets. *The British journal of ophthalmology*, 71(10):791–6, 1987.
- G. R. Patching and T. R. Jordan. Spatial Frequency Sensitivity Differences between Adults of Good and Poor Reading Ability. *Investigative Ophthalmology & Visual Science*, 46(6):2219–2224, 2005.
- P. J. Patel, F. K. Chen, L. Da Cruz, G. S. Rubin, and A. Tufail. Test-retest variability of reading performance metrics using MNREAD in patients with age-related macular degeneration. *Investigative ophthalmology & visual science*, 52(6):3854–9, 2011.
- N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan, and I. J. Constable. Retinal image analysis: concepts, applications and potential. *Progress in retinal and eye research*, 25(1):99–127, 2006.
- K. Payne, H. Wharrad, and K. Watts. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC medical informatics and decision making*, 12(1):121, 2012.
- D. Pelli and J. Robson. The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 10(1), 1988.
- C. A. Perez, J. R. Donoso, and L. E. Medina. A critical experimental study of the classical tactile threshold theory. *BMC neuroscience*, 11(1):76, 2010.
- G. M. Pérez, S. M. Archer, and P. Artal. Optical characterization of Bangerter foils. *Investigative ophthalmology & visual science*, 51(1):609–13, 2010.
- J. J. Pérez-Santonja, H. F. Sakla, and J. L. Alió. Contrast sensitivity after laser in situ keratomileusis. *Journal of Cataract & Refractive Surgery*, 24(2):183–189, 1998.
- K. Pesudovs and C. Hazel. The usefulness of Vistech and FACT contrast sensitivity charts for cataract and refractive surgery outcomes research. *British Journal of Ophthalmology*, 88(1):11–17, 2004.

- K. Pesudovs, B. Patel, J. a. Bradbury, and D. B. Elliott. Reading speed test for potential central vision measurement. *Clinical & experimental ophthalmology*, 30(3):183–6, 2002.
- R. Peterson, J. Wolffsohn, and C. Fowler. Optimization of anterior eye fluorescein viewing. *American journal of ophthalmology*, 142(4):572—575, 2006.
- D. H. Peterzell and D. Y. Teller. Individual differences in contrast sensitivity functions: the lowest spatial frequency channels. *Vision research*, 36(19):3077–85, 1996.
- S. Pflugfelder, R. Beuerman, and M. Stern. *Dry eye and ocular surface disorders*. Marcel Dekker, New York, 2004.
- S. C. Pflugfelder, S. C. Tseng, O. Sanabria, H. Kell, C. G. Garcia, C. Felix, W. Feuer, and B. L. Reis. Evaluation of subjective assessments and objective diagnostic tests for diagnosing tear-film disorders known to cause ocular irritation. *Cornea*, 17(1):38–56, 1998.
- A. Pfützner, M. Mitri, P. B. Musholt, D. Sachsenheimer, M. Borchert, A. Yap, and T. Forst. Clinical assessment of the accuracy of blood glucose measurement devices. *Current Medical Research & Opinion*, 28(4):525–531, 2012.
- PGBIZ. App store metrics by category- 2014. Technical report, 2014.
- C. Poynton. *Digital Video And HDTV Algorithms And Interfaces*. Morgan Kaufmann, Boston, 1st edition, 2003.
- P. Prasanna. Decision Support System for Detection of Diabetic Retinopathy Using Smartphones. In *7th International Conference on Pervasive Computing Technologies for Healthcare*, pages 179–179, 2013.
- J. I. Prydal, P. Artal, H. Woon, and F. W. Campbell. Study of human precorneal tear film thickness and structure using laser interferometry. *Investigative ophthalmology & visual science*, 33(6):2006–11, 1992.
- H. Pult and B. Riede-Pult. The new TF-Scan application for video keratography: its repeatability and usefulness in tear film break-up time analysis. *Contact Lens and Anterior Eye*, 34(2011):S14, 2011.
- H. Pult and B. H. Riede-Pult. A new modified fluorescein strip: Its repeatability and usefulness in tear film break-up time analysis. *Contact lens & anterior eye : the journal of the British Contact Lens Association*, 35(1):35–8, 2012.

- H. Pult, C. Purslow, and P. J. Murphy. The relationship between clinical signs and dry eye symptoms. *Eye*, 25(4):502–10, 2011.
- P. W. Radner and W. Radner. Introducing a new reading chart. *Ophthalmology Times*, 4(2), 2008.
- W. Radner, W. Obermayer, S. Richter-Mueksch, U. Willinger, M. Velikay-Parel, and B. Eisenwort. The validity and reliability of short German sentences for measuring reading speed. *Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv für klinische und experimentelle Ophthalmologie*, 240(6):461–7, 2002.
- W. Radner, E. Stifter, and S. Richter-Mueksch. Standards of reading performance and visual acuity measurements in maculopathy-reliability and validity analyses of vision tests. In *The Macula: Diagnosis, Treatment and Future Trends*, pages 253–260. 2004.
- P. Y. Ramulu, B. K. Swenor, J. L. Jefferys, and G. S. Rubin. Description and validation of a test to evaluate sustained silent reading. *Investigative ophthalmology & visual science*, 54(1):673–80, 2013.
- K. Rayner and S. Duffy. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3):191–201, 1986.
- B. Reeves, J. Wood, and A. Hill. Vistech VCTS 6500 charts-within-and between-session reliability. *Optometry and vision science*, 68(9):728–737, 1991.
- D. Regan and K. Beverley. Spatial-frequency discrimination and detection: comparison of postadaptation thresholds. *JOSA*, 73(12):1684–1690, 1983.
- B. Remeseiro, M. Penas, N. Barreira, A. Mosquera, J. Novo, and C. García-Resúa. Automatic classification of the interferential tear film lipid layer using colour texture analysis. *Computer methods and programs in biomedicine*, 111(1):93–103, 2013.
- H. W. Ren and G. Wilson. The effect of a shear force on the cell shedding rate of the corneal epithelium. *ACTA OPHTHALMOLOGICA SCANDINAVICA*, 75(4):383–387, 1997.
- M. L. Rice, E. E. Birch, and J. M. Holmes. An abbreviated reading speed test. *Optometry and vision science : official publication of the American Academy of Optometry*, 82(2):128–33, 2005.

- A. Riusala, S. Sarna, and I. Immonen. Visual function index (VF-14) in exudative age-related macular degeneration of long duration. *American journal of ophthalmology*, 135(2):206–212, 2003.
- M. Rolando and M. Zierhut. The ocular surface and tear film and their dysfunction in dry eye disease. *Survey of ophthalmology*, 45(2):S203–10, 2001.
- E. Rosen, C. Gore, and D. Taylor. Use of a digital infrared pupillometer to assess patient suitability for refractive surgery. . . . & *Refractive Surgery*, 2002.
- M. Rosenfield and A. S. Cohen. Push-up amplitude of accommodation and target size. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 15(3):231–2, 1995a.
- M. Rosenfield and A. S. Cohen. Push-up amplitude of accommodation and target size. *Ophthalmic and Physiological Optics*, 15(3):231–232, 1995b.
- M. Rosenfield and a. S. Cohen. Repeatability of clinical measurements of the amplitude of accommodation. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 16(3):247–9, 1996.
- M. Rosenfield and B. Gilmartin. Effect of target proximity on the open-loop accommodative response. *Optometry & Vision Science*, 67(2):74–79, 1990.
- B. Rosengren. On lacrimal drainage. *Ophthalmologica*, 164(6):409–421, 1972.
- J. E. Ross, D. D. Clarke, and a. J. Bron. Effect of age on contrast sensitivity function: unocular and binocular findings. *The British journal of ophthalmology*, 69(1):51–6, 1985.
- G. S. Rubin. Measuring reading performance. *Vision research*, 90:43–51, 2013.
- G. S. Rubin and K. Turano. Reading without saccadic eye movements. *Vision Research*, 32(5):895–902, 1992.
- S. Sahlin and E. Chen. Gravity, blink rate, and lacrimal drainage capacity. *American journal of ophthalmology*, 124(6):758–64, 1997.
- H. Savin. Word Frequency Effect and Errors in the Perception of Speech. *The Journal of the Acoustical Society of America*, 35(2):200–206, 1963.
- M. Scheiman and B. Wick. *Clinical management of binocular vision : heterophoric, accommodative, and eye movement disorders*. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2008.

- B. Seifert and W. Vilser. Retinal Vessel Analyzer (RVA)-design and function. *Biomedizinische Technik/Biomedical . . .*, 47:678–681, 2002.
- A. Sharma and E. Ruckenstein. Mechanism of tear film rupture and its implications for contact lens tolerance. *American journal of optometry and physiological optics*, 62(4):246–53, 1985.
- C. Sheard. Dynamic Skiametry and Methods of Testing the Accommodation and Convergence of the Eyes. In *Physiological optics*, page 108. Cleveland Press, Chicago, 1920.
- M. Shiobara and C. Schnider. Guide to the clinical assessment of on-eye wettability of rigid gas permeable lenses. *Optometry & Vision Science*, 66(4):202–206, 1989.
- J. Smith, K. Nichols, and E. Baldwin. Current patterns in the use of diagnostic tests in dry eye evaluation. *Cornea*, 27(6):656–662, 2008.
- I. Sobel. An isotropic 3 by 3 image gradient operator. *Machine Vision for three-dimensional Sciences*, 1(1):23–34, 1990.
- W. Somers and C. Ford. Effect of relative distance magnification on the monocular amplitude of accommodation. *American Journal of Optometry and Physiological Optics*, 60(11):920–924, 1983.
- B. Sterner. *Ocular accomodation. Studies of amplitude, insufficiency, and facility training in young school children*. PhD thesis, University of Gothenburg, 2004.
- B. Sterner, M. Gellerstedt, and A. Sjöström. The amplitude of accommodation in 6-10-year-old children - not as good as expected! *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 24(3):246–51, 2004.
- E. Stifter, F. König, T. Lang, P. Bauer, S. Richter-Müksch, M. Velikay-Parel, and W. Radner. Reliability of a standardized reading chart system: variance component analysis, test-retest and inter-chart reliability. *Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv für klinische und experimentelle Ophthalmologie*, 242(1):31–9, 2004a.
- E. Stifter, S. Sacu, H. Weghaupt, F. König, S. Richter-Müksch, A. Thaler, M. Velikay-Parel, and W. Radner. Reading performance depending on the type of cataract and its predictability on the visual outcome. *Journal of cataract and refractive surgery*, 30(6):1259–67, 2004b.

- E. Stifter, H. Weghaupt, T. Benesch, A. Thaler, and W. Radner. Discriminative power of reading tests to differentiate visual impairment caused by cataract and age-related macular degeneration. *Journal of cataract and refractive surgery*, 31(11):2111–9, 2005.
- C. F. Stromeyer and B. Julesz. Spatial-frequency masking in vision: critical bands and spread of masking. *Journal of the Optical Society of America*, 62(10):1221–32, 1972.
- R. Superstein, D. Boyaner, and O. Overbury. Functional complaints, visual acuity, spatial contrast sensitivity, and glare disability in preoperative and postoperative cataract patients. *Journal of cataract and refractive surgery*, 25(4):575–81, 1999.
- D. F. Sweeney, T. J. Millar, and S. R. Raju. Tear film stability: a review. *Experimental eye research*, 117:28–38, 2013.
- J. Swets, W. T. Jr, and T. Birdsall. Decision processes in perception. *Psychological review*, 68(5):301, 1961.
- D. Szczesna and D. Iskander. Lateral shearing interferometry for analysis of tear film surface kinetics. *Optometry & Vision Science*, 87(7):513–517, 2010.
- D. H. Szczesna, D. Alonso-Caneiro, D. R. Iskander, S. a. Read, and M. J. Collins. Predicting dry eye using noninvasive techniques of tear film surface assessment. *Investigative ophthalmology & visual science*, 52(2):751–6, 2011.
- D. H. Szczesna-Iskander and D. R. Iskander. Future directions in non-invasive measurements of tear film surface kinetics. *Optometry and vision science : official publication of the American Academy of Optometry*, 89(5):749–59, 2012.
- W. P. Tanner and J. a. Swets. A decision-making theory of visual detection. *Psychological review*, 61(6):401–9, 1954.
- K. Thayaparan, M. D. Crossland, and G. S. Rubin. Clinical assessment of two new contrast sensitivity charts. *The British journal of ophthalmology*, 91(6):749–52, 2007.
- A. Tomlinson and S. Khanal. Assessment of tear film dynamics: quantification approach. *The ocular surface*, 3(2):81–95, 2005.
- A. Tomlinson, M. G. Doane, and A. Mcfadyen. Inputs and Outputs of the Lacrimal system: Review of Production and Evaporative Loss. *The Ocular Surface*, 7(4):186–198, 2009.

- K. N. Toner, M. J. Lynn, T. R. Candy, and A. K. Hutchinson. The Handy Eye Check: a mobile medical application to test visual acuity in children. *Journal of AAPOS : the official publication of the American Association for Pediatric Ophthalmology and Strabismus / American Association for Pediatric Ophthalmology and Strabismus*, 18(3):258–60, 2014.
- R. Tootell and M. Silverman. Functional anatomy of macaque striate cortex. V. Spatial frequency. *The Journal of neuroscience*, 8(5):1610–1624, 1988.
- S. Trauzettel-Klosinski, K. Dietz, and L. Vision. Standardized assessment of reading performance: the New International Reading Speed Texts IReST. *Investigative ophthalmology & visual science*, 53(9):5452–61, 2012.
- K. Tsubota. Tear dynamics and dry eye. *Progress in retinal and eye research*, 17(4):565–596, 1998.
- J. Tucker and W. N. Charman. Depth of focus and accommodation for sinusoidal gratings as a function of luminance. *American journal of optometry and physiological optics*, 63(1):58–70, 1986.
- A. Turpin, D. Lawson, and A. McKendrick. PsyPad: A platform for visual psychophysics on the iPad. *Journal of vision*, 14(3):1–7, 2014.
- R. Tutt, a. Bradley, C. Begley, and L. N. Thibos. Optical and visual impact of tear break-up in human eyes. *Investigative ophthalmology & visual science*, 41(13):4117–23, 2000.
- G. Tyagi, D. Alonso-Caneiro, M. Collins, and S. Read. Tear film surface quality with rigid and soft contact lenses. *Eye & contact lens*, 38(3):171–8, 2012.
- C. Tyler. Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial vision*, 10(4):269–377, 1997.
- C. Tyler and H. Chan. Bit stealing: how to get 1786 or more gray levels from an 8-bit color monitor. In *SPIE Human Vision, Visual Processing and Digital Display 1992 Symposium on Electronic Imaging: Science and Technology*, volume 1666, pages 351–364. International Society for Optics and Photonics, 1992.
- Vaegan and B. L. Halliday. A forced-choice test improves clinical contrast sensitivity testing. *The British journal of ophthalmology*, 66(8):477–91, 1982.
- K. D. Valois. Spatial frequency adaptation can enhance contrast sensitivity. *Vision research*, 17(9):1057–1065, 1977.

- R. D. Valois, D. Albrecht, and L. Thorell. Spatial frequency selectivity of cells in macaque visual cortex. *Vision research*, 22(5):545–559, 1982.
- N. Van Haeringen. Clinical biochemistry of tears. *Survey of Ophthalmology*, 26(2): 84–96, 1981.
- P. P. van Saarloos and I. J. Constable. Improved method for calculation of corneal topography for any photokeratoscope geometry. *Optometry and vision science : official publication of the American Academy of Optometry*, 68(12):960–5, 1991.
- J. Varikooty, N. Keir, and T. Simpson. Estimating tear film spread and stability through tear hydrodynamics. *Optometry & Vision Science*, 89(8):1119–1124, 2012.
- P. Viola and M. Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.
- B. Wang and K. J. Ciuffreda. Depth-of-focus of the human eye: theory and clinical implications. *Survey of ophthalmology*, 51(1):75–85, 2006.
- Y.-Z. Wang, Y.-G. He, G. Mitzel, S. Zhang, and M. Bartlett. Handheld shape discrimination hyperacuity test on a mobile device for remote monitoring of visual function in maculopathy. *Investigative ophthalmology & visual science*, 54(8): 5497–505, 2013.
- A. Watson and D. Pelli. QUEST: A Bayesian adaptive psychometric method. *Perception & psychophysics*, 33(2):113–120, 1983.
- A. Watson and J. Robson. Discrimination at threshold: labelled detectors in human vision. *Vision research*, 21(7):1115–1122, 1981.
- R. G. Weatherhead. Use of the Arden grating test for screening. *The British journal of ophthalmology*, 64(8):591–6, 1980.
- R. N. Weinreb and P. L. Kaufman. Glaucoma research community and FDA look to the future, II: NEI/FDA Glaucoma Clinical Trial Design and Endpoints Symposium: measures of structural change and visual function. *Investigative ophthalmology & visual science*, 52(11):7842–51, 2011.
- M. Wender. Value of Pelli Robson contrast sensitivity chart for evaluation of visual system in multiple sclerosis patients. *Neurologia i neurochirurgia polska*, 41(2): 141 – 143, 2006.
- S. Whittaker and J. Lovie-Kitchin. Visual requirements for reading. *Optometry & Vision Science*, 70(1):54–65, 1993.

- B. Wick and P. Hall. Relation among accommodative facility, lag, and amplitude in elementary school children. *American journal of optometry and physiological optics*, 64(8):593–8, 1987.
- L. Wieder, G. Gäde, L. M. Pech, H. Zimmermann, K.-D. Wernecke, J.-M. Dörr, J. Bellmann-Strobl, F. Paul, and A. U. Brandt. Low contrast visual acuity testing is associated with cognitive performance in multiple sclerosis: a cross-sectional pilot study. *BMC neurology*, 13(1):167, 2013.
- F. Wilkinson, H. R. Wilson, and C. Habak. Detection and recognition of radial frequency patterns. *Vision Research*, 38(22):3555–3568, 1998.
- H. Wilson, D. McFarlane, and G. Phillips. Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision research*, 23(9):873–882, 1983.
- J. E. Wold, A. Hu, S. Chen, and A. Glasser. Subjective and objective measurement of human accommodative amplitude. *Journal of Cataract & Refractive Surgery*, 29(10):1878–1888, 2003.
- R. M. Wold. The spectacle amplitude of accommodation of children aged six to ten. *American journal of optometry and archives of American Academy of Optometry*, 44(10):642–64, 1967.
- E. Wolff. *The anatomy of the eye and orbit*. HK lewis and Co, London, fourth edition, 1954.
- J. S. Wolffsohn, O. A. Hunt, S. Naroo, B. Gilmartin, S. Shah, I. A. Cunliffe, M. T. Benson, and S. Mantry. Objective accommodative amplitude and dynamics with the 1CU accommodative intraocular lens. *Investigative ophthalmology & visual science*, 47(3):1230–5, 2006.
- J. S. Wolffsohn, L. N. Davies, N. Gupta, S. A. Naroo, G. A. Gibson, T. Mihashi, and S. Shah. Mechanism of action of the tetraflex accommodative intraocular lens. *Journal of refractive surgery (Thorofare, N.J. : 1995)*, 26(11):858–62, 2010.
- J. S. Wolffsohn, L. N. Davies, S. A. Naroo, P. J. Buckhurst, G. A. Gibson, N. Gupta, J. P. Craig, and S. Shah. Evaluation of an open-field autorefractor’s ability to measure refraction and hence potential to assess objective accommodation in pseudophakes. *The British journal of ophthalmology*, 95(4):498–501, 2011.
- C. Wong. iPhone ECG monitoring—the gateway to the new paradigm of STEMI therapy. *International journal of cardiology*, 168(3):2897–2898, 2013.

- R. Woods and W. Thomson. A comparison of psychometric methods for measuring the contrast sensitivity of experienced observers. *Clinical vision sciences*, 8(5):401–415, 1993.
- R. L. Woods and W. D. Thomson. Effects of exercise on aspects of visual function. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 15(1):5–12, 1995.
- C. E. Wright and N. Drasdo. The influence of age on the spatial and temporal contrast sensitivity function. *Documenta Ophthalmologica*, 59(4):385–395, 1985.
- Y. Ye, J. Wang, Y. Xie, J. Zhong, Y. Hu, B. Chen, X. He, and H. Zhang. Global Teleophthalmology With iPhones for Real-Time Slitlamp Eye Examination. *Eye & contact lens*, 40(5):297–300, 2014.
- T. Yedidya, R. Hartley, J.-P. Guillon, and Y. Kanagasingam. Automatic dry eye detection. In *Medical image computing and computer-assisted intervention: MICCAI 2007*, volume 10, pages 792–9, 2007a.
- T. Yedidya, R. Hartley, J.-P. Guillon, and Y. Kanagasingam. Automatic dry eye detection. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 10(Pt 1):792–9, 2007b.
- T. Yedidya, R. Hartley, and J.-P. Guillon. Automatic Detection of Pre-ocular Tear Film Break-Up Sequence in Dry Eyes. In *Digital Image Computing: Techniques and Applications (DICTA), 2008*, pages 442–448. Ieee, 2008.
- S.-Y. Yu. Reliability of smartphone-based electronic visual acuity testing- (applications in remote monitoring and clinical research of macular pathology). *Investigative ophthalmology & visual science*, 55(9), 2014.
- Z.-t. Zhang, S.-c. Zhang, X.-g. Huang, and L.-y. Liang. A pilot trial of the iPad tablet computer as a portable device for visual acuity testing. *Journal of telemedicine and telecare*, 19(1):55–9, 2013.
- E. Zvornicanin, J. Zvornicanin, and B. Hadziefendic. The Use of Smart phones in Ophthalmology. *Acta informatica medica : časopis Društva za medicinsku informatiku BiH*, 22(3):206–9, 2014.