
Designing For An Individual's Eyes: Human-Computer Interaction, Vision And Individual Differences

by

Mike Bennett

A Thesis submitted to
University College Dublin, Ireland
for the degree of Ph. D.
in the
College of Engineering,
Mathematical & Physical Sciences

June 2009

School of Computer Science & Informatics
J. Carthy, Ph. D. (Head of School)
Under the supervision of
Aaron Quigley, Ph. D.

CONTENTS

Abstract	xii
Acknowledgements	xiv
List of Publications	xv
1 Introduction	1
1.1 Motivation & Purpose	1
1.2 Mass Customisation & Personalisation	5
1.2.1 Adapting At Source	5
1.2.2 What Adapts	6
1.3 Application & Design Scenarios	7
1.3.1 Self-Adapting Signage	8
1.3.2 Visual Popout Paint For Car Safety	8
1.3.3 Audience Aware Presentations	9
1.4 Approach: Model, Predict & Adapt	10
1.4.1 Model	10
1.4.2 Predict	10
1.4.3 Adapt	12
1.5 Hypothesis	12
1.6 Contributions	13
1.6.1 Simplified Models Of Physiological Eye Function For HCI	14
1.6.2 Degree And Location Of Perceptual Stability Due To Spatial Perception	14
1.6.3 Degree And Location Of Perceptual Stability Due To Colour Perception	14
1.6.4 Optometry Models Of Individual Eye Function Integ- rated With Perception	15
1.6.5 Demonstrated Predictors Used To Adapt Information Visualisations & Interfaces To Individual Eye Function	15
1.7 Summary	15
2 Background	17
2.1 Physiology As A Design Context & Constraint	17

2.2	Human-Computer Interaction & Interaction Design	20
2.2.1	Information Visualisation	21
2.2.2	Individual Differences Applied To HCI	22
2.3	Individual Differences In Eye Function	23
2.3.1	Differences In Spatial Perception	25
2.3.2	Differences In Color Perception	27
2.4	Modelling Human Vision	29
2.4.1	Optical Aberrations & Refractive Errors	29
2.4.2	Point Spread Functions	30
2.4.3	Wavefronts	31
2.4.4	Zernike Polynomials Overview	32
2.4.5	Modelling Colour Deficient Vision (CDV)	35
2.5	Sensation & Perception In Vision	36
2.6	Summary	37
3	Modelling & Simulating Human Vision	38
3.1	Introduction	38
3.2	Human Visual System Models	40
3.2.1	Model Requirements	41
3.3	What Was Modeled	44
3.3.1	Differences in Spatial Visual Acuity	44
3.3.2	Differences in Colour Perception	44
3.4	How Did We Model	45
3.4.1	Feature Independent	45
3.4.2	By Reference	47
3.5	Modelling Spatial Visual Acuity	49
3.5.1	Optical Aberrations & Refractive Errors	50
3.5.2	Aberration Maps As Point Spread Functions	51
3.5.3	Biological Fidelity With Zernike Polynomials	52
3.5.4	Simulating Retinal Images: Visual Acuity	53
3.6	Modelling Colour Perception	56
3.6.1	N-Dimensional Colour Spaces	57
3.6.2	Colour Appearance Models	58
3.6.3	Perceptually Uniform Colour Spaces	58
3.6.4	Colour-Appearance Phenomena	61
3.6.5	Measuring Colour Differences	62
3.6.6	Point / Pixel Measures	63
3.6.7	Generating Retinal Images: Colour Perception	64
3.7	Individual Differences	66
3.7.1	Visual Acuity	66
3.7.2	Colour Deficient Vision	67
3.8	Conclusions	69
4	Visual Stability Predictor: Visual Acuity	70
4.1	Measuring Perceptual Stability	70
4.1.1	Predictions As Constraints: How Visually Rich Can A Design Be?	72

4.1.2	Correlating With Human Performance	72
4.2	Stability Measure Of Spatial Visual Acuity	73
4.2.1	Perceptual Measure	74
4.2.2	Perceptual Stability Algorithm: STRIVE	74
4.2.3	Measuring Visual Change With Shannon Entropy	76
4.2.4	Measuring Visual Change With Non-Normalised Shannon Entropy	81
4.2.5	Calculating Perceptual Stability: Visual Acuity	84
4.3	Conclusions	87
5	Experimental Evaluation Of Visual Acuity Predictor	88
5.1	Experiments: Visual Acuity Stability	88
5.2	Eye Charts As A Gold Standard Of Visual Function	89
5.3	ETDRS, Pelli-Robson & Campbell-Robson Eye Charts	90
5.3.1	ETDRS Eye Chart	91
5.3.2	Pelli-Robson Eye Chart	93
5.3.3	Campbell-Robson Contrast Sensitivity (CRCS) Eye Chart	94
5.4	Virtual Eyes	94
5.5	Experiments	95
5.5.1	Overall Approach & Details	95
5.5.2	Hypothesis	97
5.5.3	Experiment Design	98
5.5.4	Experiment 1: PER & LocPER of ETDRS Eye Chart	101
5.5.5	Experiment 2: PER & LocPER of Pelli-Robson Eye Chart	105
5.5.6	Experiment 3: Row Regions PER & LocPER of Campbell-Robson Eye Chart	110
5.5.7	Experiment 4: Grid Regions PER & LocPER of Campbell-Robson Eye Chart	115
5.6	Conclusion	120
6	Visual Stability Predictor: Colour Perception	121
6.1	Predicting Colour Stability	121
6.2	Measuring Colour Perception Visual Stability	122
6.2.1	Perceptual Measure	122
6.2.2	Perceptual Stability Algorithm: STRICODI	123
6.2.3	Calculating Perceptual Stability: Colour Perception	125
6.3	Conclusion	129
7	Experimental Evaluation of Colour Stability Predictor	130
7.1	Experiments: Colour Perception Stability	130
7.2	Ishihara Eye Charts	131
7.3	Experiments	132
7.3.1	Overall Approach & Details	133
7.3.2	Hypothesis	135
7.3.3	Experiment Design	135
7.3.4	Experiment 5: Protanopia	139
7.3.5	Experiment 6: Deuteranopia	148

7.4	Conclusion	149
8	Adapting Visual Designs To Individual Differences	150
8.1	Adaptions Demonstrated	150
8.2	Physiological Versus Semantic Adaptions	151
8.3	Visual Acuity Adaptions	154
8.3.1	Text & Font Style	155
8.3.2	Network Graphs	163
8.4	Colour Perception Adaptions	174
8.4.1	Maps	176
8.5	Conclusions	183
9	Conclusions & Future Work	184
9.1	When Individual Differences First Matter	184
9.2	Contributions	185
9.2.1	Simplified Models Of Physiological Eye Function For HCI	186
9.2.2	Degree And Location Of Perceptual Stability Due To Spatial Perception	187
9.2.3	Degree And Location Of Perceptual Stability Due To Colour Perception	187
9.2.4	Optometry Models Of Individual Eye Function Integ- rated With Perception	188
9.2.5	Demonstrations Of Predictions Used To Make Design Adaption Decisions	189
9.3	Future Directions	189
9.3.1	Related Possibilities	189
9.3.2	Building Upon Thesis Findings	190
9.4	Summary	193

LIST OF TABLES

1.1	Refinements of the question “What do people see in a design?”	13
3.1	Tristimulus values for standard illuminants and observers. Used in Equation 3.1 to account for the colour of light in the environment (Illuminant) and the stimulus size (visual angle) of the area the colour light waves are reflected or emitted from.	62
3.2	Zernike coefficients for Zernike modes in normal population of human observers with 6 mm pupils.	68
4.1	STRIVE Curve Interpretations	85
5.1	List of experiments arranged by eye charts and region divi- sions along with standard human visual acuity judgements. .	100
5.2	Results of analysing the ETDRS eye chart for LocPER.	102
5.3	Averaged result of analysing the Pelli-Robson eye chart for LocPER.	105
5.4	Breakdown of results of analysing the Pelli-Robson eye chart for LocPER, when chart divided into 2 by 2 regions.	107
5.5	Breakdown of results of analysing the Pelli-Robson eye chart for LocPER, when chart divided into 2 by 3 regions.	108
5.6	Effect of increases in Noise Threshold on $PERS_{va}$ values for Campbell-Robson eye chart divided into 1 by 4 regions. . . .	112
5.7	Averaged result of analysing the Campbell-Robson eye chart for LocPER when the eye chart is divided into rows.	112
5.8	Breakdown of results of analysing the Campbell-Robson eye chart for LocPER, when chart divided into 1 by 3 regions. . .	112
5.9	Breakdown of results of analysing the Campbell-Robson eye chart for LocPER, when chart divided into 1 by 4 regions. . .	114
5.10	Averaged result of analysing the Campbell-Robson eye chart for LocPER, when chart divided into grids.	116
5.11	Breakdown of results of analysing the Campbell-Robson eye chart for LocPER, when chart divided into 2 by 2 regions. . .	118
5.12	Breakdown of results of analysing the Campbell-Robson eye chart for LocPER, when chart divided into 3 by 3 regions. . .	119

7.1	How optotypes and lines on Ishihara eye charts are seen by protanopes and deuteranopes.	137
8.1	List of network graph design adaptations evaluated with $PERS_{va}$	163
8.2	$PERS_{cp}$ scores for each version of each deuteranope map. Location in table corresponds to locations of maps in Figure 8.30.	182
8.3	$PERS_{cp}$ scores for each version of each protanope map. Location in table corresponds to locations of maps in Figure 8.31.	182
9.1	Refinements of the question "What do people see in a design?"	185

LIST OF FIGURES

1.1	An example of how conference presentations are often difficult to see.	2
1.2	An example of colour used to create three groups of dots and a visual pop out effect.	3
1.3	Different stages light-waves undergo processing that alters how they are perceived.	5
1.4	An example of an audience aware self-adapting advert from Amnesty International.	9
2.1	Three examples of how people with different spatial visual acuities could see the same eye test chart.	25
2.2	Examples of how different forms of colour blindness effect perceived colours.	27
2.3	Example of how a single point of light can deviate as it traverses an optical system (lens).	29
2.4	Example of ideal and aberrated wavefronts generated by light rays traversing an optical system.	31
2.5	Pyramid of Zernike modes where each mode is an aberration.	32
3.1	Flowchart showing how the models of visual function are used to adapt visual designs.	39
3.2	An example of a stable and unstable percept.	42
3.3	Example of side by side feature dependent analysis of potential designs for Information Visualisations.	46
3.4	Examples of how colour blindness effects the perception of colour in a photograph.	48
3.5	Three simulated retinal image examples of how people with different visual acuities could see the same eye chart.	50
3.6	Example of the point spread of a single point light source as it travels through an optical system.	51
3.7	Example of three different Zernike Mode PSFs applied to a photograph of shoes on grass.	52
3.8	Steps in algorithm for simulating visual acuity to generate a retinal image of a visual design.	54

3.9	Examples of visual acuity simulated retinal images generated from a photograph of red shoes on green grass.	55
3.10	CIELAB 1976 L a* b* is a widely used and validated [32] perceptually uniform colour space. From [55].	58
3.11	Steps in algorithm for simulating colour perception to generate perceptually valid colour representation of a visual design.	65
4.1	Steps in STRIVE (SimulaTed Retinal Images with Visual Entropy) algorithm to generate measure of visual change.	75
4.2	Simulated retinal images of a vertical bar.	76
4.3	Examples of optical blur on combined visual acuity eye chart.	78
4.4	Two graphs showing how increases in RMS for optical aberration defocus Z_2^0 effect measures of entropy.	79
4.5	Increase in Non-Normalised Shannon Entropy of the simulated retinal images with increasing RMS values.	81
4.6	Increase in Non-Normalised Shannon Entropy of the simulated retinal images with increasing RMS values.	82
4.7	STRIVE curve slopes for combined eye chart.	84
4.8	$PERS_{va}$ perceptual stability score for Figure 4.3.	86
5.1	Example of ETDRS Eye Chart.	90
5.2	Example of Campbell-Robson Contrast Sensitivity Eye Chart.	91
5.3	Example of Pelli-Robson Contrast Sensitivity Eye Chart [66].	92
5.4	Orientation of Campbell-Robson eye chart during experiments.	96
5.5	Depiction showing how the ETDRS and Campbell-Robson eye charts were divided into 1 by 2, 1 by 3 and 1 by 4 regions.	102
5.6	Heatmap of ETDRS eye chart when divided into 1 by 2 regions. Red text shows normalised $PERS_{va}$ scores.	103
5.7	Heatmap of ETDRS eye chart when divided into 1 by 3 regions. Red text shows normalised $PERS_{va}$ scores.	103
5.8	Heatmap of ETDRS eye chart when divided into 1 by 4 regions. Red text shows normalised $PERS_{va}$ scores.	104
5.9	2 by 1, 2 by 2 and 2 by 3 regions the Pelli-Robson chart is divided into to test for LocPER contrast sensitivity.	104
5.10	Heatmap of Pelli-Robson eye chart when divided into 2 by 1 regions.	106
5.11	Heatmap of Pelli-Robson eye chart when divided into 2 by 2 regions. Red text shows normalised $PERS_{va}$ scores.	106
5.12	Heatmap of Pelli-Robson eye chart when divided into 2 by 3 regions. Red text shows normalised $PERS_{va}$ scores.	107
5.13	Results of first four lines of Pelli-Robson contrast chart, when divided into 2 by 4 regions to test effect of optotype shape without contrast.	110
5.14	Heatmap of Campbell-Robson eye chart when divided into 1 by 2 regions. Red text shows normalised $PERS_{va}$ scores.	113
5.15	Heatmap of Campbell-Robson eye chart when divided into 1 by 3 regions. Red text shows normalised $PERS_{va}$ scores.	113

5.16	Heatmap of Campbell-Robson eye chart when divided into 1 by 4 regions. Red text shows normalised $PERS_{va}$ scores. . . .	114
5.17	3 by 3 and 4 by 4 regions the Campbell-Robson chart is divided into to test for the CSF curve.	115
5.18	Heatmap of Campbell-Robson eye chart when divided into 2 by 2 regions. Red text shows normalised $PERS_{va}$ scores. . .	117
5.19	Heatmap of Campbell-Robson eye chart when divided into 3 by 3 regions. Red text shows normalised $PERS_{va}$ scores. . .	117
5.20	Heatmap of Campbell-Robson eye chart when divided into 4 by 4 regions. Red text shows normalised $PERS_{va}$ scores. . .	118
6.1	Steps in STRICODI (SimulaTed Retinal Images with Colour Differences) algorithm to generate measure of change.	124
6.2	Example of STRICODI with $PERS_{cp}$ thresholding an image.	126
6.3	Heatmap of colour perceptual stability of photo of red shoes on green-yellow grass.	128
7.1	Examples of two Ishihara colour plates, when viewed with normal vision and protanopia.	132
7.2	Ishihara 24 colour plates. x indicates start and end points of lines that observers may or may not trace depending on their visual function.	134
7.3	Ishihara 24 colour plates as seen with protanopia.	140
7.4	Ishihara 24 colour plates with perceptually unstable regions for protanopia highlighted.	141
7.5	Heatmaps showing perceptual stability of Ishihara 24 colour plates when viewed by protanopes.	142
7.6	Ishihara 24 colour plates as seen with deuteranopia.	145
7.7	Ishihara 24 colour plates with perceptually unstable regions for deuteranopia highlighted.	146
7.8	Heatmaps showing perceptual stability of Ishihara 24 colour plates when viewed by deuteranopes.	147
8.1	Example of how the $PERS_{va}$ score is a physiological rather than semantic metric.	152
8.2	Effect of Z_2^0 optical aberration on vertical bars set at different distances from each other.	153
8.3	Image showing the English alphabet lower-case letters that are compared with $PERS_{va}$	156
8.4	Mean and standard deviation of $PERS_{va}$ scores for Courier-Bold versus Times-Roman alphabet text rows.	157
8.5	Scatter plot distribution of $PERS_{va}$ for each font size and font type, when analysing alphabet text rows.	157
8.6	Distribution of $PERS_{va}$ score for each virtual eye for each font style and font size, when analysing alphabet text.	158
8.7	Image showing the paragraphs of text that were compared with $PERS_{va}$	160

8.8	Mean and standard deviation of $PERS_{va}$ scores for Courier-Bold versus Times-Roman text paragraphs.	161
8.9	Scatter plot distribution of $PERS_{va}$ for each font size and font type, when analysing text paragraphs.	161
8.10	Distribution of $PERS_{va}$ score for each virtual eye for each font style and font size, when analysing paragraph text. . . .	162
8.11	Graph Set 1 - Node size increases.	164
8.12	Graph Set 2 - Edge width increases.	164
8.13	Graph Set 3 - Hollow node size increases.	164
8.14	Graph Set 4 - Hollow node border width increases.	164
8.15	Graph Set 1 - $PERS_{va}$ mean and standard deviation.	166
8.16	Graph Set 1 - Distribution of $PERS_{va}$ for each virtual eye for each region.	166
8.17	Graph Set 2 - $PERS_{va}$ mean and standard deviation.	167
8.18	Graph Set 2 - Distribution of $PERS_{va}$ for each virtual eye for each region.	167
8.19	Graph Set 3 - $PERS_{va}$ mean and standard deviation.	168
8.20	Graph Set 3 - Distribution of $PERS_{va}$ for each virtual eye for each region.	168
8.21	Graph Set 4 - $PERS_{va}$ mean and standard deviation.	169
8.22	Graph Set 4 - Distribution of $PERS_{va}$ for each virtual eye for each region.	169
8.23	Four network graphs that vary along multiple visual dimensions.	172
8.24	Graph Set 5 - $PERS_{va}$ mean and standard deviation.	173
8.25	Graph Set 5 - Distribution of $PERS_{va}$ for each virtual eye for each region.	173
8.26	Perceptual graph layout - Example of a seven node seven edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.	175
8.27	Pseudocoloured satellite maps showing NOAA/NASA AVHRR SST for November 1999.	177
8.28	Deuteranope versions of maps.	178
8.29	Protanope versions of maps.	179
8.30	Thresholded perceptual change map for deuteranope versions of maps.	180
8.31	Thresholded perceptual change map for protanope versions of maps.	181
9.1	Perceptual graph layout - Example of a six node three edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.	194
9.2	Perceptual graph layout - Example of a five node five edge circular graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.	195

9.3	Perceptual graph layout - Example of a seven node seven edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.	196
-----	---	-----

ABSTRACT

When a user interface, information visualisation or graphic designer is conceiving and creating design mockups how does the designer know whether the intended audience is able to perceive the design? When a designer does know how well an intended audience can or cannot see, such as with a design targeted at an aging audience, how does that knowledge influence the visual layout of the design?

There are rules of thumb about font size, contrast, and the interaction between unused space that are learnt and handed down as design lore. If a designer follows good use of font size, with good contrast then a proposed design should be readable as long as its not too cluttered. Unfortunately “good usage”, “good contrast” and “not too cluttered” are subjective measures. What one designer defines as good another could find distinctly lacking, though experience and training do help a designer acquire knowledge of what visually works.

This thesis is concerned with examining and showing how the experience of seeing a design can be non-subjectively quantified. Then it demonstrates how the quantifications tied together with individual differences in the Human Visual System (HVS) can be used to evaluate and adapt the designs, such that they are customised to individual eye sight.

In order to non-subjectively quantify the experience of seeing a design we introduce, evaluate and demonstrate two measures of perceptual stability. Perceptual stability is defined by us as a measure of how stable or unstable a visual design or image is due to differences in a perceiver’s perception. The first measure $PERS_{va}$ evaluates how easy or difficult it is for people to see visual detail in a design. While the second measure $PERS_{cp}$ evaluates how different forms of colour perception effect the legibility of a visual design.

Objective quantifications which are capable of modelling individual differences are useful for automating design judgements, i.e. automatically compare a range of potential interface designs and make a decision about which is best for a specific user. Demonstrated in this work are automatic evaluations of text and font styles, network graph designs and layouts, and the pseudocolouring of scientific visualisations. In the longer term, as we move into a world where Mass Customisation and Product Personalisation become common place, objective design quantifications are useful for adapting and customising designs to suit individual physiologies, capabilities and preferences.

Dedicated to Rois, Michael and my parents.

ACKNOWLEDGEMENTS

Big thanks to Dr. Aaron Quigley for enabling me to pursue challenging research at the intersection of disciplines. Thanks to Dr. Fred Cummins, and also to everyone in the Systems Research Group in University College Dublin, Ireland for stimulating discussions and advice.

Thanks to UCD for the scholarship and the seed funding award.

Over the last few years many individuals have provided support, friendship and inspiration. Of those I am particularly indebted to Dr. Sile O'Modhrain, Dr. Carol Strohecker, Jonah Brucker-Cohen and others in the ex-MIT Media Lab Europe community, whose innovative projects, design thinking and diversity of views thought me the depth, challenges, and fun of Human-Computer Interaction and Interaction Design.

...most of all - thanks Rois (and parents and in-laws), I couldn't be luckier...and to baby Michael, your slap-the-keyboard one letter addition to this thesis may make you the youngest thesis contributor ever!

LIST OF PUBLICATIONS

[11] (*Awarded Best Poster*) Mike Bennett and Aaron Quigley. Understanding Distance & How Humans See Interfaces & Designs In *VGV Irish Graduate Student Symposium on Vision, Graphics and Visualisation*, Dublin, Ireland, June 2008.

[10] Mike Bennett and Aaron Quigley. Perceptual usability: Predicting changes in interfaces & designs due to visual acuity differences. In *AVI 2008 Conference on Advanced Visual Interfaces*, Napoli, Italy, May 2008.

[9] Mike Bennett and Aaron Quigley. A method for the automatic analysis of colour category pixel shifts during dichromatic vision. In *2nd International Symposium on Visual Computing (ISVC)*, volume 2 of *Lecture Notes in Computer Science*, pages 457–466, Lake Tahoe Nevada, USA, Nov 6-8 2006. Springer Verlag.

[8] Mike Bennett and Aaron Quigley. Automatically evaluating the impact of colour blindness on information visualisations. In *CASCON 2006 Dublin Symposium*, Dublin, Ireland, October 2006. IBM Centre for Advanced Studies.

Introduction

1.1 Motivation & Purpose

When a user interface, information visualisation, or graphic designer is conceiving and creating design mockups how does the designer know whether the intended audience is able to perceive the design? When a designer does know how well an intended audience can or cannot see, such as with a design targeted at an aging audience, how does that knowledge influence the visual layout of the design?

For example, often conference presentations are difficult to see and read, due to an interplay of the visual design of the presentation, the size of the presentation display, the viewer's position and how good the viewer's visual acuity is (see Figure 1.1).

There are rules of thumb about font size, contrast, and the interaction between unused space that are learnt and handed down as design lore. If a designer follows good use of font size, with good contrast then a proposed design should be readable as long as its not too cluttered. Unfortunately "good usage", "good contrast" and "not too cluttered" are subjective measures. What one designer defines as good another could find distinctly lacking, though experience and



Figure 1.1: An example of how conference presentations are often difficult to see, due to the visual design of the presentation and the viewer's position.

training do help a designer acquire knowledge of what visually works.

Focusing in on new media graphic design, specifically the areas around the visual design of information visualisations and user interfaces, we face the same questions. How do we know what people can see when they look at a complex information visualisation, user interface or any kind of computationally driven design? How do we know what impacts differences in how people can see has upon the experience of seeing visual designs?

These questions are made more complex because the graphical designs are not static pictures printed with ink on paper. These dynamic visual designs change in response to many factors, including changes in the underlying data used to generate the visual designs, such as a graph or treemap generated by changing stock prices.

This thesis is concerned with examining and showing how the experience of seeing a design can be non-subjectively quantified. Then it demonstrates how the quantifications tied together with individual differences in the Human

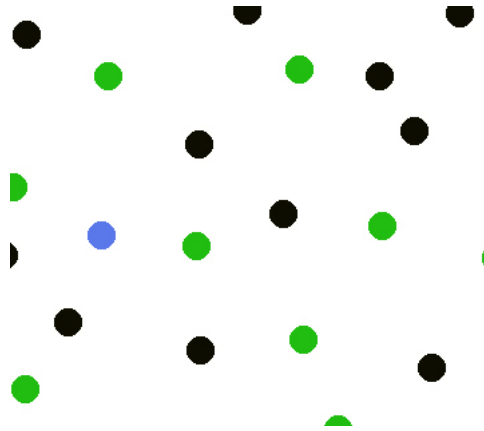


Figure 1.2: An example of colour used to create three groups of dots (green, black, blue). Blue is also used to create a visual pop out effect.

Visual System (HVS) can be used to evaluate and adapt the designs - thereby improving human performance.

The experience of seeing is multifaceted with multiple interactions between sensation, perception and awareness [63]. To make the question approachable it is limited to two facets of seeing. Firstly, the effect of colour perception on graphical designs, and secondly, the effect of visual acuity (spatial perception) on our ability to see a design.

Colour is widely used in graphical designs [90], often to denote categorical distinctions and groupings within data, e.g. green signifies a node representing a male, black signifies a node representing a female (see Figure 1.2). Colour is also a fundamental building block of more complex visual components, such as textures. By focusing on colour perception we are examining the effects of a very important low level unit of composition in graphical interfaces and visual designs.

Visual acuity dictates our ability to see detail. *Spatial visual acuity is the smallest spatial detail that can be visually detected, discriminated, or identified* [61]. If objects are too small or too cluttered we cannot properly see them in a design. Often when designing a visualisation or interface there is a trade off between the amount of information we have to communicate versus the amount of display

real-estate we have to communicate with. By focusing on visual acuity we are examining the interplay of the size of graphical objects used in designs and our experience of seeing them, also known as spatial perception.

For example if we have a massive network graph consisting of thousands of nodes and edges it may be impossible to display that graph in a meaningful manner on a small screen. This can be impossible because all the nodes and edges would form a visually indistinct cluster on-screen with nodes and edges layered on top of each other and obscuring each other.

In the rest of this chapter the motivation for the research is established, while also tying together the various fields of knowledge and perspectives that informed our approach to quantifying the experience of seeing and then using the quantifications to adapt designs to individual eye function. The core contributions of and motivations for this research are:

- modelling the human visual system (spatial perception and colour perception) (see Chapter 3),
- using the models to make feature independent predictions about the experiences of seeing designs (see Chapter 4 and Chapter 6),
- then using the predictions to adapt graphical designs to suit individual eye function (see Chapter 8).

The techniques introduced and evaluated in this thesis are useful for automating design judgements, i.e. automatically compare a range of potential interface designs and make a decision about which is best for a specific user. Demonstrated in this work are automatic evaluations of text and font styles, network graph designs and layouts, and the pseudocolouring of scientific visualisations. In the longer term, as we move into a world where Mass Customisation and Product Personalisation become common place, objective

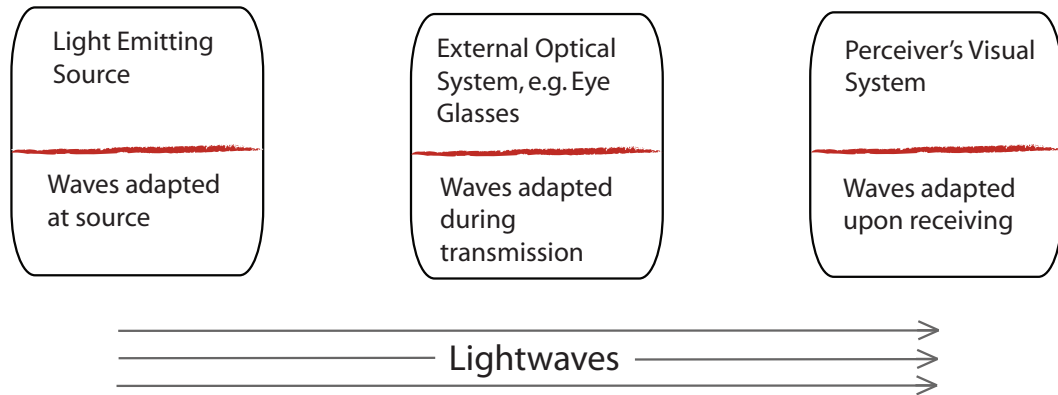


Figure 1.3: Different stages light-waves undergo processing that alters how they are perceived.

design quantifications are useful for adapting and customising designs to suit individual physiologies, capabilities and preferences.

1.2 Mass Customisation & Personalisation

By focusing on the role of the perceiver's visual system we are faced with the fact that visual systems differ between individuals. Individual differences in eye function have been well established for a considerable period of time going back centuries.

1.2.1 Adapting At Source

We are all familiar with corrective eye glasses that bring a person's eye sight within normal ranges. Eye glasses do not fix flawed eye sight, though advanced eye surgery can now often adjust the physiology of the eye to correct for flaws, e.g. LASIK eye surgery [48]. See Chapter 3 and Chapter 2 for more details on individual eye function and related physiological structures.

Eye glasses can be thought of as simple information processing systems that carry out transformative functions on the paths light-waves take. Eye glasses

carry out information processing while sitting between a light emitting or reflecting source and the perceiver's visual system (see Figure 1.3, External Optical System). Eye surgery adjusts the perceiver's visual system and by doing so adjusts the information processing of light-waves in the eye (see Figure 1.3, Perceiver's Visual System).

With computational driven designs and displays we have the opportunity to adapt light emitting or reflecting surfaces to individual eye function. That is, we have the opportunity to correct for limits, differences and flaws in human vision at the point light-waves are emitted (see Figure 1.3, Light Emitting Source).

As with eye glasses we also have the opportunity to adapt the emission and reflection of light-waves to individual eye function when we have measures of individual eye function. Such measures are often carried out by optometrists using various standardised eye charts and tests. See Chapter 4 for more details.

1.2.2 What Adapts

A question we can then ask is what should adapt to users? Since this work is focused on visual perception that which adapts has to be visible, whether this is through it emitting or reflecting light-waves.

In the immediate term we know that we can adapt what appears on existing displays, particularly displays involving graphical designs. To constrain the broadness of the research question we limit the application of the research to graphical designs within the domain of Human-Computer Interaction, predominately to Information Visualisations.

In the longer term as emerging research in technology, materials, robotics and nanotechnology indicate we can expect a future where the physical world can become real-time adaptable. That is, light sources and reflective surfaces won't

only be traditional computer displays, whether large wall sized or small hand-held displays. Exactly how and whether these technologies will work raises many wide open research questions [69]. For instance there are now conferences focused on Mass Customisation and Product Personalisation [46]. There are also research tools enabling researchers to experiment with algorithms to control programmable matter composed of virtual miniature robotic swarms [98, 49] and catoms [67]. Sample experiments include experimental techniques for enabling the swarms to self-organise into virtual physical objects such as cups and saucers.

As people become more technologically literate we can expect them to try to do more with technology (technology as an enabler), e.g. many people create websites with tools that simplify the problem of creating websites. As more people begin to design complex interactive interfaces and information visualisations [26, 86] we need to help them design the "best" visualisations possible. The techniques developed during the course of this research can be used as part of automated techniques for aiding designers in adapting and improving their visual Human-Computer Interaction and Information Visualisation design work.

1.3 Application & Design Scenarios

By having predictors of the effects of spatial perception and colour perception we can begin thinking about and creating the following new applications and interaction innovations.

1.3.1 Self-Adapting Signage

Possible design scenarios include signage that automatically adapts what is shown on the display depending on the location of the signage. For example place a self-adapting sign in a hospital corridor, then input average length of corridor and input average age profile of the people who would be looking at the sign. Then send the sign a design to display and let the sign determine the appropriate size and contrast to display the design at.

The design sent to the sign could be loosely defined so the self-adapting sign could determine the optimal layout for the design. Rather than explicitly stating that the text font needs to be 50px high and a picture needs to be 1000 by 500px the design could have guidelines such as "Text message A is most important and must be shown, picture B is reasonably important and if shown needs to maintain a size relationship of 1:0.75 to the size of text A".

1.3.2 Visual Popout Paint For Car Safety

A more futuristic design scenario is where the paint on cars automatically adapts to improve drivers ability to see the cars. Imagine cars coordinate an ad-hoc dynamic self-organising network where they swap information about drivers visual acuity and colour perception. As a driver moves behind a different car the car in front has a model of the following driver's visual acuity and colour perception. Then the leading car has its paint job shift colour and texture to make it easier for the following driver to spot and make judgements about the distance between the cars.

These self-repainting cars could also have other smart algorithms where when they pass a school for young children they further adapt their paint jobs so the cars visually pop out for the school children. Ideally improving road and traffic safety around the school.

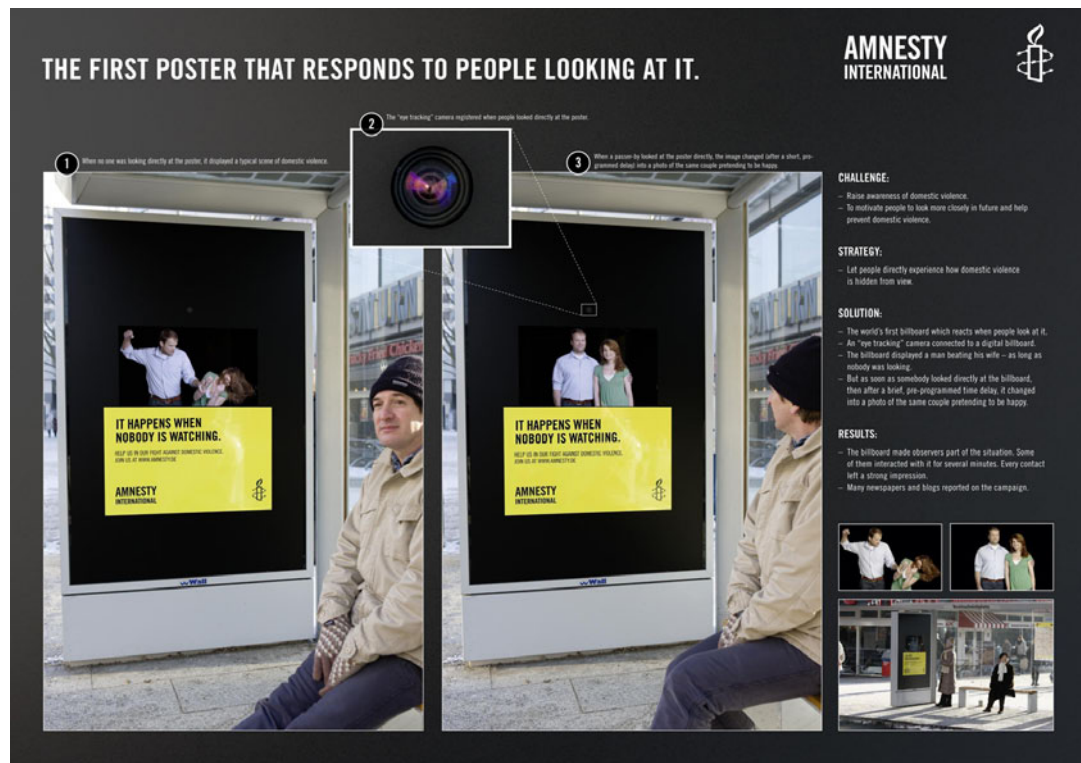


Figure 1.4: An example of an audience aware self-adapting advert from Amnesty International. The couple in the advert appear happy when directly looked at, but appear to engage in violence when not directly looked at.

1.3.3 Audience Aware Presentations

More mundane but a potentially very useful immediate application of the research is mentioned in Section 1.4.3. Tools that help people improve the legibility of slideshows and conference presentations. We have all experienced presentations that are difficult to see and read. Presentations could be measured and adapted to suit average attendee age, display size and expected range of viewer distances.

Shown in Figure 1.4 is an example of an audience aware self-adapting advert from Amnesty International. When the advert is directly looked at the couple in the advert appear happy, but when not directly looked at the couple appear to engage in violence.

1.4 Approach: Model, Predict & Adapt

Succinctly put the approach taken is to model, predict and adapt. To achieve our aim of non-subjectively evaluating an interface or visualisation we require metrics capable of making predictions about how easy or difficult designs are to see.

These metrics must consistently agree with human judgements. That is, they should be capable of saying which of two designs are easier to see, or because of individual differences in colour perception the metrics should tell us to what degree colour confusions could arise. For example, this can indicate whether a colourful interface appears perceptual different between perceivers who have some form of colour blindness.

1.4.1 Model

In order to develop the metrics we model the perception of colour and the perception of visual detail. The HVS is a highly complex structure about which many ambiguities remain; so in creating our models we have to decide what is feasible to model (see Section 2.3). To this end we create models focused on low-level vision. Low-level vision occurs in the early stage of the HVS, predominately due to the physiological structure of the eye. What we model and how we model is covered in-depth in Chapter 3.

1.4.2 Predict

Once we have models we can use them to make predictions about visual interfaces and designs. We look at two kinds of predictions. The first kind of prediction is a measurement of the degree of change in a design due to changes in spatial acuity perception and colour perception. We are referring to this

measure as perceptual stability and defining it as *a measure of how stable or unstable a visual design or image is due to differences in a perceiver's (spatial or colour) perception.*

The perceptual stability measure enables us to do a number of interrelated things. Firstly, compare the perceptual stability of a range of potential designs and make a decision about which is best, i.e. most perceptually stable. Secondly, predict the impact of changes in perceivers eye function on how easy or difficult something is to see. An example of this is where normal age related differences in visual acuity, or the effect of variations in viewer distance on visual acuity, lead to differences in the experience of seeing a design. The further away the viewer is the less visual detail seen, and the older the viewer is typically less visual detail is seen.

Building on our first predictor, perceptual stability, we then develop a second metric that is designed to identify the location of changes. This helps us understand what parts of a design are most or least perceptually stable. By knowing this a designer or automated graphical layout technique can decide to put lower priority information in locations that may be more difficult to see. Another benefit is that by testing the perceptual stability of each sub-part of a design we can ensure that all parts of the design meet certain minimum perceptual stabilities. Taking a single measure of perceptual stability for a whole design is informative but it may miss local minimas and maximas.

For the purposes of this thesis the second predictor is referred to as *Location of Perceptual Stability (LocPERS)*, and the first predictor is referred to as *Perceptual Stability (PERS)*. The predictors and how they function is elaborated upon in Chapter 4 and Chapter 6.

1.4.3 Adapt

Working predictors can be used in automated graphical design techniques to generate improved designs. These adaptations can be carried out to cater for individual differences in eye function, or different contexts of use for information visualisations and interfaces.

For example one could imagine the adaption techniques built into a slide show presentation tool, such as PowerPoint or KeyNote. The adaptations that occur can be based on the audience's average age, from which we know the expected normal age-related visual acuity. By knowing normal visual acuity we can use the predictors to test variations of a set of slides and then make graphical alteration suggestions based on how the predictors rate the slides.

Within this work we demonstrate evaluations and adaptations of text and fonts styles (see Section 8.3.1), network graphs (see Section 8.3.2), and pseudocolouring of scientific visualisations (see Section 8.4.1). Refer to Chapter 8 for more details on the adaptations.

1.5 Hypothesis

The core hypothesis of this thesis is that models of human vision integrated with image quality predictors are able to simulate and agree with human judgements about the legibility of visual designs. Legibility is a measure of how easy or hard a design is to see.

In order to test the hypothesis a number of assumptions and constraints are made. Listed in Table 1.1 in the Refinements column are the top level constraints, while the nature of and basis for the constraints are discussed in Chapter 2 and Chapter 3.

The following top level assumptions are made:

Table 1.1: Refinements of the question "What do people see in a design?"

Facet	Refinement
What	Computationally driven visual designs.
People	Individual differences in low-level vision.
See	Spatial Perception (primarily spatial visual acuity), Colour Perception.
Design	HCI focused on Information Visualisations.

- the current state of the art techniques from vision science and colour theory are sufficiently developed for application to HCI research problems,
- we do not require semantic knowledge about the analysed images (see Section 3.4.2), i.e. we do not have to recognise and compare shapes or features in an analysed image to develop techniques that agree with human judgements of legibility,
- point / pixel based measures of image quality are effective as a basis for techniques to measure a visual design's legibility (see Section 3.6.6 and Section 3.4).

1.6 Contributions

In the previous sections our motivating question is "What do people see in a design?". Further refinements are outlined in Table 1.1. The core contributions of this thesis are techniques for modelling and evaluating the effects of spatial and colour perception on computationally driven designs, along with demonstrations of how the predictors can be used to adapt designs to suit individual eye function.

1.6.1 Simplified Models Of Physiological Eye Function For HCI

The first core contribution of the work are simplified models of eye function for HCI. Rather than attempting to develop or utilise research-in-progress models with high biologically fidelity [27] we constructed and tested simplified mathematical models. Unlike research in Vision Science here we focus in on the effects of seeing designs, rather than the process of seeing itself.

1.6.2 Degree And Location Of Perceptual Stability Due To Spatial Perception

In Section 1.4.2 we introduced our definition of Perceptual Stability, and in Section 2.1 we refine what is meant by stable and unstable. Another core contribution of this work is the development of metrics based on the criteria of the Perceptual Stability definition.

For Spatial Perception we developed metrics for measuring Perceptual Stability. These metrics consist of a measure of the degree of change a design perceptually undergoes when subject to varying amounts of normal optical aberrations.

1.6.3 Degree And Location Of Perceptual Stability Due To Colour Perception

Similarly with colour perception we create metrics that can measure the degree of perceptual change a design undergoes when viewed by an individual with varying colour perception. A related contribution is a technique for helping identify the locations in a design that are subject to perceptual change.

1.6.4 Optometry Models Of Individual Eye Function Integrated With Perception

When creating the models and predictors of eye function we design the models such that they can be adapted to model differences in physiological function. The approaches taken are informed by knowledge of optometry and eye function, specifically wavefront models for modelling optical aberrations in the human eye.

An unexpected optometry contribution from this work is presented in Section 5.5.5.1, where we experimental show how $PERS_{va}$ may have found a bias in the Pelli-Robson eye chart. Within optometry a considerable amount of research is focused on detecting and removing biases in eye charts.

1.6.5 Demonstrated Predictors Used To Adapt Information Visualisations & Interfaces To Individual Eye Function

Further we demonstrate how the models and predictors we constructed can be applied to improving the designs of interfaces and information visualisations. Finally the predictions are shown to be useful for aiding the automatic adaption of designs, so the designs are improved for spatial and colour perception.

1.7 Summary

In the following Chapters the question "What can people see in a design?" is examined in multiple related steps. Firstly, we develop models of eye function for analysing visual designs. The models are then refined to cater for differences in individual eye function. Then the models are integrated with predictors to provide feedback on the perceptual stability of designs. After which

the predictors are validated by testing them on a range of eye charts, and are further examined by testing them on a range of information visualisations and font styles. Then the predictors are used to adapt designs to individual eye function, and the resulting adaptations are presented.

Background

This chapter provides background discussion on techniques, research questions and approaches that are relevant to this thesis. We motivate the work by discussing the relevancy of physiology for design, then relate that to an individual differences approach to Human-Computer Interaction.

Following this is background material providing a brief introduction to modelling and simulating facets of human vision, spatial perception and colour perception, optical aberrations, wavefronts and using Zernike Polynomials to model early stage human vision.

2.1 Physiology As A Design Context & Constraint

In the Chapter 1 it is observed that the dynamic nature of computationally driven visual designs can make designing more challenging, but it also presents a significant opportunity. We can adapt the designs to context.

One context we are already familiar with is the technical capabilities of displays, which place limits such as the maximum display resolution and presentation size on our designs. Other less well defined contexts include the environment the display is placed in, ambient lighting conditions and perceiver's

perceptual capabilities.

By better understanding and defining the contexts we can better understand the potential design space from which we can create designs. Designers already implicitly take contextual information into consideration when designing. For example they do not design visualisations that use infra-red as a colour. Not because infrared cannot be used to encode meaning but because unaided human vision does not convert the infrared wavelengths of light into a colour percept.

The central question of "What do people see in a design?" has various contexts. Environmental factors and physical display contexts are essential but of particular interest to us is the human context. What impact does human perception have on the effectiveness and experience of seeing visual designs and patterns? This is a very broad question that has been indirectly and directly empirically examined numerous ways over the centuries; ranging from specific work on issues such as representational determinism [101, 100] to broader studies of visual function [61] and vision science [63, 16].

There are many possible ways of looking at the human context in perception of designs, for example taking a high level approach and starting at conscious visual awareness. Rather than assuming that one aspect of the human visual system is more important than another we ask the question: At what point does the human perception become a relevant context? Or putting it another way: When and at what point does the act of seeing first potentially alter a design?

It is easy to think of a design as a finished thing, created by a designer and broadcast to an audience. This is overly simplistic and implicitly uses the physical world and original designer as the frame of reference for a design, i.e. the units of measurement used to dictate the size of the fonts, the range of light-waves reflected by different surfaces of the design, the limits and capab-

ilities of the original designer's visual system. A design has at least two frames of reference. The second frame of reference is the perceiver and the biases they introduce to the design. Bias can arise from many phenomena including but not limited to cultural background, cognitive structures, relevant education and prior expectations.

By asking at what point does the act of seeing first potentially alter a design we are asking about biases introduced in perception due to the physiological function of the human eye. How the eye transduces light-waves from stimulus into sensation introduces biases. The physiological structure of the eye determines eye function which leads to biases in perceptions, i.e. variations on the norm of eye function can lead to differences in interpretation due to differences in what is or is not seen in a design.

For example human eyes have peak sensitivities to different ranges of light-waves. Light-waves that fall outside these ranges are ignored, therefore they never become percepts. Furthermore within expected ranges different human eyes have somewhat varying peak sensitivities. These individual variations lead to differences in perceptions. A prime illustration of this is colour blindness, which is most often caused by signification variations in light-wave photoreceptor function and reduced photoreceptor sensitivities. Since specific ranges of light-waves become percepts those light-waves which cannot be transduced cannot be used when creating designs.

Expanding on the introduction of perceptual stability in Section 1.4.2 we can now refine what we mean by stable and unstable. A stable percept is one which is subject to the least amount of perceived change in a design - due to how the physiological structure and function of the eye and optical system biases the perception. An unstable percept is one where the design is significantly affected by how the low-level physiology and function of the eye biases the perceptions.

2.2 Human-Computer Interaction & Interaction Design

Human-Computer Interaction (HCI), or Interaction Design for those coming from a predominately design background, is the art and science of designing technological artifacts for enhancing people's abilities and enriching their lives, both as individuals and as social creatures. HCI is a very diverse field requiring experts from Art & Design, Cognitive & Vision Science, Materials Science, Psychology, Physiology, Ethnography, Social Science, Artificial Intelligence, Computer Science, Engineering, Mathematics and many other domains. This diversity of knowledge and perspectives makes it simultaneously challenging and deeply interesting.

There are many approaches to HCI [21, 68, 31] ranging from less empirical design and art inspired creators to experimentalists who seek a science of HCI. Whether the field of HCI has a scientific basis and what value there is in giving it one remains an open question. Is it possible to develop a "science" of HCI / Interaction Design? Or is our scientific understanding of humans and human groups still too under-developed?

The reality is that there is no choice about waiting for a science of HCI to fully emerge and be developed. Society currently demands artifacts that enhance human abilities - the need is there, it must be filled, and later on maybe understood and refined. From this perspective aspects of designing may be viewed as a heuristic first approximation of a science of HCI. What design does, how it does it and what its limits are can inform us about design approaches that are worth researching to establish whether they can be given a useful empirical grounding.

Measurements and models of eye function give us empirical stepping stones

to quantifying the impact of seeing on designs. For that reason in order to answer the central question "What do people see in a design?" we sought to acquire and apply knowledge about human visual function from optometry [61], ophthalmology [48, 52], optics [40] and vision science [63, 57].

2.2.1 Information Visualisation

Information Visualisation is concerned with the effective representation of data in a visual form [20, 85]. Effective means the representation aids problem solving and provides insights. In theory there are an infinite set of mappings between data and what visual representation could be created, though a good representation should not mislead.

A large body of information visualisation research exists that is concerned with innovative graphical representations. Novel visual representations, or significant variations on existing visualisations, are a staple source for conference and journal publications. As research in the area has progressed there has been a growing awareness and appreciation [22] for the need to understand what does and does not make one visualisation more effective than another. Cleveland's [24] work on statistical graphs is one of the earlier examples of examining what makes an information visualisation perceptually effective or not. More recently Colin Ware has brought together and synopsized a range of work from perceptually psychology and vision science that informs readers about perception for the design of information visualisations [90].

Within vision science there is a pre-existing and growing body of work on how the human visual system functions. By acquiring and using an understanding of the vision science literature and related domains we may be able to place aspects of Information Visualisation and HCI on a beneficial empirical basis.

By focusing in on Information Visualisations we can further refine what we

are asking in this thesis. When we ask the question “What do people see in a design?” we are specifically asking it about Information Visualisations displayed on computationally driven displays.

2.2.2 Individual Differences Applied To HCI

Within HCI and Information Visualisation there has been an intermittently ongoing debate about the benefit or negative aspects of automatic interface adaptations, interface consistency and catering to individual differences. In this work we have taken it as a given that adaption to individual differences in physiological function is an advantageous requirement for interfaces and visualisations. By completely ignoring individual differences we would be overly simplifying the problem of creating effective visual designs.

Elsewhere another danger in ignoring individual differences has been recognised [20]. As information visualisations become more prevalent then more people may be regularly and consistently mistakenly misled due to perceptual differences. Consistently misleading people would mean that differences between perceivers potentially turn from an inconvenience into a disability.

In 1996 Dillion and Watson [28] observed parallels between the development of two branches of psychology going back 100+ years and the development of HCI as a field. The two branches of psychology they wrote about were experimental psychology and differential psychology. Differential psychology is concerned with the importance of differences between people, and does not ignore outliers as anomalies or errors in experiments. Experimental psychology is focused on the median.

They observed that HCI has tended to only focus on one branch, that of experimental psychology. Individual differences and differential psychology was not receiving the attention it deserved, especially where approximately 25%

variation in task performance could be accounted for due to individual differences [28].

For this work moving away from a one sized fits all approach and accounting for individual differences increases the complexity of knowing what users see in designs. Any predictor of what users can see must be capable of modelling different ways of seeing.

Modelling different ways of seeing "at the point at which seeing first potentially alters a design" (see Section 2.1) implies that individual physiological differences in eye function need to be taken into consideration when building our models and predictors. Fortunately, as initially discussed in Section 1.2 and covered in more depth in Chapter 3, Chapter 5 and Chapter 7, the methodology for measuring and predicting the effects of individual differences in eye function has been extensively studied in optometry and ophthalmology research [4, 34, 29, 66, 92, 36, 37, 70, 61].

2.3 Individual Differences In Eye Function

People's visual system vary in capabilities. In some cases these differences are due to normal age related changes, such as decreasing visual acuity because of lens within the eye becoming less flexible with age [48, 61]. Other times these differences are brought on because of medical issues like diabetic retinopathy, or age related macular degeneration [61, 63]. Variations may also be due to genetic differences, for example colour blindness is most often due to genetic mutations carried by X chromosomes in females [59].

Individual differences are not necessarily negative, for example increased cone density and photoreceptor distribution may lead to better spatial visual acuity [25]. There are specific norms of performance that we can expect people to have but these norms are guidelines. For example standard observer human

visual acuity is commonly rated as 20/20, which is the ability to accurately read an eye chart with standardised size fonts and layout at 20 feet. It is not unusual for people to have vision rated above 20/20, i.e. 25/20 where they can accurately read the standardised eye chart from 25 feet rather than the normal 20 feet.

There are numerous other ways individual differences in eye structure can lead to performance differences in people's ability to see, e.g. differences in flicker sensitivities, differences in rates of light and dark adaptation.

Since seeing is a complicated process the question "What do people see in a design?" is made more challenging when thinking about individual differences. For example when looking at a dynamic visualisation rapidly changing in time how do we know the minimum time a block of colour must be displayed for our eyes to pick up that colour before we can change it to a separate colour? Answering that question fully implies accounting for the effects and interplay of stimulus size, perceptual distance between colours, intensity of light, allowing for effects of nearby colours, super-acuities, and other factors.

In order to make the question "What do people see in a design?" approachable we have to place limits on what aspect of seeing we ask the question about. Also desirable is that the parts of seeing we focus on are those that have large impacts upon the perception of designs. For instance the effect of individual differences in pupil size and pupil accommodation could have been examined but we would not necessarily expect these differences to have significant impacts on how people perceive designs.

As is briefly mentioned in Section 1.4.1 we focus on the effects of colour perception and the effects of spatial perception on the perception of information visualisations. Differences in colour perception are worth exploring because colour is a basic building block of graphical designs, often denoting separation and grouping (see Figure 1.2). Spatial perception is also worth exploring



Figure 2.1: Three examples of how people with different spatial visual acuities could see the same eye test chart.

because the ability to see detail is a basic building block in graphical designs. Previously others have examined aspects of colour perception on information visualisation and user interfaces [58, 87, 89, 47, 42, 90], but the effect of visual acuity has not received a great deal of attention [1].

Focusing on both spatial perception and colour perception enables us to seek more general models and predictors of the effects of seeing on designs. Finding commonalities between these two different facets of seeing could provide a framework and insights into accounting for the impact of other aspects of seeing upon designs.

2.3.1 Differences In Spatial Perception

For spatial perception we are concentrating on the effects of differences in people's ability to see detail in designs. Specifically on the effects of defocus on the designs (see Figure 2.1). Defocus can be introduced by the following interrelated variables:

- Person's spatial visual acuity and spatial vision
- Distance person is perceiving stimuli from

Spatial perception is the term used to describe our ability to see details clearly, consistently and accurately. Spatial perception differs between individuals

depending on the interplay of a range of factors (see Figure 2.1). Contributing factors include the types and degrees of optical aberrations, refractive errors experienced by light-waves as they traverse the eye's optical system, luminance of the stimuli, exposure duration and retinal location of receiving photoreceptors [61].

Spatial visual acuity is the smallest spatial detail that can be visually detected, discriminated, or identified [61]. Generally it is recognised that there are four kinds of spatial visual acuity, i.e. Detection, Localisation, Resolution and Identification Acuties. In this work we are focused on Identification Acuity, as that is widely used by optometrist and clinicians as a measure of individual differences in spatial visual acuity [4, 34]. Identification acuity is also known to be a good predictor of real world visual acuity and performance at life tasks [62, 92].

Spatial vision is another important aspect of spatial perception, and it is well established that it strongly influences the ability to see clearly [66, 62, 36]. Spatial vision is the contrast sensitivity of the perceiver's visual system. Contrast sensitivity is the ability to distinguish between the relative luminance of stimuli. For example if there are two greyscale lines drawn side by side: What is the minimum grey colour difference between the two lines for the lines to be perceived as different shades of grey?

A person's spatial visual acuity and spatial vision are reasonably stable though they are known to change, usually decreasing, with age. Within an individual spatial perception can differ significantly between each eye [71]. Measures of spatial visual acuity do not necessarily correlate with measures of spatial vision [39].

Significant differences in the ability to see detail can often be due to the perceiver's position relative to what they are looking at. That is, the visual angle of the target visual stimulus varies as a function of distance, i.e. objects in the



Figure 2.2: Examples of how different forms of colour blindness effect perceived colours in a photograph. First image full colour original. Second (protanope), third (deutanope) and fourth (tritanope) images were generated with the Vischeck colour blindness simulator.

world are perceptually reduced in size as the viewer moves away from them.

Spatial visual acuity is covered in more depth in Section 2.4, while Section 3.5 of Chapter 3 provides more details about spatial vision modelling.

2.3.2 Differences In Color Perception

Colour perception and colour theory is a deeply studied field [59, 43]. Fundamental questions remain about how colour perception functions. Questions include how many colour categories exist between cultures [12, 50]? Why do these categories exist [7, 97]? How are differences between colours calculated [54, 76, 96]? What, if its even possible to create, is the optimal perceptually uniform colour space [32, 53]?

In this work we are focusing on the effects of differences in colour perception. Colour deficient vision (dyschromatopsia), or colour blindness as it is more colloquially known, affects approximately 8 to 10% of males and 0.4% of females [43, 45]. Recent research indicates that a mild form of colour blindness may occur in heterozygous females [13], i.e. mothers carry and pass colour blindness on to their children, but it was thought they did not experience any form of colour blindness.

A common misconception about people with colour deficient vision (CDV) [59] is that they only see in shades of gray, as though continuously seeing the world as a black-and-white photograph. This is not the case, the vast majority of people with CDV perceive a range of colours (see Figure 2.2).

The human eye consists of approximately 6 to 7 million cones and 120 million rods [16, 63]. There are known to be three classes of cones, each of which has a peak sensitivity to a different wavelength of light. One way of thinking of these classes of cones is that they classify light into three frequency ranges which correspond to the colours perceived as red, green and blue. Rods enable humans to see at night and contribute little to colour vision. Rods are a lot more sensitive to light and are somewhat overwhelmed in normal daylight. Thus cones enable photopic (colour) vision and rods enable scotopic (low light) vision.

Colour vision deficiencies are the result of a reduced light sensitivity of the cones or a complete lack of one or more classes of the cones. When there is a reduced sensitivity in one class of cones a person is said to have anomalous trichromatic colour vision. When one class of cone is completely missing, or non-functioning, a person has dichromatic vision. Cone monochromacy is when only one class of cones are functioning, and if no cones are functioning or present a person has achromatic vision.

Inherited variations in colour vision cannot currently be ‘fixed’. These variations have a physiological basis, which is most often associated with X chromosome genes. Colour vision deficiencies can be acquired due to the normal aging process and may also develop due to a range of illnesses, such as cataracts, diabetes and age-related macular degeneration.

Colour perception, colour spaces and colour deficient vision are covered in more depth in Section 2.4.5 and Section 3.6 of Chapter 3.

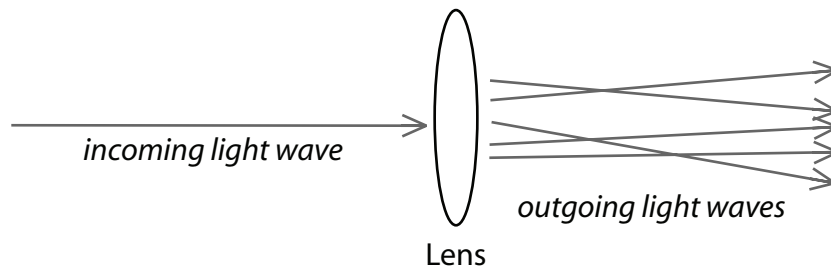


Figure 2.3: Example of how a single point of light can deviate as it traverses an optical system (lens).

2.4 Modelling Human Vision

This section provides a brief introduction to modelling and simulating facets of human vision, spatial perception and colour perception, optical aberrations, wavefronts and using Zernike Polynomials to model early stage human vision. Many of the details here are elaborated upon and built upon in Chapter 3.

2.4.1 Optical Aberrations & Refractive Errors

In order to model individual differences in visual acuity, we model and simulate optical aberrations and refractive errors in individual human eyes (see Section 3.5.1 and Section 3.7.1).

Optical aberrations occur due to light waves traversing the human eye undergoing numerous alterations in paths taken; often deviating from an ideal path only realisable in a perfect optical system (see Figure 2.3). These optical aberrations and refractive errors occur in many sub-parts of the human eye. Optical aberrations and refractive errors occur due to optical deflection, scatter, interference, absorption, refraction, diffraction [52, 40]. For example the cornea partially absorbs wavelengths of light commonly perceived as yellow, while the lens and cornea take advantage of differences in optical densities to create refractive indexes that focus incoming light waves.

Optical aberrations and refractive errors in one part of the eye may cancel or

increase the affect of an aberration in another part of the eye. The sum total of errors and aberrations affects the final light wavefront reaching the light receptors in back of the eye. Ideally the light waves are perfectly focused on the fovea thereby forming a clear retinal image. Variations in fovea receptor functions lead to variations in colour perception, which we take into consideration later when modelling colour perception (see Section 3.6).

Differences in spatial visual acuity can be modelled as optical aberrations and refractive errors in the optical system of the eye. One approach to simulating differences in spatial visual acuity involves creating a high fidelity model of each of the optical sub-systems of the human eye and then simulating light waves traversing each optical sub-system in turn [6, 65, 57]. For example CWhatUC simulates visual acuity with a high fidelity model of early stage optical and biological processing of light waves in the cornea [35]. High fidelity modelling of the human eye is a complex multi-stage process, which is an open research area and computationally expensive [27, 1].

2.4.2 Point Spread Functions

In this work we use Point Spread Functions (PSF) extensively when modelling spatial human vision, as detailed in Section 3.5.2 and Section 3.5.

First conceived and used at the turn of the last century to test the optical quality of lens used in telescopes. PSFs have since proven to be invaluable as a measurement and descriptor of optical aberrations in the human eye and other optical systems. In the mid-60s Shack-Hartmann developed the Shack-Hartmann Wavefront Sensor which enabled the accurate measurement of the PSF of a living human eye. This aberrometer technology, and more recent advances, are widely used by ophthalmologists when planning corneal ablations, e.g. LASIK for corrective eye surgery [48].

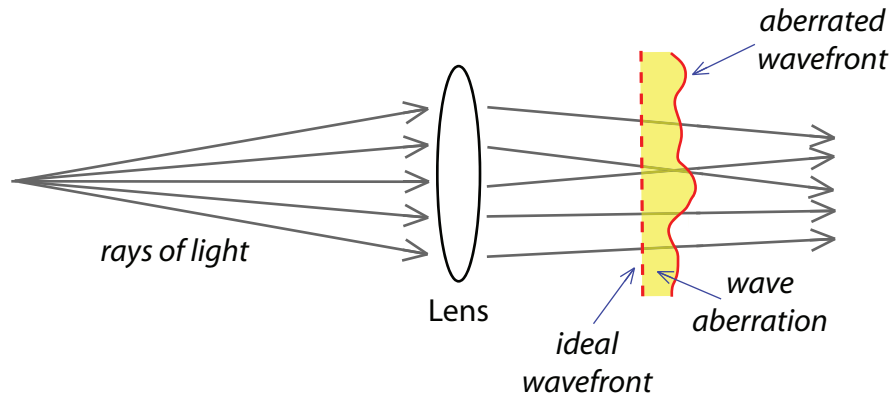


Figure 2.4: Example of ideal and aberrated wavefronts generated by light rays traversing an optical system.

2.4.3 Wavefronts

As described in Section 2.4.1, modelling the human eye requires techniques that describe how optical aberrations effect light travelling through individual human eyes. Within Optometry and Ophthalmology wavefronts are used to describe how light travelling through an optical system (the eye) is affected by optical aberrations. Wavefronts encode the effects of optical aberrations on light waves passing through an optical system.

Wavefronts can be considered as perpendicular to the direction light rays travel in (see [48], Chapter 2). A wavefront effectively describes the properties of multiple rays of light at a point along the rays of light, as shown in Figure 2.4.

For optical systems there is a perfect wavefront, that which occurs when no optical aberrations are present. When optical aberrations are present the wavefront is said to be aberrated. The Shack-Hartmann Wavefront Sensor mentioned in Section 2.4.2 records the wavefront aberrations of human eyes.

Wave aberration is defined as the difference between the actual aberrated wavefront and the ideal or intended wavefront [48]. In order to mathematically describe wavefront aberrations a set of basis functions called Zernike Polynomials are

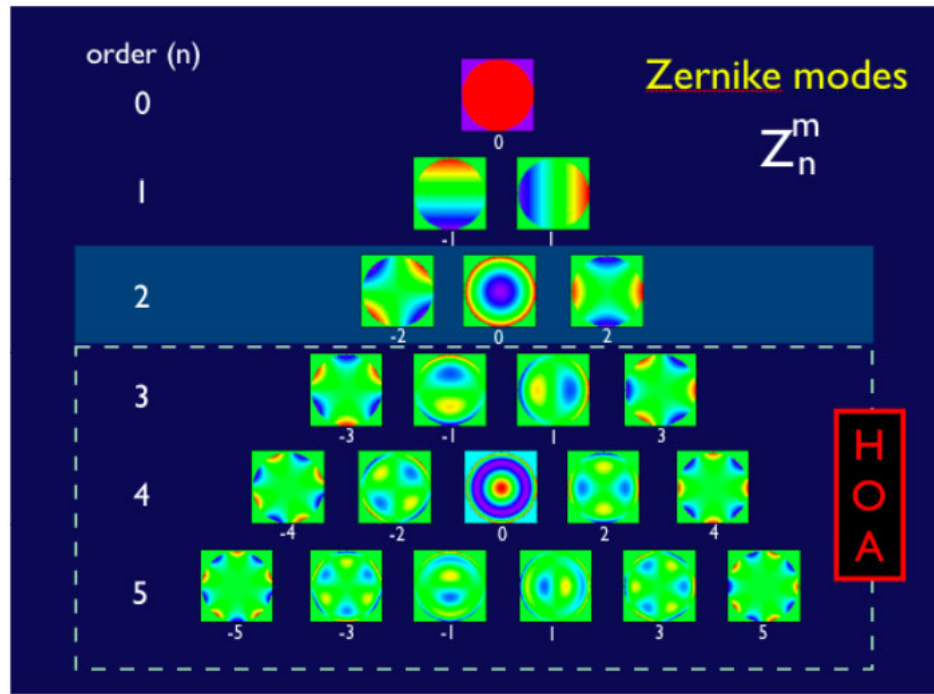


Figure 2.5: Pyramid of Zernike modes where each mode is an aberration (from [72]).

employed. The basis functions are referred to as Zernike modes (see Section 2.4.4).

Zernike Polynomials mathematically describe the difference between a perfect wavefront and an aberrated wavefront, as depicted in Figure 2.4.

2.4.4 Zernike Polynomials Overview

Optometrists and Ophthalmologists have adopted Zernike Polynomials to describe the aberrations of a wavefront [80, 81, 99, 48].

Zernike Polynomials resemble Fourier Transforms in that a complex wavefront can be decomposed into a set of component Zernike modes, just as a Fourier Transform can decompose a complex wave into a simplified set of sine and cosine waves.

The wave aberration function W_{eye} (Equation 2.1) can be decomposed into Zernike modes Z_n^m (Equation 2.2). Each mode describes a specific optical ab-

Equation 2.1: Wavefront aberration function as weighted sum of Zernike Polynomials [82].

$$W_{eye}(p, \theta) = \sum_{n,m} C_n^m Z_n^m(p, \theta) \quad (2.1)$$

where

C_n^m is Zernike coefficient in microns μm
 Z_n^m is double indexed Zernike mode (Equation 2.2)

and

p is normalized pupil radius
 θ is azimuthal component from 0 to 2π radians

erration, such as the second-order modes defocus (Z_2^0) and astigmatism (Z_2^{-2} , Z_2^2), which are shown in Figure 2.5. Third-order modes and up describe higher order aberrations (HOA), also shown in Figure 2.5.

To obtain W_{eye} an aberrometer measures the wavefront of an eye, alternatively W_{eye} can be created by summing multiple Zernike modes (Z_n^m) weighed with Zernike coefficients (C_n^m), as shown in Equation 2.1.

Zernike coefficients control how much each Zernike mode contributes to the total wavefront aberration (Equation 2.3). Equal value Zernike coefficients do not mean Zernike modes contribute equally to the degree of perceptual aberration experienced, i.e. $C_2^0 = 0.5$ has more effect on the degree of optical aberrations in an eye than $C_5^{-5} = 0.5$. Thibos et al., have demonstrated that the “*magnitude of aberration coefficients in any given individual tends to be smaller for higher order modes than for lower order modes*” [82], while also showing for a sample of two hundred normal eyes that the “*average wavefront variance falls exponentially with radial order n* ” [84]. Order n is as shown in Figure 2.5.

Total aberration in a wavefront is described with the total wavefront RMS (Root Mean Squared), as listed in Equation 2.3. When reading research involving Zernike coefficients it is important to note that the terms RMS and

Equation 2.2: Zernike Polynomial as defined in [81] with a double indexing scheme. Generates Zernike modes as illustrated in Figure 2.5.

$$Z_n^m(p, \theta) = \begin{cases} N_n^m R_n^{|m|}(p) \cos m\theta & m \geq 0 \\ -N_n^m R_n^{|m|}(p) \sin m\theta & m < 0 \end{cases} \quad (2.2)$$

where

$$R_n^{|m|}(p) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! [0.5(n+|m|-s)]! [0.5(n-|m|-s)]!} p^{n-2s}$$

$$N_n^m = \sqrt{\frac{2(n+1)}{1+\delta_{m0}}} \quad \begin{array}{l} m = 0 \text{ then } \delta_{m0} = 1 \\ m \neq 0 \text{ then } \delta_{m0} = 0 \end{array}$$

and

p is normalized pupil radius

θ is azimuthal component from 0 to 2π radians

Equation 2.3: Total wavefront RMS (root mean squared).

$$RMS = \sqrt{\sum_{n,m} C_n^m{}^2} \quad (2.3)$$

where

C_n^m is Zernike coefficient in microns μm

the Zernike coefficient are sometimes used interchangeably, which can lead to confusion.

2.4.4.1 Visual Acuity As A Function Of Zernike Coefficients

It has been demonstrated that visual acuity decreases as the value of RMS (Equation 2.3) increases [84], that is as RMS increases our ability to see decreases. This correspondance between visual acuity and RMS also applies to individual Zernike coefficients. That is, within a single Zernike mode increases in the Zernike coefficient correspond to a decrease in visual acuity.

Optometrist commonly measure the visual acuity of a human perceiver using a range of psychophysical tests, such as ETDRS eye charts (see Figure 3.5). The tests generate LogMAR measures of a person's visual acuity.

Drawing a relationship between LogMAR and amount of aberrations in a wavefront is an ongoing focus of research [91, 83, 38, 23, 48]. For this thesis the implications are that a LogMAR measure of visual acuity, which is easily obtained with an eye chart, is not yet sufficient to re-create a simulated wavefront aberration corresponding to a specific individual's eye function. This has implications for the capabilities of the visual acuity predictor, which is discussed in Chapter 4. Additionally it also impacts the experimentally design for testing how well the predictor correlates with human judgements about perceptual stability (see Chapter 5).

2.4.5 Modelling Colour Deficient Vision (CDV)

In the human eye the Long-, Medium- and Short-wave (LMS) colour receptors create a three dimensional colour space. Variations in the sensitivity and presence of the colour receptors change the shape of each perceiver's colour space.

Relating a normal observer's LMS colour space to CDV reduced LMS colour spaces was achieved by finding and measuring unilateral CDV perceiver's colour experiences [15]. Unilateral colour perceivers are individuals with one eye exhibiting normal trichromatic colour perception, and the other eye a variant of CDV. By having unilateral observers identify perceptually equivalent colours between each eye colour transforms could be created for converting from a standard three dimensional colour space into CDV colour spaces.

For modelling CDV we use the simulation technique outlined in [87, 15] and implemented in Vischeck [30], as this is one of the established approaches for simulating CDV. Recent research based on molecular genetic analysis in conjunction with colour identification testing suggests that the perception of colour during CDV is a more complex and dynamic process than captured in existing models [88, 13].

2.5 Sensation & Perception In Vision

As has been presented in the previous sections the simple question "What do people see in a design?" involves many complex facets. One final facet which is worth emphasising is that working with low-level vision requires dealing with crossing the boundary between sensation and perception. A dictionary definition of Sensation and Perception tells us:

- Sensation is a physical stimulation that occurs when something happens to or comes in contact with the body.
- Perception is the conversion of a stimulus into that which an organism becomes aware of.

The sensation and perception boundary makes many aspects of understanding what people see in a design as an "experience" a challenging question. It

is an issue the constantly raises its head during analysis. Understanding what we are analysing and what the underlying data means in both sensation and perception terms is critical. Especially because of the non-linearities inherent in sensation and perception, e.g. Steven's Law and the relationship between luminance (cd/m^2) and perceived intensity, visual acuity decreasing as perceiver approaches stimuli [41].

Colour is a prime example of a perceptual quality. It does not occur in the physical world, wavelengths of light do. Our brain and visual system maps different wavelengths of light into different colours. In theory there is a purely arbitrary relationship between the wavelengths of light and the colour precept.

Modelling the relationships between perceptions and sensations is particularly relevant when dealing with colour. Perceptually acceptable colourspaces are required for analysing the colour used in designs; whether colour is used to generate details and textures or colour is used to denote categories. The encoding and colourspaces used for images stored on computers do not easily translate to colour percepts. For instance when working with a triple RGB (Red, Green, Blue) value for a pixel how do we know what colour percept corresponds to a triple [50, 74, 7, 97, 12]?

2.6 Summary

In this Chapter the implications of a relationship between design, physiology and individual differences is presented. The history of individual differences in HCI and how it relates to research in psychology is touched upon.

Further details were then provided on the physiological aspects of individual differences in visual function, with an elaboration on differences in spatial and colour perception.

Modelling & Simulating Human Vision

This chapter presents two models of eye function for generating simulated retinal images. The models are used to simulate the effects of individual differences in spatial and colour perception on visual designs.

How the models operate and the underlying basis for the models is elaborated upon. With models and simulations of eye function we are able to simulate how different eyes see a design.

3.1 Introduction

With models of visual function a visual design can be evaluated to establish what parts of a design are easy or difficult to see, see Figure 3.1. For example when viewing a visual logo from a long distance away, how small must the visual features of the logo be before they are impossible to clearly and easily see and interpret?

The Human Visual System (HVS) is a highly complex structure about which many ambiguities remain; so in creating our models we had to decide what is

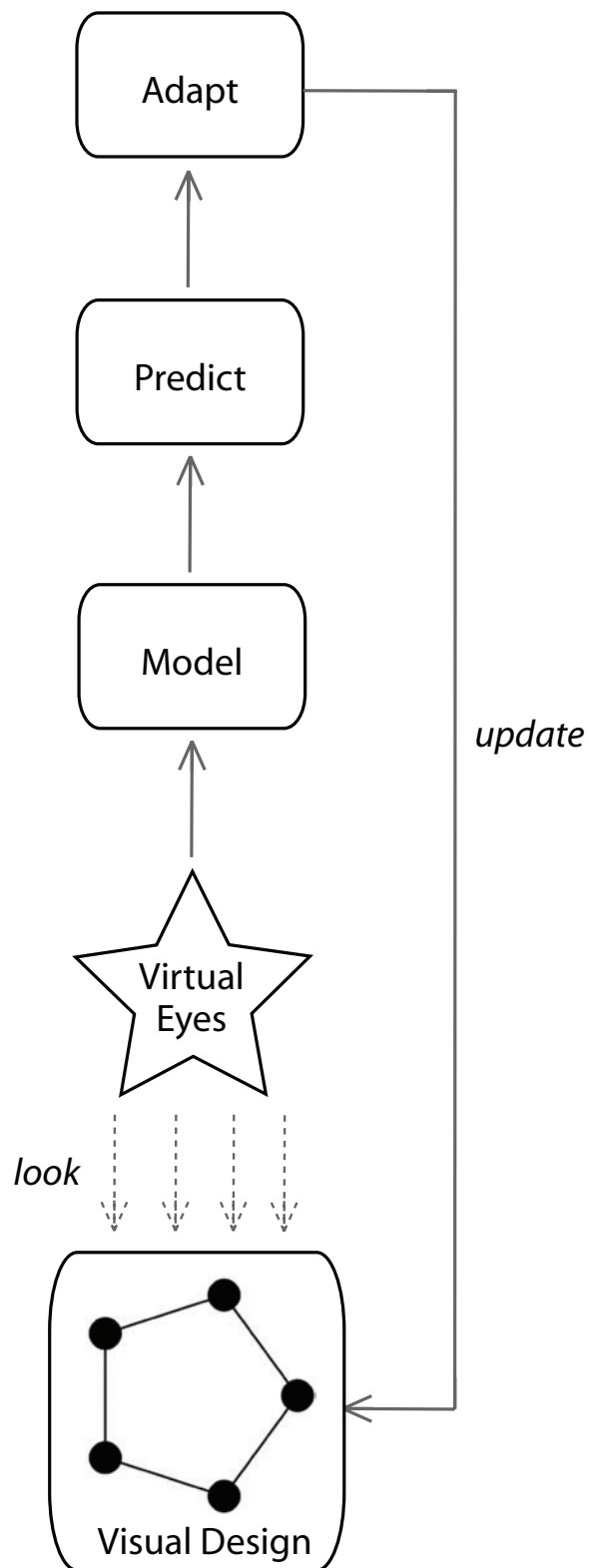


Figure 3.1: To adapt a visual design models of visual function are employed. The models integrate with predictors, which feed into adaption techniques for improving the layout and presentation of visual designs.

feasible to model. To this end we create models focused on low-level vision. Low-level vision occurs in the early stage of the HVS, predominately due to the physiological structure of the eye.

Also desirable is that the parts of seeing we focus on are those that have large impacts upon the perception of designs. For instance the effect of individual differences in pupil size and pupil accommodation could be examined but we would not necessarily expect these differences to have significant impacts on how people perceive designs.

As is briefly mentioned in Section 1.4.1 and discussed Section 2.3 we focus on the effects of colour perception and the effects of spatial perception on the perception of information visualisations. Differences in colour perception are worth exploring because colour is a basic building block of graphical designs, often denoting separation and grouping (see Figure 1.2). Spatial perception is also worth exploring because the ability to see detail is a basic building block in graphical designs.

3.2 Human Visual System Models

The aim here is to construct usable models that can be applied to simulating low-level early stage vision. Such models are primarily concerned with the effects of optics and photoreceptors in the human eye. However, more complex models that simulate aspects of vision such as neural defocus were outside the scope of this thesis.

The decision to limit the scope allows for a tractable approach to the question "What do people see in a design?". While there is a large body of knowledge about human vision there remains many unanswered questions and issues [63, 91, 3].

In other domains such as research into video codecs there are numerous Human Visual System (HVS) metrics for automatically measuring the perceived quality of video streams [95, 94]. Applications of metrics include improving video encoding to remove visual features that are less important for human vision.

The central approach in this thesis is that the same applied vision science approach can be taken to quantifying the perceived quality of interfaces and information visualisations. Models based on knowledge of vision science present a stepping stone to empirical quantifications of the experience of seeing designs.

3.2.1 Model Requirements

The required outcome for the models developed here are that they can be built upon to create general purpose predictors of the perceptual stability of visual interfaces and designs. Recall, Section 1.4.2 defines perceptual stability as:

A measure of how stable or unstable a visual design or image is due to differences in a perceiver's (spatial or colour) perception.

We define stable and unstable as:

A stable percept is one which is subject to the least amount of perceived change in a design - due to how the physiological structure and function of the eye and optical system biases the perception.

An unstable percept is one where the design is significantly affected by how the low-level physiology and function of the eye biases the perception.

An example of a stable percept is the word "Human" on the left hand side of Figure 3.2; the right hand side shows an unstable percept. In Figure 3.2 the



Figure 3.2: An example of a stable and unstable percept. From top to bottom optical blur increases. The large version of Human is more perceptually stable than the small version.

amount of optical blur increases from the top of the figure to the bottom of the figure. As the amount of blur increases, the large version of the word "Human" remains legible while the small version of "Human" quickly becomes impossible to read. As people with varying levels of visual acuity view the large word the legibility of text is more stable between viewers. Differences in normal visual acuity will not lead to significant differences in the perception of the large text. When people view the small text the small differences in visual acuity have a significant effect on the perception of the word.

A further requirement of the models is that they correlate with human experience while also being usable for critiquing the perceptually induced changes in designs.

However a high biological fidelity from the models is not a requirement. For examples of research into high biologically models, CWhatUC [35] is an example of simulating corneal visual acuity, while elsewhere researchers have taken the approach of creating a photo accurate model of the human eye [27]. The purpose of our models is to help understand the effects of individual differences in perception on designs, rather than help understand the interplay of individual differences in vision.

Models that are constructed here should not be brittle, such that advances in vision science do not quickly render them inapplicable. Vision science knowledge is imperfect hence attempting to construct and apply models that are highly dependent on specific aspects of current known visual function could quickly render such models unusable. To achieve models capable of remaining applicable with advances in vision science the models are developed with a modular multi-stage compartmentalised structure, as shown in Figure 3.8 and Figure 3.11.

A final essential requirement is that the models can be built upon to create predictors capable of providing metrics of both degree of perceptual change and location of perceptual change.

In summary our requirements for the models are that they:

- are capable of being built upon to generate perceptual stability predictors
- are usable for critiquing the perceptually induced changes in designs
- help understand the effects of individual differences in visual function
- are modular and so capable of adapting to increasing knowledge in vis-

ion science

- integrate with predictors to provide degree and location measures of perceptual change.

3.3 What Was Modeled

As was previously discussed in Section 2.3 the two aspects of vision that are modeled in this thesis are individual differences in colour perception and spatial perception, primarily spatial visual acuity.

3.3.1 Differences in Spatial Visual Acuity

If objects are too small or too cluttered they cannot properly be seen in a design. Often when designing a visualisation or interface there is a trade off between the amount of information we have to communicate versus the amount of display real-estate we have to communicate with.

By modelling spatial visual acuity and spatial perception we are enabling the examination of the interplay of the size of graphical objects used in designs and our experience of seeing them. For more background see Section 2.3.1.

3.3.2 Differences in Colour Perception

Colour is widely used in graphical designs [90], often to denote categorical distinctions and groupings within data, e.g. green signifies a node representing a male, black signifies a node representing a female (see Figure 1.2). Colour is also a fundamental building block of more complex visual components, such as textures.

By modelling colour perception we are enabling the examination of the effects

of a crucial low level unit of composition in graphical interfaces and visual designs. For more background see Section 2.3.2.

3.4 How Did We Model

Our models are constructed to be feature independent and use by reference analysis. Feature independent means the analysis techniques do not analyse and require specific visual features in order to function correctly. This contributes to creating general purpose predictors, which are capable of analysing a wide range of visual designs. While by reference analysis contributes to making tractable the problem of measuring visual information loss due to optical defocus. Further details are provided in Section 3.4.1 and Section 3.4.2.

3.4.1 Feature Independent

A model which is feature independent does not rely on complex visual features for its analysis. For example the spread of a line has been used as a feature for metrics to analyse, e.g. zero-crossings [57] (see Figure 3.3). Line spread is a measure of how a visual line changes under visual transforms. In perceptual video metrics it was used as an early measure of video quality.

3.4.1.1 Issues

Problems arise with line spread as a general purpose measure as it relies on the visual features of the design. If a design does not consist of readily identifiable lines then the line spread cannot be easily calculated.

Another more serious complication is that a metric such as line spread is too simplistic. In this particular case the simplicity arises because multi-channel theories of human vision suggest that there are different feature detectors de-

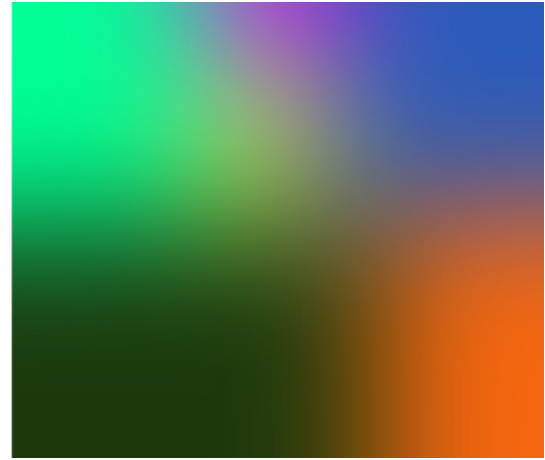


Figure 3.3: Example of side by side feature dependent analysis of potential designs for Information Visualisations. With a line spread metric it is possible to analyse the design on the left but not the design on the right.

pendent on attributes including line rotation and line scale. A response by one feature detector can be tuned to a line orientation of 60 degrees and may ignore a line rotated at 50 degrees [63]. Questions then arise about whether line spread matters more or less at different orientations and at varying scales in human vision.

Feature dependent metrics can be fundamentally incorrect in unexpected and potentially misleading ways when the psychophysical function of the HVS is not adequately taken into consideration.

3.4.1.2 Design Space Constraint

When constructing the model we want a more generic solution that is less dependent on specific features in visual designs. Another important motivation for this is that the analysis techniques and models should not place constraints on the design space that potential visual designs can be created from. Constraints are placed on the design space because feature based models limit the side-by-side comparison of designs.

For an example of a potential constraint, attempt to use line spread to compare

the two potential information visualisation designs as shown in Figure 3.3 - when the designs are subject to blur. The two information visualisations in Figure 3.3 cannot be measured and compared by using a line spread metric as the heatmap based visualisation (right image) cannot be analysed using line spread. Since they cannot be measured with the same metrics and models they must be measured with different metrics, which raises further issues and questions about the equivalence of various metrics and models [57].

3.4.2 By Reference

By reference analysis requires a starting “perfect” image for doing comparisons against. That is, the model has a reference image against which to make judgements about the effects of perception on designs. A comparison of the reference image with the changed image occurs when generating the measures, where the updated image is changed due to differences in perception.

For example, version 1 of the letter chart show in Figure 3.5 is the starting reference image. Version 2 is the same chart subject to optical blur, and version 3 is the chart subject to a greater degree of optical blur. For the perceptual stability metric version 2 and version 3 are compared with version 1, and each other, to quantify the perceptual stability for spatial perception.

3.4.2.1 Spatial Perception

In this work for measuring the effects of differences in spatial perception reference images are required for the analysis. Without reference images it is challenging to gauge what parts of an image are susceptible to changes in spatial visual acuity.

For example consider the case of no-reference analysis of a single image that has already undergone a significant amount of optical defocus. Without a ref-



Figure 3.4: Examples of how colour blindness effects the perception of colour in a photograph. First image full colour original. Second (protanope), third (deutanope) and fourth (tritanope) images were generated with the Vischeck colour blindness simulator [30].

erence image it is difficult to accurately estimate which parts of the image have experienced significant information loss.

Consider images experiencing blurring due to defocus, where information is lost from the images. Without a reference image restoring some of the information lost is possible but to do so extra information is required, such as the Point Spread Function (PSF) (see Section 3.5.2) of the blurring operation. With a PSF available the inverse of the PSF can be used to reconstruct some, though not all, of the information lost. Adaptive optics employed in telescopes are an example use of inverse PSFs.

3.4.2.2 Colour Perception

For colour perception a by reference image is required because during colour perception information is lost from the designs when perceived by individuals with colour deficient vision (see Section 2.3.2). If there is only an image showing what a colour deficient perceiver would perceive it would be impossible to re-colour the image back to what people without colour deficient vision see.

For example, Figure 3.4 shows a photograph as seen by normal colour perceivers (the by reference image), also shown are corresponding versions of the

photograph as perceived by individuals with colour deficient vision. Without the by reference image we are unable to identify which visual features in the photograph are effected by colour deficient vision.

Colour deficit vision reduces the dimensionality and shape of colour spaces, e.g. a non-linear conversion of a 3 dimensional colour space to a 2 dimensional colour space. Reconstructing the original 3 dimensional colour space, then assigning pixels to original locations is not feasible without extra information. Without knowing the colour usage in the original images it is impossible to measure the degree or location of perceptual changes.

One could consider an alternate approach to by reference, which is the development of an abstract representation of normal colour distribution in the world [79, 7, 97]. In this case significant differences from the normal colour distributions introduced by colour deficient vision may be detectable. This theorised approach differs from by reference in that there is an external model of colour distribution. Though in both cases extra information is present that can be used to make comparisons and measures of the changed image. Developing empirically valid external models of normal colour distributions is a significant undertaking in itself, with relevant questions such as "What are normal colour distributions in visual designs versus in natural scenes?"

3.5 Modelling Spatial Visual Acuity

Spatial visual acuity is the smallest spatial detail that can be visually detected, discriminated, or identified [61]. Building on the background to visual acuity presented in Section 2.3.1 of Chapter 2 we now present how human visual acuity is modelled.

To model visual acuity we use a physiologically valid means of deforming the visual designs, which simulates how differences in visual acuity affect the per-



Figure 3.5: Three simulated retinal image examples of how people with different visual acuities could see the same eye chart.

ception of the designs (see Section 3.5.1). Physiologically validity is achieved through using Zernike Polynomials that encode optical aberrations in the human eye (see Section 3.5.3.1). Point Spread Functions (PSF) are generated with the Zernike Polynomials (see Section 3.5.2), then the PSFs are used as convolution kernels to deform the visual designs and interfaces into how people see the visual designs (see Section 3.5.4).

The spatial visual acuity model enables the creation of simulated retinal images, which depict how different perceivers see the same visual design. For example Figure 3.5 shows a simulated retinal image of how three different perceivers could see the same eye chart.

Figure 3.8 shows a flowchart depicting how the Zernike Polynomials and PSFs integrate to generate the simulated retinal images.

3.5.1 Optical Aberrations & Refractive Errors

The approach we have adopted for modelling optical aberrations is to use aberration maps of the optical system in the eye [48] (see Section 2.4.1). Aberration maps describe how a wavefront of light deviates from an intended wavefront. For example when measuring the optical quality of a lens we can examine and record details about the physical material of the lens, alternatively we can use an aberration map which records details about the effect the

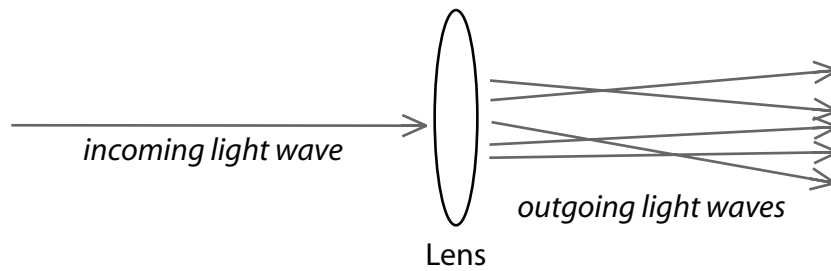


Figure 3.6: Example of the point spread of a single point light source as it travels through an optical system.

material of the lens has on how light travels through the lens.

Aberration maps offer the advantage of rapid computation, and within Ophthalmology and Optometry aberrations maps are a well established tool for modelling and the focus of continuing research [48], e.g. used for LASIK pre-operative surgery planning.

A significant advantage of aberration maps is that they enable modelling of individual eye function, because aberrations maps can be generated by directly measuring an individual's eyes. We take advantage of individual maps when modelling the effects of individual differences in spatial visual acuity (see Section 2.4.4.1).

3.5.2 Aberration Maps As Point Spread Functions

Point Spread Functions (PSF) are a description of the point spread of a point source of light after it travels through an optical system (see Figure 3.6). PSFs describe the impulse response of an optical system, see Section 2.4.2 for background on PSFs.

Simulated retinal images are needed for the perceptual stability metric, i.e. to compare the original visual design with the visual design as it would be perceived by different viewers. A PSF can be used as a convolution kernel for transforming an image, i.e. the style of transform applied to the original image

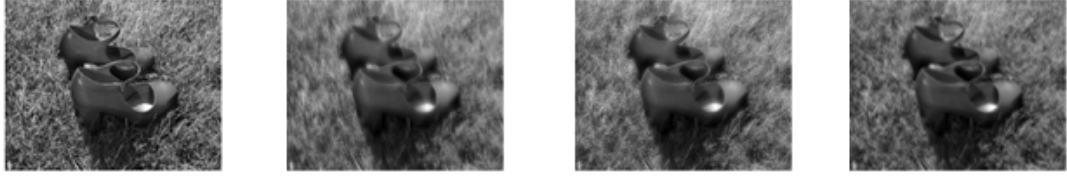


Figure 3.7: Example of the effects of three different Zernike Mode PSFs applied as convolution filters to a photograph of shoes on grass. First photograph is the original version.

is encoded by the PSF (see Figure 3.7). PSFs that encode aberration maps of a human eye enable the creation of simulated retinal images.

A 2D Gaussian can be used to create a blurring kernel PSF, where increases in the Gaussian σ increases the amount of image blur. Some of our early research used a 2D Gaussian to generate the image blurs [10], but in terms of eye function the biological fidelity of the 2D Gaussian for simulating visual acuity defocus is unknown.

A further limitation with using a 2D Gaussian is that it approximates only one kind of optical aberration commonly found in human eyes, i.e. defocus. There are other common optical aberrations such as Coma and Spherical Aberration, which must be simulated otherwise the PSFs lack biological validity. A lack of biological validity limits the correctness of the simulated retinal images.

3.5.3 Biological Fidelity With Zernike Polynomials

To create biologically valid PSFs we use Zernike Polynomials to generate wavefront aberration maps. Zernike wavefronts are then transformed into PSFs. Converting a wavefront to a PSF is a standard technique in Ophthalmology and Optics, details of which can be found elsewhere [48].

Show in Figure 3.7 are three biologically valid simulated retinal images. Each of the images shows how a photograph of shoes on grass (column 1) is visually transformed due to the optical aberrations of Coma (column 2), Spherical

Aberration (column 3) and Quadrafoil (column 4).

3.5.3.1 Zernike Polynomials, Zernike Coefficients & RMS

Background on Zernike Polynomials, Zernike Coefficients and RMS with relevant definitions and equations is provided in Section 2.4.4 and Section 2.4.3 of Chapter 2.

Using Zernike Polynomials to simulate human vision offers the ability to simulate different pupil sizes. For this thesis we standardise on a pupil dilation of 6 mm with monochromatic light at 555 nm wavelength, the approximate peak of photopic sensitivity. Smaller pupil dilation sizes such as 3 mm in diameter are less susceptible to HOA, while larger pupil sizes like 8 mm incur greater effects from HOA. Transforming wavefronts measured at one pupil size and wavelength to another pupil size and wavelength is possible [73, 18], which is useful when performing comparative analysis of multiple eyes.

3.5.4 Simulating Retinal Images: Visual Acuity

With the Zernike Polynomials and PSFs we can create simulated retinal images, which take into consideration the effects of differences in visual acuity.

As shown in Figure 3.8, with examples in Figure 3.9, the retinal image simulation technique works as follows:

1. Visual design is captured as a grayscale image. At this stage a colour image is not used as simulating chromatic aberrations introduces another level of complexity.
2. Zernike coefficients are set to create different types of wavefronts, which simulate different types of visual aberrations. Multiple Zernike wavefronts W_{eye} of aberrations in an eye may be generated.

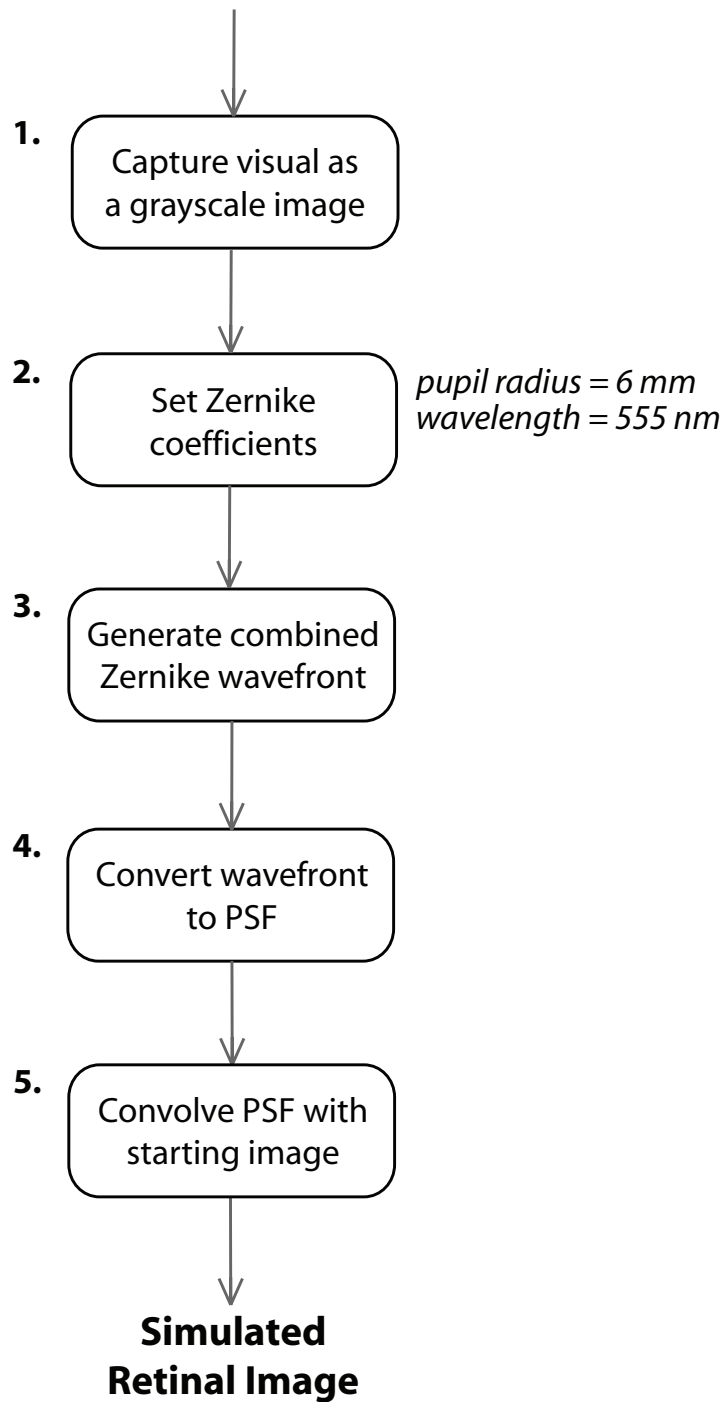


Figure 3.8: Steps in algorithm for simulating visual acuity to generate a retinal image of a visual design. See Section 3.5.3.1 for more details. Examples of simulated retinal images are shown in Figure 3.9.

Original Image



Grayscale Version Of Image



Examples Of Simulated Retinal Images

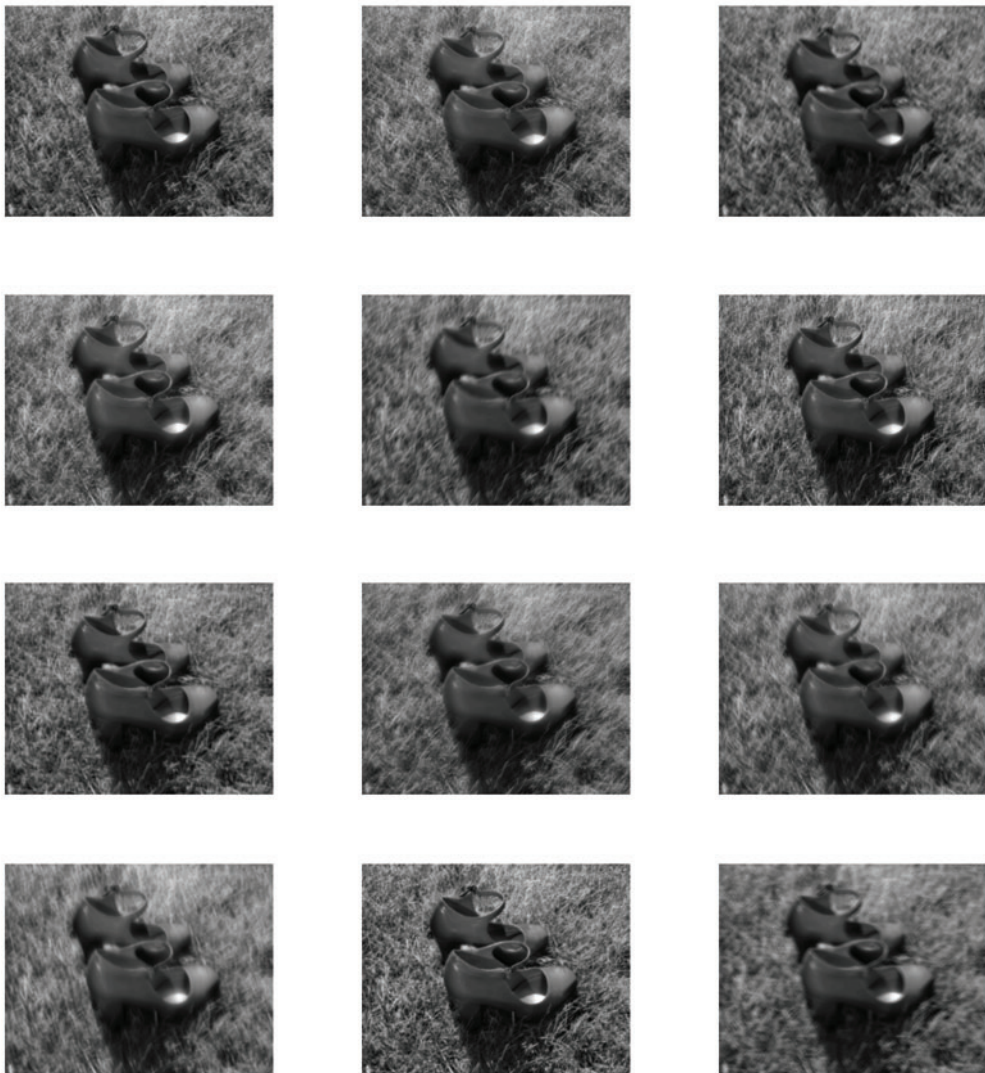


Figure 3.9: Examples of visual acuity simulated retinal images generated from a photograph of red shoes on green grass. Simulated retinal images are within normal random visual acuity distribution (see Section 3.7.1). Differences between the simulated retinal images are often subtle.

3. Combined wavefront W_{total} is generated, which is wavefront due to each of the eye aberrations W_{eye} .
4. The wavefront is converted to a PSF via a transform, which includes a Fourier Transform [48].
5. PSF is used as a convolution kernel to transform input image into a simulated retinal image.

The simulated retinal image is used when generating the perceptual stability measure for visual acuities. Within Ophthalmology and Optometry this is an established approach for generating simulated retinal images [48, 6, 38, 23, 2, 91, 84]. Chapter 4 presents how the simulated retinal images are analysed to establish the perceptual stability of the visual designs.

3.6 Modelling Colour Perception

In Section 2.3.2 of Chapter 2 we provide background on colour perception. Building on that knowledge we now look at how colour perception is modeled for this work.

Colour can be easily thought of as absolute, that is it exists independent of observers. This is not the case because colour is a perceptual quality; a mapping of lightwaves from sensation to colour percepts (see Chapter 2).

When modelling colour perception in computing terms we need to understand who or what is doing the observing, i.e. what creates the mapping between a numerical description of a colour and a colour percept (see Section 3.6.1)? Once this is clarified we can begin treating differences in colour perception as differences in colour spaces (see Section 3.7.2). Understanding what changes and what is perceptually equivalent between two or more N-

dimensional colour spaces enables the creation of perceptual colour stability metrics (see Chapter 7).

Colour deficient vision is an example of the effects of individual differences in colour perception. With a model and simulations of colour perception we evaluate the impact of individual differences in colour perception (see Section 3.6.7 and Chapter 6).

3.6.1 N-Dimensional Colour Spaces

When considering images in computing terms the values of image pixels indicate the colour of a pixel. These pixel colour values are references within N-dimensional colour spaces (see Figure 3.10). Locations within a particular colour space dictate the colour of a pixel.

If the N-dimensional colour space is translated, transformed or altered then the colour of a pixel typically changes. Pixels thereby “look-up” colours in colour spaces. Without a reference colour space, pixel values have no inherent colour.

For example with a RGB triplet [136, 45, 213], where each triplet entry has a range between 0 and 255. What colour does RGB[136, 45, 213] refer to? In the case of the standard computer based RGB colour space: Purple. There is no mathematical function that can generate correct colour descriptions from a tristimulus without a reference colour space.

In this work we want to understand the effects of perceptual differences in colour perception on graphical designs. Framing that based on the previous paragraphs it is clear that we are asking about the effects of altered colour spaces. Specifically we are asking whether pixel values still reference the same perceptual colours, when colour spaces are altered due to differences in perceiver’s visual systems.

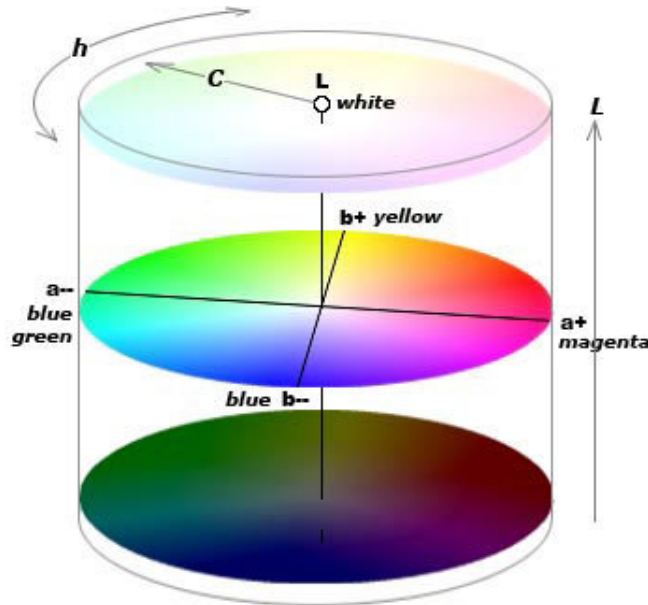


Figure 3.10: CIELAB 1976 $L^* a^* b^*$ is a widely used and validated [32] perceptually uniform colour space. From [55].

3.6.2 Colour Appearance Models

Mark Fairchild in “Color Appearance Models: CIECAM02 and Beyond” [33] page 21 wrote a “*Color Appearance Model provides mathematical formulae to transform physical measurements of the stimulus and viewing environment into correlates of perceptual attributes of colour (e.g., lightness, chroma, hue, etc.)*”.

When testing whether colour tristimuli are perceptually equivalent or different we required a colour appearance model which provides a colour space that:

- is perceptually uniform
- accounts for colour appearance phenomena.

3.6.3 Perceptually Uniform Colour Spaces

Perceptually uniform colour spaces are colour spaces where the distance between colours is equivalent and linear. With this property we can use colour difference measures of distance between two points to establish whether two

or more colour tristimuli are perceptually close or distant. Section 3.6.5 covers the ΔE colour difference measurements we use.

Perceptually uniformity can be achieved through mathematical techniques that perform perceptually valid non-linear transforms of colour spaces before measuring colour differences. Alternatively pre-created colour spaces that are designed to exhibit perceptual uniformity may be used.

For this thesis the perceptually uniform colour space CIELAB 1976 was selected and used (see Figure 3.10). This colour space is widely deployed in research and commercial applications. CIELAB 1976 has been proven to perform well in daylight and near-daylight, even when compared with recently developed colour spaces [32, 53].

CIELAB 1976 is designed to take into consideration chromatic adaption, which is the automatic change in sensitivity of the LMS cones in the eye. CIELAB 1976 also models some of the non-linear multi-dimensional variability in the eye in response to stimuli, such as is due to Steven's Law (see Section 2.5) or the Hunt Effect, e.g. perceived colourfulness is a non-linear function of luminance. Furthermore CIELAB 1976 is designed to account for the interdependencies between lightness, chroma and hue, while also providing a simple colour difference measure [33, 32, 44, 96].

The weak points of CIELAB 1976 include that its colour adaption transform is overly simplistic, e.g. it does not account for temporal adaption for longer exposure time to stimuli. Nor does it consider stimuli context, i.e. background and surround effects. There is also no accounting for Luminance (L^* is Lightness) and related effects, nor are Brightness and Colourfulness variables of the CIELAB 1976 equations (see Equation 3.1).

Equation 3.1: CIELAB 1976 equation for converting from CIE XYZ.

$$\begin{aligned}
 L^* &= 116f\left(\frac{Y}{Y_n}\right) - 16 \\
 a^* &= 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\
 b^* &= 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \\
 C^* &= \sqrt{a^{*2} + b^{*2}} \\
 h^* &= \arctan\left(\frac{b^*}{a^*}\right)
 \end{aligned} \tag{3.1}$$

where

$$f(\omega) = \begin{cases} 7.7871\omega + \frac{16}{116} & \omega \leq 0.008856 \\ \omega^{1/3} & \omega > 0.008856 \end{cases}$$

and

X, Y and Z are CIE XYZ tristimulus values

X_n, Y_n and Z_n are CIE XYZ tristimulus values of reference illuminant

3.6.4 Colour-Appearance Phenomena

As has been noted CIELAB 1976 accounts for aspects of chromatic adaption, which can lead to different percepts for the same LMS tristimuli. Significant causes of variation in the perception of colour can be due to contextual effects, such as environmental lighting. Often when the viewing conditions differ two or more stimuli can have different colour-appearances, though the tristimuli are equivalent. The opposite is also true, tristimuli which are not metamers can appear perceptually equivalent.

Chromatic adaption is one example of a colour-appearance phenomena. Numerous other colour-appearance phenomena exist, such as Simultaneous Contrast, Colour Constancy and Hue Shift [63, 59, 90].

For this work we limit ourselves to CIELAB 1976 for taking account of colour-appearance phenomena. This choice came from the realisation that some of the more recent models, such as the more sophisticated CIECAM02, require the use of manually set variables for calculation, i.e. viewing context, background luminance, surrounding luminance, degree of adaption based on whether self-luminous or illuminated object [33, 32]. Measuring these variables for interfaces and visual designs is not clear cut, and without rigorous validation for HCI could introduce inconsistencies and incorrect results when measuring colour differences.

As it is CIELAB 1976, like most colour-appearance models, was created based on a standard set of stimuli sizes, viewing conditions and standard observers. The implications of this is that more complex stimuli, such as visual designs, are on the edge of what CIELAB and CIECAM were originally designed for, i.e. aiding the measurement of colour differences between large uniform blocks of colour.

When using CIELAB 1976 we set X_n , Y_n and Z_n to the D60 reference illumin-

Table 3.1: Tristimulus values for standard illuminants and observers. Used in Equation 3.1 to account for the colour of light in the environment (Illuminant) and the stimulus size (visual angle) of the area the colour light waves are reflected or emitted from.

Illuminant	Observer	X_n	Y_n	Z_n
D60	2°	96.422	100.0	82.521
D60	10°	96.720	100.0	81.427
D65	2°	95.047	100.0	108.883
D65	10°	94.811	100.0	107.304

ant (see Table 3.1) for a stimulus size of 2 degrees (see Equation 3.1). A 2 degree stimuli implies we expect visual objects to subtend a visual angle of less than 2 degrees. For complex interfaces and visual designs this is a reasonable assumption, as viewing distance from computer displays tends to be greater than 1 foot. At this distance a 12pt letter on a 15 inch screen subtends less than 0.5 degrees.

3.6.5 Measuring Colour Differences

When distinguishing colours we implicitly make colour difference judgements; whether two colours are the same, how they match and how they differ. Over the years a significant focus of empirical research in colour theory has been on developing perceptually valid colour difference measures, such as CIE Delta E 1976 [96] and CIEDE 2000 [54, 76].

Colour spaces and colour difference measures are often inter-related, that is a colour difference measure is designed to work for a particular colour space and the unique properties of that colour space. For example after the CIELAB 1976 colour space was designed, follow on research revealed that the region around the colour blue is especially non-uniform [54], thereby making the Euclidean CIE Delta E 1976 (ΔE_{ab}^*) measure less accurate when making colour difference

measurements about blue and related colours.

For this work we use the CIEDE2000 Delta E (ΔE_{00}^*) colour difference measure [54, 76, 53]. ΔE_{00}^* is a follow on to ΔE_{CMC}^* (Color Measurement Committee for the English Society of Dyers and Colourists) and CIE Delta E 1994 (ΔE_{94}^*). ΔE_{00}^* is designed as an improvement on existing colour difference measurements, especially for dark and near neutral colours while also improving performance for CIELAB blues and grays. Improvements were partially achieved by introducing variables for lightness, chroma and hue - the ratios of which can be set to take into consideration the fact that human vision is more sensitive to changes in chroma and less sensitive to changes in lightness.

A ΔE_{00}^* of less than 1 is usually taken as indicating two tristimuli do not exhibit a perceptual just noticeable difference (ΔV), while a ΔE_{00}^* greater than 6 is often taken to mean two colours are significantly different.

When considering what colour difference measure to use we also explored using the ΔE_{ab}^* and ΔE_{94}^* colour difference measures. Finally we settled on ΔE_{00}^* , as a result this is our colour difference measure of choice in this thesis. ΔE_{00}^* has been shown to be the most accurate colour difference measure to date [54], though it is important to note ΔE_{00}^* is considered one of the more complex colour differences measures to implement correctly [76]. There is also ongoing debate and research into whether and under what circumstances ΔE_{00}^* performs statistically better than ΔE_{94}^* , ΔE_{CMC}^* and ΔE_{ab}^* .

3.6.6 Point / Pixel Measures

When considering the colour difference ΔE_{00}^* measurement, and techniques in this thesis which use the measurement (see Chapter 6 and Equation 6.1), it is important to note that the colour difference measure is point / pixel based. That is when comparing whether two colour patches are the same or different,

only point representations of the colour patches are used, which in the case of computer images correspond to pixels.

Treating colour judgements as point based measures is the standard approach but it does have inherent limitations. Point based measures do not fully account for all the facets of human colour perception. For example, it is known that when one colour fully surrounds another colour, the central colour appears perceptually different depending on the surrounding colour [63, 59, 90]. With point based colour difference measures no account is taken for the effects of surrounding and close by pixel colours.

While it would be desirable to use colour spaces and colour difference equations which are non-point based, the appropriate way of doing so and the validity of the colour difference measures remains an open vision science and colour theory question, e.g. CIECAM02 [33, 60] and S-CIELAB [102].

3.6.7 Generating Retinal Images: Colour Perception

Bringing together a perceptually valid colour model with techniques for simulating colour deficient vision enables the generation of simulated retinal images, which can be analysed to establish perceptual stability.

Figure 3.11 shows the steps in the algorithm for simulating colour perception and generating a perceptually valid colour representation of a visual design:

1. An image of the visual design to be analysed is captured. Often these images are encoded in the RGB colour space.
2. Pick which type of colour deficient vision to simulate, i.e. protanopia or deuteranopia.
3. Simulate colour deficient vision by performing colour space conversions

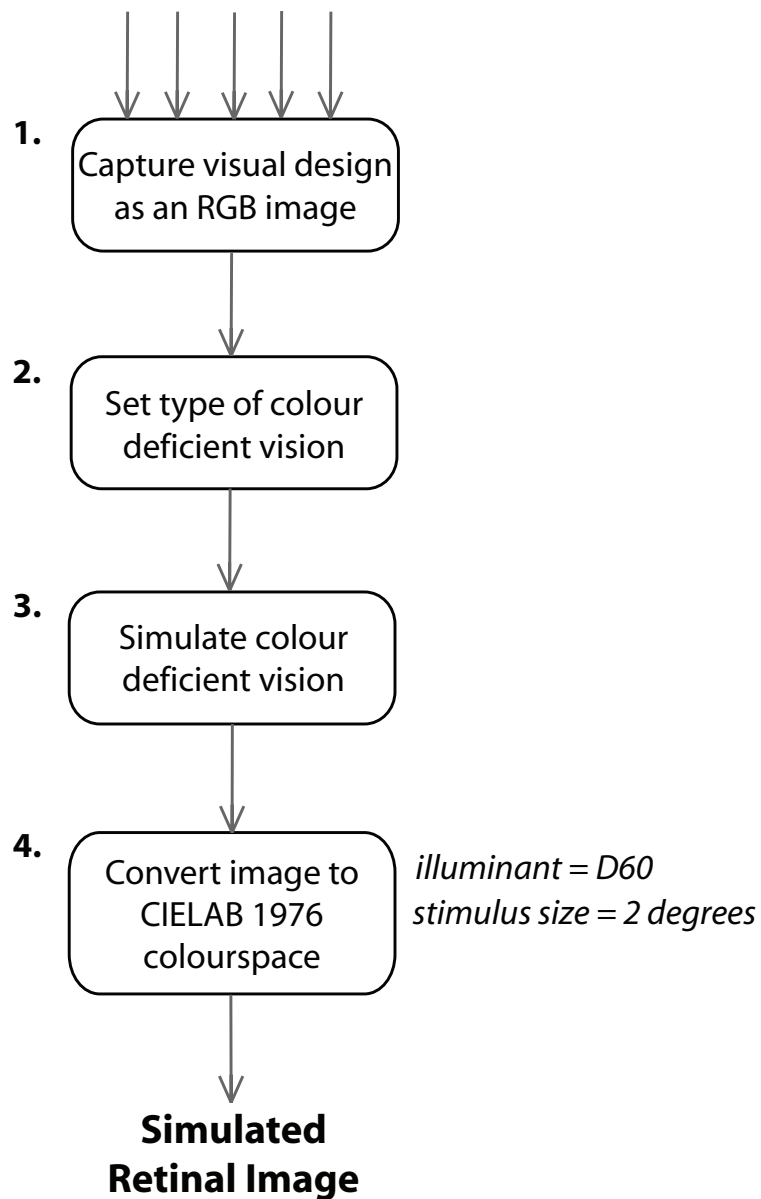


Figure 3.11: Steps in algorithm for simulating colour perception to generate perceptually valid colour representation of a visual design.

of the RGB image. Section 3.7.2 covers the colour space conversion process for colour deficient vision.

4. Transform the colour deficient image's colour space from RGB to CIELAB 1976, with a 2 degree standard observer and the D60 reference illuminant.

In Chapter 6 measures of perceptual stability for colour perception are presented, which are based on the simulated retinal image and the CIELAB ΔE_{00}^* colour difference measure.

3.7 Individual Differences

Adapting the spatial visual acuity and colour perception models to individual differences in eye function is achieved as follows.

For background on individual differences in spatial visual acuity see Section 2.3.1, while for details on individual differences in colour perception see Section 2.3.2.

3.7.1 Visual Acuity

For visual acuity the Zernike Polynomial approach taken to modelling optical aberrations is capable of accounting for individual differences in eye function and structure.

When simulating individual differences in optical aberrations there is a need to understand what Zernike modes are more prevalent, and within Zernike modes what is the standard distribution of Zernike coefficients. By knowing the prevalence and distribution we can construct Zernike wavefronts that are physiologically valid for a normal population of human perceivers. That is,

the constructed wavefronts are not outliers of normal eye function.

With normal eye function, as described in Section 2.4.4.1, it was noted that visual acuity decreases as RMS increases, with the same inverse relationship applying to individual Zernike modes. While in Section 2.4.4 it was noted that Thibos et al., have established that Zernike coefficients tend to be smaller for higher order modes [82], and wavefront variance falls exponentially with n [84]. Further, an analysis of 2560 Shack-Hartmann aberrometry measurements of normal healthy adult eyes, taken at 10 different laboratories [73], found that Z_2^0 and Z_4^0 spherical aberration tend to be positive. Also especially useful is the finding that W_{eye} can be modelled as a *multivariate, Gaussian, random variable with known mean, variance and covariance* [82]. Bringing these findings together enables the creation of simulations of normal individual differences in W_{eye} .

Table 3.2 shows the Zernike modes and Zernike coefficients we use and examine when simulating individual differences in normal visual acuity. As is demonstrated in Chapter 4 and Chapter 5 the benefit of examining multiple Zernike modes is that it enables more rigorous testing of the perceptual stability predictor for visual acuity.

With different Zernike modes we can simulate different types of eyes, with different levels of visual acuity. This enables measuring how accurately the visual acuity predictor makes judgements about the legibility of eye charts. By testing how accurately the visual acuity predictor (see Chapter 4) works on eye charts we establish whether its judgements are in agreement with human vision judgements (see Chapter 5).

3.7.2 Colour Deficient Vision

As has been covered in Section 2.3.2 dyschromatopsia, better known as colour blindness (Colour Deficient Vision - CDV), occurs when a person's perceptual

Table 3.2: Zernike coefficients for Zernike modes in normal population of human observers with 6 mm pupils. Values based on estimates from Figure 1 and Figure 2 in [82], where Min and Max are within ± 2 standard deviations from the mean (95% probability of aberrations falling within Min and Max). Sigma estimated with distribution fitting.

Mode	Min μm	Max μm	Mean μ	Sigma σ
Z_1^{-1}	-1.05	1.04	-0.005	0.603
Z_1^1	-0.73	0.81	0.040	0.445
Z_2^{-2}	-0.43	0.38	-0.025	0.234
Z_2^0	-0.41	1.09	0.340	0.433
Z_2^2	-0.68	0.35	-0.165	0.297
Z_3^{-3}	-0.25	0.22	-0.015	0.136
Z_3^{-1}	-0.36	0.30	-0.030	0.191
Z_3^1	-0.21	0.22	0.005	0.124
Z_3^3	-0.18	0.18	0.000	0.104
Z_4^{-4}	-0.08	0.06	-0.010	0.040
Z_4^{-2}	-0.05	0.07	0.010	0.035
Z_4^0	-0.10	0.29	0.095	0.113
Z_4^2	-0.11	0.10	-0.005	0.061
Z_4^4	-0.10	0.08	-0.010	0.052

experience of colour is significantly different from the normal.

Framing colour deficient vision in terms of N-dimensional colour spaces (see Section 3.6.1) we note that differences in colour perception can be treated as differences in colour spaces. In the case of CDV the CDV observer has a reduced or altered colour space compared to the standard observer's colour space (see Section 2.4.5).

We use the CDV simulation technique outlined in [87, 15] and implemented in Vischeck [30], as this is one of the established approaches for simulating CDV. Though recent research based on molecular genetic analysis in conjunction with colour identification testing suggests that the perception of colour during

CDV is a more complex and dynamic process than captured in existing models [88, 13].

With the ability to simulate different types of CDV, we can test the perceptual stability colour perception measure (see Chapter 6) for a range of different eyes and related colour spaces. By testing how accurately the colour perception predictor works for a range of different eyes we establish whether its judgements are as expected for human vision (see Chapter 7).

3.8 Conclusions

In this chapter we presented models of spatial visual acuity and colour perception, while discussing how the models can be used for generating simulated retinal images. Two algorithms were introduced to generate simulated retinal images for visual acuity and colour perception.

The importance of colour theory and N-dimensional colours spaces for modelling colour perception is highlighted, and the relevance and usefulness of Zernike Polynomials for modelling visual acuity is emphasised. Individual differences in visual function and how they are accounted for in the models when generating the simulated retinal images is also presented.

Visual Stability Predictor: Visual Acuity

In this Chapter we introduce the visual acuity perceptual stability measure $PERS_{va}$. With $PERS_{va}$ we can evaluate a visual design to establish what parts of the design are easy or difficult to see.

Also introduced is the STRIVE (SimulaTed Retinal Images with Visual Entropy) algorithm. STRIVE generates entropic measures of changes in a visual design, due to optical aberrations in the eyes perceiving a design. $PERS_{va}$ is built upon STRIVE.

4.1 Measuring Perceptual Stability

Chapter 3 describes the models developed for generating the simulated retinal images, which depict the effects individual differences in spatial visual acuity and colour perception have on visual designs.

The following sections cover how the simulated retinal images and related models are used for creating predictors of perceptual stability. Following this are experiments to test whether the predictors make predictions consistent

with human performance.

As previously discussed in Section 1.4.2 and Section 3.2.1 the core functionality requirement for the predictors is that they provide a measure of how stable or unstable a visual design or image is due to differences in a perceiver's (spatial or colour) perception.

We consider two facets of perceptual stability:

- Degree of perceptual change (PERS), i.e. how stable or unstable a visual design is.
- Location of perceptual change (LocPERS), i.e. what areas in a visual design are least or most stable.

PERS enables us to measure of number of interrelated things. Firstly, compare the perceptual stability of a range of potential designs and make a decision about which is best, i.e. which is most perceptually stable. Secondly, predict the impact of changes in perceivers eye function on how easy or difficult something is to see. An example of this is normal age related differences in visual acuity, which lead to differences in the experience of seeing a design. The further away you are the less visual detail seen, and the older you are the less visual detail seen.

Building on the PERS predictor we develop a second metric to identify the location of changes (LocPERS). The goal of this measure is to help us understand which parts of a design are most or least perceptually stable. By knowing this a designer or automated graphical layout technique can decide to put lower priority information in locations that may be harder to see. Another benefit is that by testing the perceptual stability of each sub-part of a design we can ensure that all parts of the design maintain or achieve certain minimum perceptual stabilities. Taking a single measure of perceptual stability for an entire design

is informative but it may miss variations in the local minima's and maxima's in the perceptual measure.

4.1.1 Predictions As Constraints: How Visually Rich Can A Design Be?

When displaying a visual design on a graphical display it is trivial to establish how many physical pixels are available and the range of colours each pixel can display. This provides the upper bound on how much physical information can be statically displayed on-screen.

In this instance the use of the term information is referring to the maximum number of bits displayable on-screen at a single moment of time. This notion applies equally to data visualisation.

We do not know exactly how small something should be for us to be able to see it. This perspective is simplified because it is not taking into consideration the perceptual effects of super-acuities and temporal integration.

In addition it is unknown what is the upper bound on how much information can be displayed on-screen because of eye function. There is an interplay of eye function with display device limits to create a different, and potentially dynamic, upper bound on display "information" density.

An ability to distinguish between information content on-screen at a purely numerical level and the information content on-screen due to perceptual effects can lead to a perceptual information theory for visual displays.

4.1.2 Correlating With Human Performance

A prime requirement of the measures is that they make measurements consistent with humans. For example a person viewing a design with 20/20 (stand-

Equation 4.1: Non-Normalised Shannon Entropy (NNSE).

$$H(X)_{NonNorm} = - \sum_{i=1}^n s_i^2 \log s_i^2 \quad (4.1)$$

where

s_i is pixel colour value adjusted to range 0.0 to 1.0

and

$$0 \log 0 = 0$$

ard) vision is able to state which part of an interface is easier to see than another, and we can expect their observations to be consistent with other normal observers. For example if the font used for text is smaller in one area than another the smaller text will be harder to read; presuming text contrast and colour is equivalent.

Measuring the degree to which one set of visual features is potentially harder to see than another set is clearly desirable. A magnitude measure enables a user to develop a sense of how much more difficult one part of a design is to see than another. In addition it enables an automated layout / design technique to explore a set of potential designs, and then rank the potential designs along a scale from easiest to most difficult to see (see Chapter 8).

4.2 Stability Measure Of Spatial Visual Acuity

Spatial visual acuity is the smallest spatial detail that can be visually detected, discriminated, or identified [61]. As differences in the perception of designs occur we want to establish whether designs are perceptually changed. That is, whether the visual designs perceptually appear to lose visual detail as visual acuity decreases.

Our method for measuring the degree of change in a visual design is based on Non-Normalised Shannon Entropy (Equation 4.1), which is a measure of information structure.

4.2.1 Perceptual Measure

When considering the eye and HVS as a communication system we hypothesize that elements of Shannon's Information Theory [75] can form the basis of a potential measure.

In related fields, information theory for sensory coding has been researched and applied to vision modelling and statistical image analysis [3]. We are taking a similar approach, in addition we wish to avoid approaches that involve feature specific analysis (see Section 3.4.1). Such approaches have the potential for undue reliance on image structure, e.g. zero-crossing / edge width of a line increasing due to increasing optical aberrations (see Figure 4.2).

4.2.2 Perceptual Stability Algorithm: STRIVE

The goal of the STRIVE perceptual stability algorithm is to measure whether the entropy of information changes, due to optical aberrations. As the aberrations increase or decrease in a visual design, due to optical aberrations and individual differences, we examine whether the entropy of the information in the images changes in a consistent, predictable and useful manner. Recall we hypothesize that the entropy of a visual design changes predictably with changes in optical aberrations. From [5] we find a justification for this assumption, i.e. the continuous entropy of a wavefront PSF increases with increases in wavefront aberrations.

Depicted in Figure 4.1 is the resultant algorithm for generating the changes in entropy due to optical aberrations. For the purpose of this thesis the algorithm

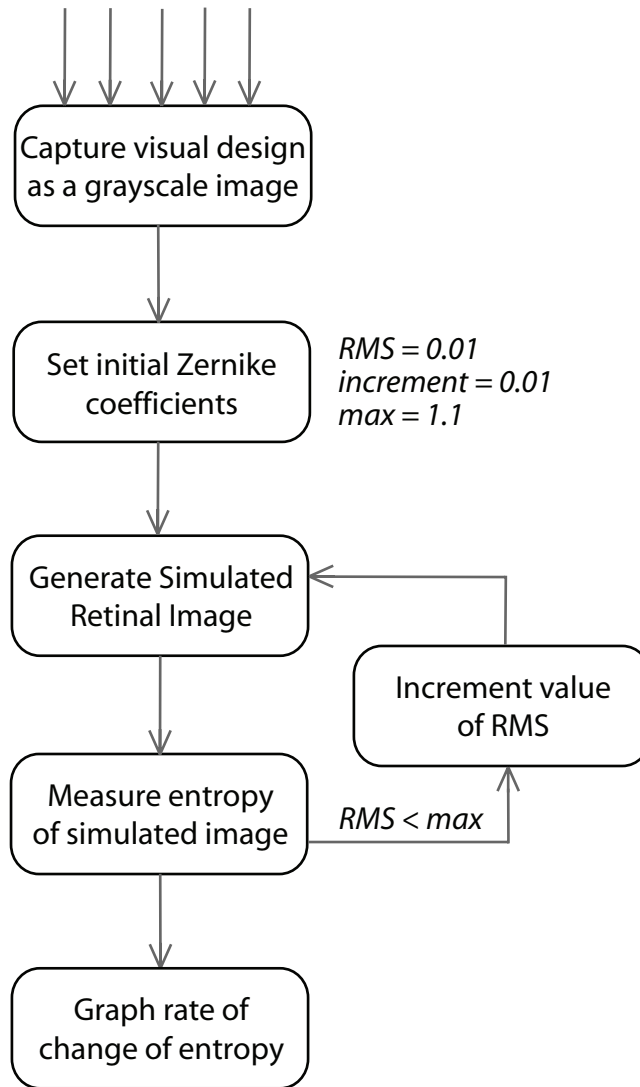


Figure 4.1: Steps in STRIVE (SimulaTed Retinal Images with Visual Entropy) algorithm to generate entropic measure of change due to optical aberrations in a visual design.

is referred to as STRIVE (SimulaTed Retinal Image with Visual Entropy). The steps within STRIVE are:

1. Capture grayscale image of visual design to analyse.
2. Initial Zernike coefficients for contributing Zernike modes are set so there are no optical aberrations. For rest of algorithm we use the RMS (see Equation 2.3) of the Zernike coefficients.
3. Simulated retinal images are generated, based on algorithm shown in



Figure 4.2: Simulated retinal images of a vertical bar. 1st row shows the effect of Z_2^0 from 0.0 to 1.09 RMS, 2nd row shows the effect of Z_3^{-1} from 0.0 to 0.36 RMS, and 3rd rows shows the effect of Z_4^4 from 0.0 to 0.11 RMS. Figure 4.5 shows related changes in Non-Normalised Shannon Entropy.

Figure 3.8 and discussed in Section 3.5.4.

4. An entropy measure of the simulated retinal image is made. In Section 4.2.3 the basis for the measure is elaborated upon.
5. To establish perceptual stability the entropy measure as a function of RMS increasing is graphed and analysed. This is elaborated on in Section 4.2.4.

4.2.3 Measuring Visual Change With Shannon Entropy

Our initial measure [10] was based on Shannon's measure of entropy as given in Equation 4.2, where the count of occurrences of each unique pixel colour is a discrete symbol x_i . The change of entropy as a function of changes in optical

Equation 4.2: Shannon's measure of Entropy.

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i) \quad (4.2)$$

where

i refers to a pixel
 n is total number of pixels
 x_i is count of occurrences of a pixel colour
 p is probability of x_i occurring out of all pixels n

aberrations was used for analysis, see Equation 4.3.

Using Equation 4.3 to analyse the effects of optical aberrations on natural images gave what seemed meaningful predictions. However problems arose when using it to analyse images of interfaces, due to the general structure of the images. Interfaces are often sparse images, e.g. Figure 4.3 Region 1 Original Image, while natural images tend to be complex, e.g. Figure 4.3 Region 2 Original Image. We refer to images as sparse where the count of unique colours is low in the image, while complex images are where the number of unique colours is large.

With complex images the entropy tends to decrease due to increasing optical aberrations. However, with sparse images the entropy increases as the degree of aberrations increases. Eventually, the entropy decreases but at a point which depends on the starting image. Particularly ambiguous are sparse interface images that are close to complex, as aberrations can quickly turn them into

Equation 4.3: Erroneous entropy predictor of perceptual stability.

$$PERS_{err} = \frac{d(H(X))}{d(RMS)} \quad (4.3)$$

where

RMS is root mean squared for Zernike mode

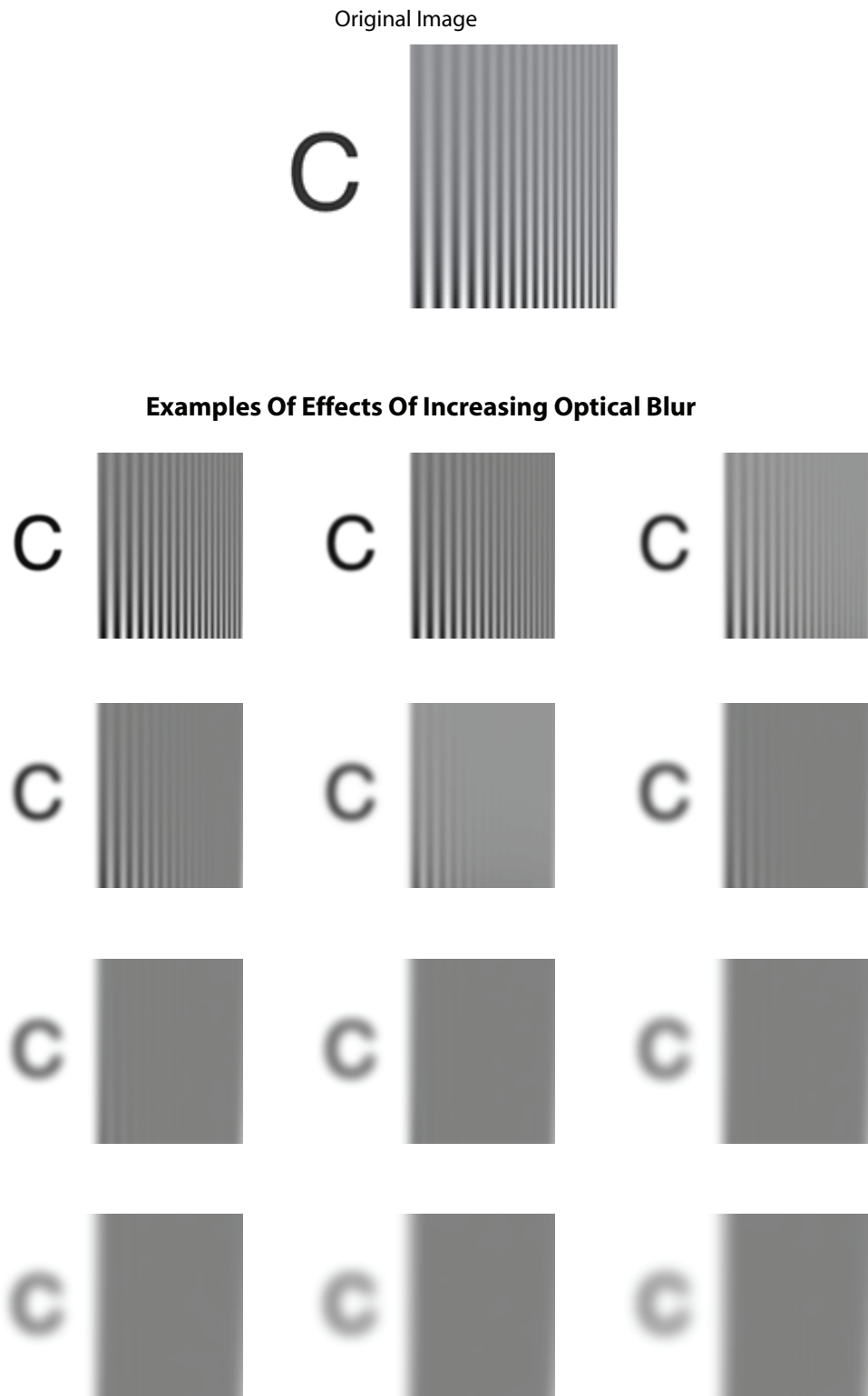


Figure 4.3: Original image is a simplified combined visual acuity eye chart. On the left is the Sloan letter C, commonly used to measure visual acuity with optotypes, and on the right is an area from a CSF eye chart. Examples of effects of increasing optical blur on eye chart are shown, with blur increasing from upper left to bottom right.

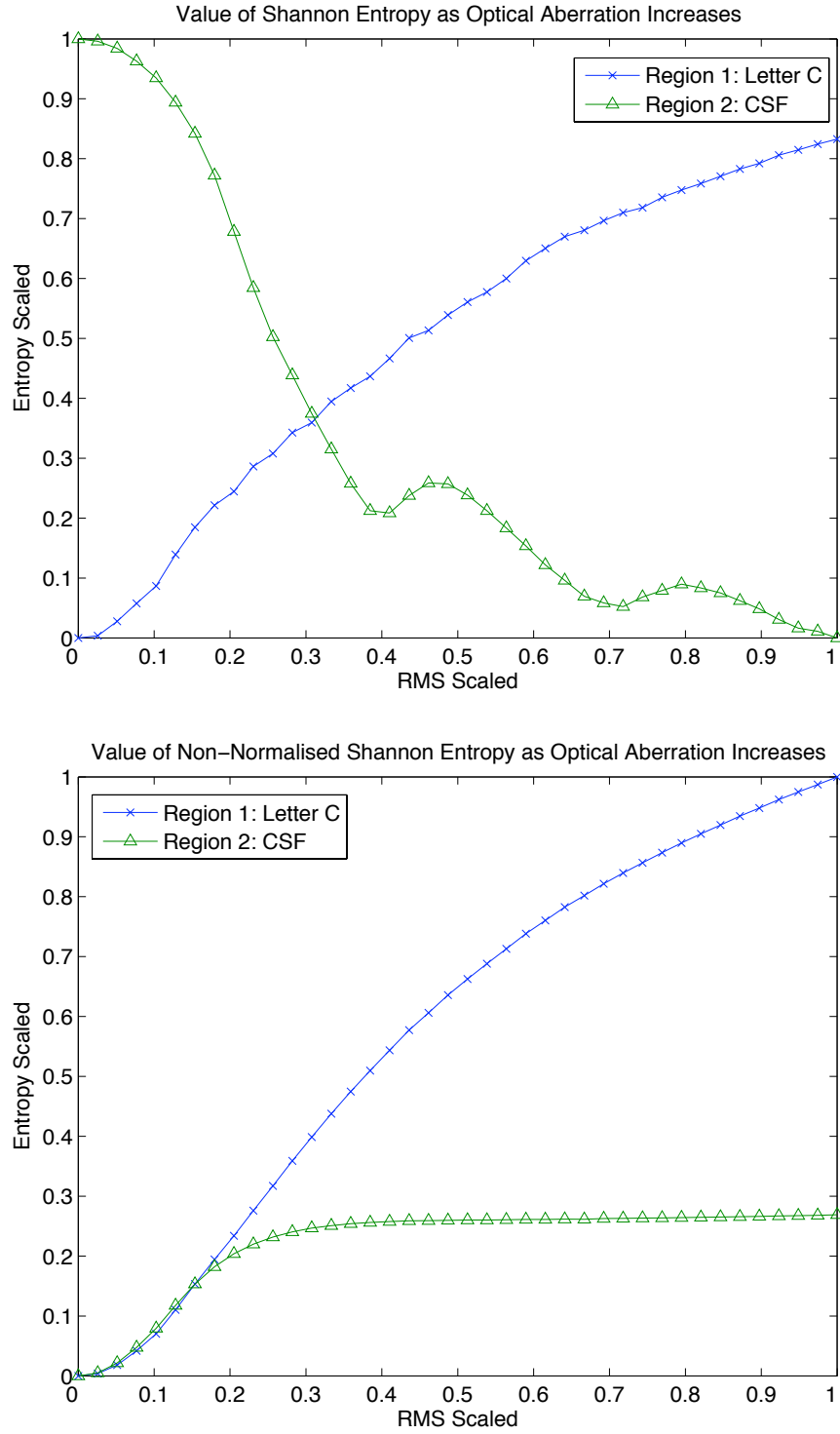


Figure 4.4: Two graphs showing how increases in RMS for optical aberration defocus Z_2^0 effect measures of Shannon Entropy and Non-Normalised Shannon Entropy. Entropy measures were generated using the algorithm STRIVE shown in Figure 4.1. RMS went to a maximum of 1.09 for a pupil of 6 mm with wavelength set to 555 nm. Original Image in Figure 4.3 is used to generate the simulated retinal images, and is divided in half into left (Region 1) and right (Region 2) regions.

complex images. The direction of change of entropy over multiple aberrations with sparse images is image content specific, which makes it challenging to use for comparisons of the effect of aberrations on potential interface designs and information visualisations.

An example of the behaviour of Shannon Entropy with increasing optical aberrations can be seen with the combined eye chart shown in Figure 4.3, Original Image. On the left hand side of the chart is the Sloan font letter C [66], while on the right is an area from a CSF eye chart [61]. When the eye chart is used in the STRIVE algorithm with Shannon Entropy we can see how the entropy changes in the top graph of Figure 4.4. For the letter C the Shannon Entropy increases, while the Shannon Entropy of the CSF region decreases.

After examining why the Shannon Entropy in sparse images, e.g. letter C, was increasing we found one of the contributing factors is because there is an increase in the number of colours. That is, the entropy changes due to a change in the number of unique colours (x_i) in an image as well as a change in the distribution $p(x_i)$ of the colours. For example in Figure 4.2 you can see a black line that undergoes increasing amounts of optical aberrations from the left to the right, with a corresponding increase in the number of shades of gray in the line.

This also applies to complex images, though it is less obvious. The decreased entropy of complex images is influenced by a reduction in the number of colours as well as the distribution of the colours.

Shannon Entropy is overly sensitive to the change in colours in an image as optical aberrations increase. When an image has lost a lot of visual detail Shannon Entropy measures can still change quickly due to increases or decreases in the number of colours, even when visual structure is changing slowly.

More details on our initial use of Shannon Entropy, along with a normalised

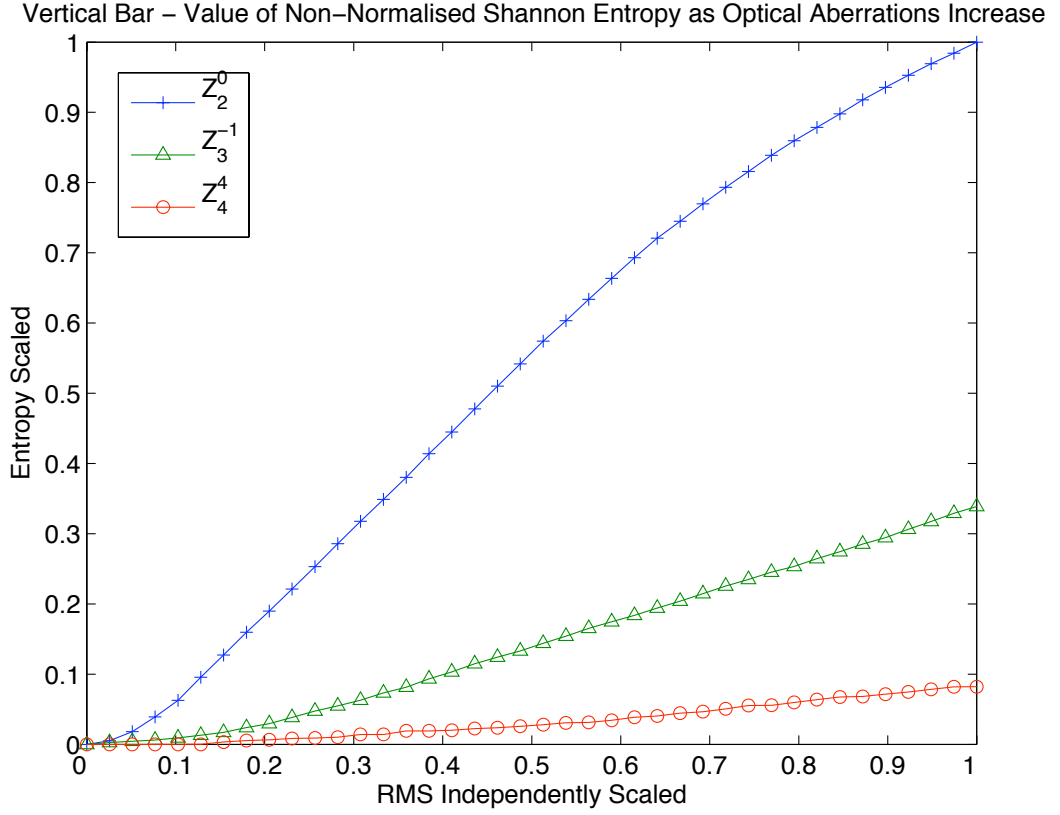


Figure 4.5: Shown is the increase in Non-Normalised Shannon Entropy of the simulated retinal images with increasing RMS values for optical aberrations Z_2^0 , Z_3^{-1} and Z_4^4 . Pre-scaled RMS values were selected from Table 3.2. Simulated retinal images are shown in Figure 4.2.

variant of Shannon Entropy to control for increases or decreases in the number of colours can be found in [10]. In that earlier work Gaussian 2D functions were used to generate the aberration PSFs.

4.2.4 Measuring Visual Change With Non-Normalised Shannon Entropy

Finding that measures of Shannon Entropy as a function of RMS are overly sensitive to colour occurrence lead us to examine other entropy measures, which could be used as building blocks for the perceptual stability measure.

Equation 4.1 is Non-Normalised Shannon Entropy (NNSE), which is the en-

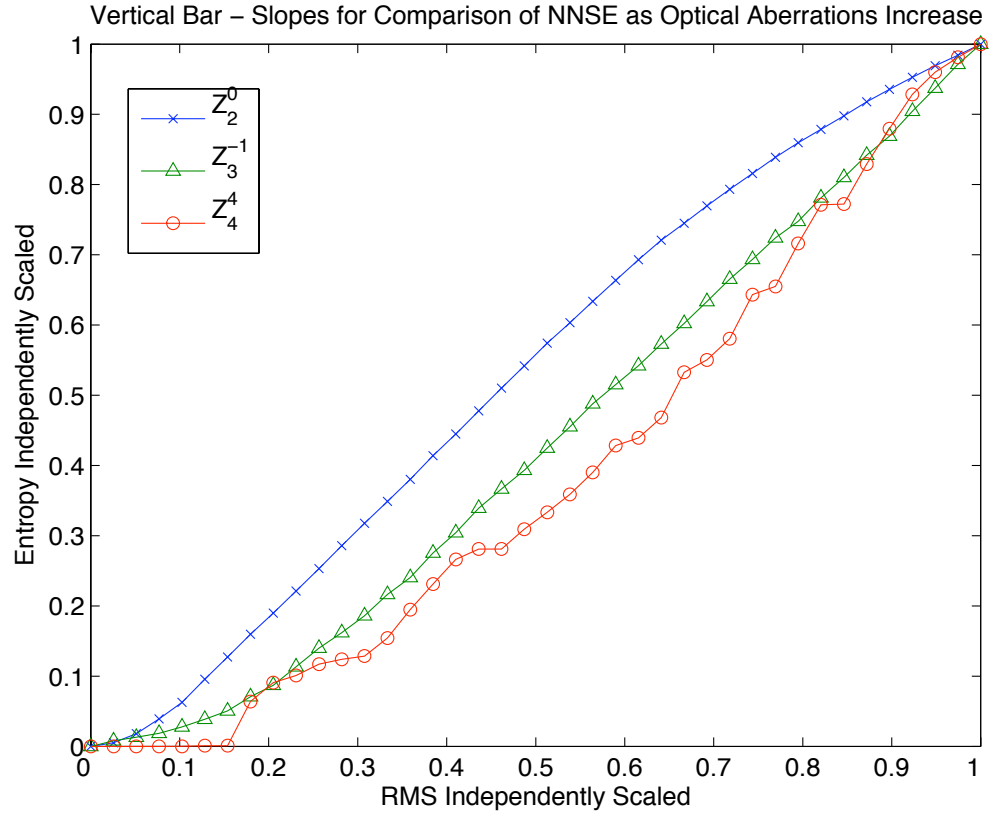


Figure 4.6: Each STRIVE curve set of entropy values are independently normalised to range from 0.0 to 1.0. Shown is the increase in Non-Normalised Shannon Entropy of the simulated retinal images with increasing RMS values for optical aberrations Z_2^0 , Z_3^{-1} and Z_4^4 . Pre-scaled RMS values were selected from Table 3.2. Simulated retinal images are shown in Figure 4.2.

trophy measure that we found to be effective for the perceptual stability measure. Shown in the bottom graph of Figure 4.4 is the result when the combined Sloan C and CSF eye chart (see Figure 4.2) are subjected to the STRIVE algorithm using Non-Normalised Shannon Entropy for step 4. In Figure 4.4 the bottom graph shows how Region 1 continues to change for longer, that is as the degree of optical aberration increases the letter C continues to visually change, while Region 2 initially changes for the first 3rd of increases in RMS and then changes very little with continued increases in optical aberrations. Clearly, Region 2 loses a lot of visual detail when subjected to a lower amount of optical aberrations than Region 1.

From the graph in Figure 4.5 we see that STRIVE with NNSE can also be used to make predictions with regards to comparisons between different Zernike modes. For the simulated retinal images of the vertical bar shown in Figure 4.2 and the maximum RMS values based on Zernike coefficients from Table 3.2; the graph indicates that the vertical bar does sustain continuous change as the RMS values and related optical aberrations increase to maximum.

Furthermore Zernike mode Z_2^0 has a steeper slope than Z_3^{-1} , which has a steeper slope than Z_4^4 (see Figure 4.6). The steepness of the slopes indicate the amount of change the vertical bar undergoes with increases in RMS across different Zernike modes. From the simulated retinal images in Figure 4.2 we can see that slopes do correspond with degree of visual change.

Features in a visual design are more perceptually stable when the design continues changing with increasing amounts of aberrations. Looking at Figure 4.6 we see that Z_4^4 changes less than Z_3^{-1} , which changes less than Z_2^0 . With these three Zernike modes the change is continuous but the rate of change distinguishes the perceptually stability. Z_4^4 is most perceptually stable, followed by Z_3^{-1} and finally Z_2^0 .

A visual design which is capable of sustaining continuous increases in optical aberrations without losing all visual detail is more perceptually stable, i.e. increases in RMS lead to increases in NNSE when a visual design is perceptually stable. Framing it slightly differently, a visual design is more robust when it is capable of experiencing high amounts of optical aberrations without losing all visual details, where NNSE is functioning as a measure of visual detail. For example in Figure 4.7 we can see that the STRIVE curve for Region 2 is initially very steep then flattens, while Region 1 has a continuous change in NNSE over the full range of RMS. Region 1 is not as perceptually stable as Region 2.

In the experiments in Chapter 5 these observations and initial results are experimentally tested and validated, by testing the performance of STRIVE with

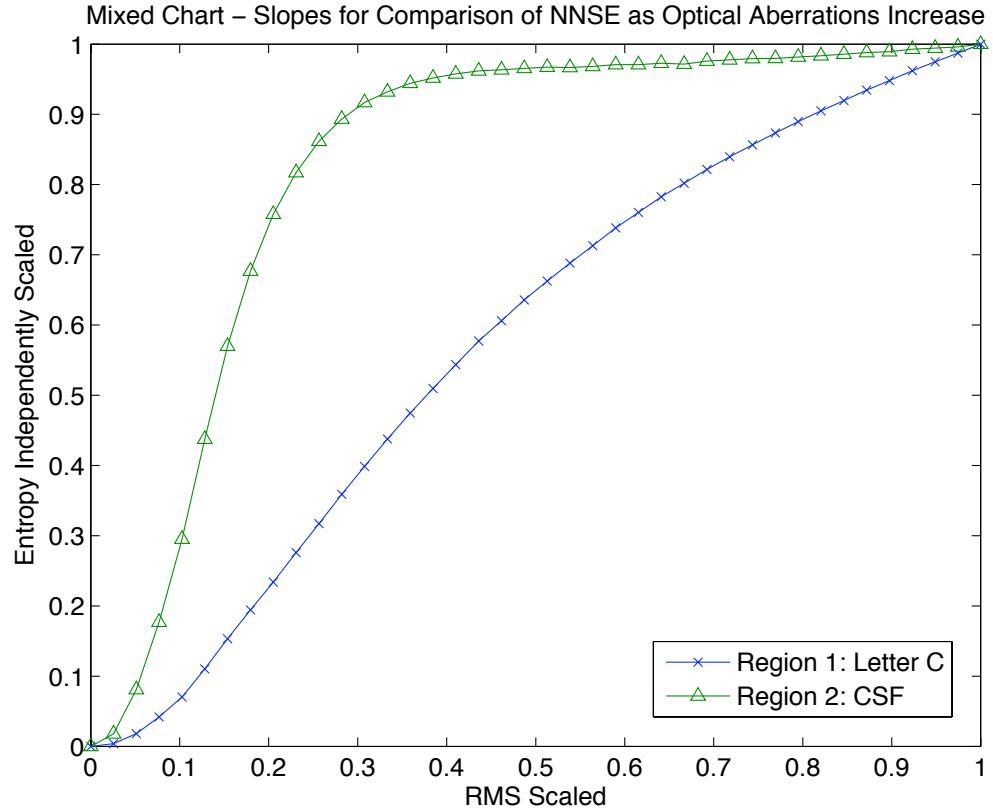


Figure 4.7: Simplified combined eye chart (see Figure 4.3) STRIVE curve slopes independently normalised to range from 0.0 to 1.0. Maximum RMS value is 1.09 for Zernike mode Z_2^0 .

NNSE on a range of eye charts.

4.2.5 Calculating Perceptual Stability: Visual Acuity

The output of the STRIVE algorithm is a series of data points each of which is an NNSE measurement. Together the data points describe a curve, which is referred to as the STRIVE curve. Estimating perceptual stability is achieved by analysing the STRIVE curves.

In this work we analyse the STRIVE curves by measuring the size of the area under the STRIVE curve. Area under the STRIVE curve is used because of the interpretations of the shape of the curve as listed in Table 4.1. A number of these interpretations were identified in Section 4.2.4 and are experimentally

Table 4.1: STRIVE Curve Interpretations

Slope	Interpretation
Steep	A steeper sloped STRIVE curve means more visual changes are occurring due to increased optical aberrations caused by increments in RMS values.
Flat	A flatter STRIVE curve means little to no more significant visual changes are occurring.
Smooth	A STRIVE curve which continues smoothly changing over the full range of RMS values means a design is visually robust, in the sense that it does not lose all visual features straight away.
Rough	A roughly changing STRIVE curve means a visual design loses different amounts of visual detail at different rates, depending on the range of RMS, i.e. as optical aberrations increase the visual design does not lose all visual detail at the same rate. An important caveat is where the NNSE range is very small, which is dealt with by the Noise Threshold.

validated in Chapter 5.

Data fitting of the i NNSE data points is used to generate an n th degree polynomial S_i , which describes the STRIVE curve. Before data fitting the maximum value of i is found, if this is less than the Noise Threshold (NT) all the values of i in NNSE are set to 0. Then the NNSE data points are normalised to range between 0.0 and 1.0, which makes different STRIVE curves comparable because often the NNSE minimum and maximum differ between visual designs.

A Noise Threshold is used because NNSE values with a tiny range create misleadingly steep and bumpy STRIVE curves, which do not reflect visual change due to optical aberrations. An example of when the Noise Threshold is important is where analysing a very slight visual gradient. Optical aberrations do cause very slight changes in such a gradient, with resulting NNSE values that when normalised between 0.0 and 1.0 generated a very rough and bumpy STRIVE curve, i.e. the STRIVE curve can change radically between RMS in-

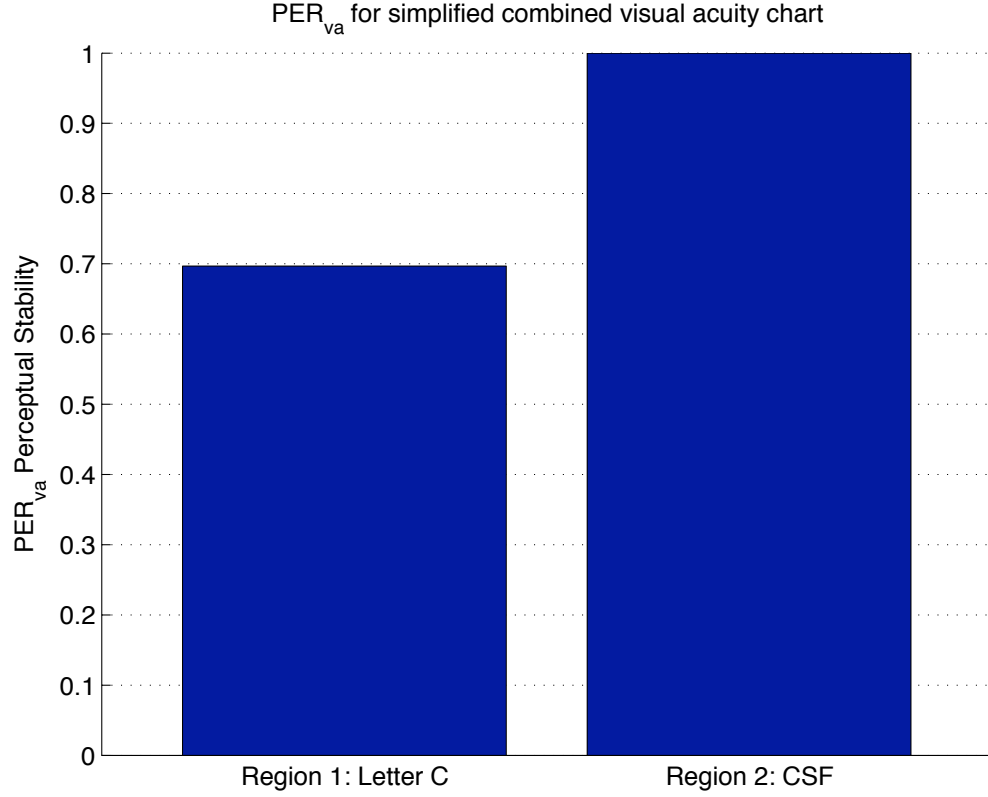


Figure 4.8: $PER S_{va}$ perceptual stability score for Figure 4.3. A lower score is better. Slopes of STRIVE curves shown in Figure 4.7, and range of NNSE values graphed in bottom graph of Figure 4.4. Perceptual stability scores normalised to maximum value of 1. Region 1: Letter C is more perceptually stable than Region 2: CSF, when Z_2^0 is analysed for an RMS of 1.09.

creases because of a tiny range, rather than fundamental visual changes.

An integral of the S_i polynomial is used to measure the area under a STRIVE curve (Equation 4.4). The area under a curve is $PER S_{va}$, the perceptual stability measure for visual acuity measured by normalised area.

For example the bar chart in Figure 4.8 shows the perceptual stability of the simplified combined visual acuity eye chart shown in Figure 4.3. Region 1: Letter C is more perceptually stable than Region 2: CSF, that is Region 1 can undergo more optical aberrations than Region 2.

A benefit of having a polynomial is it enables sensitivity analysis. By using different RMS ranges with the polynomials the sensitivity of a visual design to

Equation 4.4: Generate $PERS_{va}$ (perceptual stability for visual acuity) score by measuring the normalised area under a STRIVE curve.

$$PERS_{va} = \begin{cases} A_n & \max\{NNSE_0^i\} \geq NT \\ 0 & \max\{NNSE_0^i\} < NT \end{cases} \quad (4.4)$$

where

$$A_n = \int_0^{RMS} S_i$$

and

S_i is polynomial generated by data fitting i NNSE data points
 NNSE data points are normalised to range 0.0 to 1.0
 NT is Noise Threshold (see Section 4.2.5)

optical aberrations can be examined in detail, without requiring regeneration of multiple simulated retinal images and Zernike polynomials.

4.3 Conclusions

In this chapter we have introduced STRIVE and $PERS_{va}$ for measuring the perceptual stability of visual designs due differences in visual acuity. Also compared is the use of Non-Normalised Shannon Entropy versus Shannon Entropy as the basis for measuring visual change.

How these techniques work is demonstrated and explained, while the implications of and motivations for the Noise Threshold is discussed. With $PERS_{va}$ we can evaluate a visual design to establish what parts of the design are easy or difficult to see.

Experimental Evaluation Of Visual Acuity Predictor

In this Chapter *STRIVE* and $PERS_{va}$ are experimentally tested and validated on eye charts, which demonstrates that they function as an effective predictor of perceptual stability for visual acuity.

5.1 Experiments: Visual Acuity Stability

The criteria for success of the visual acuity predictions is that they should successfully and consistently make predictions that correlate with human performance. As mentioned in Section 4.1 of Chapter 4 the visual acuity predictor should be capable of making predictions about:

- Degree of perceptual stability (*PERS*), i.e. how stable or unstable a visual design is.
- Location of perceptual stability (*LocPERS*), i.e. what areas in a visual design are least or most stable.

When considering what experiments could be created to test the visual acuity predictor there is the option of creating a range of interfaces and information visualisations, then testing how well the predictor functions on them. While this approach may initially seem valid, from a vision science perspective it would be empirically very weak. Weakness arises because the design of the information visualisations and interfaces used in an experiment could easily and non-obviously radically impact the experimental outcomes.

5.2 Eye Charts As A Gold Standard Of Visual Function

Using information visualisations and interfaces for the experiments would be valid only where the perceptual characteristics of the visual designs are extremely well studied and the perceptual interactions of visual components are clearly understood and established. With eye charts there have been numerous studies of human performance which show subtle but important perceptual interactions between optotype layout, style, design, contrast, and other visual features. These results have led to a steady evolution of eye charts, such that modern eye charts are empirically known to provide accurate predictions of visual acuity and can consistently produce repeatable measures of eye function.

Within vision science, especially optometry, there is a large body of research into human performance on eye charts. For the experiments we tested the performance of the visual acuity predictor on eye charts. By doing so we were testing whether the visual acuity predictor made predictions consistent with a large body of well established human performance results.

Studies have also experimentally demonstrated that there is a correspondence

between measures of a person's visual acuity and their ability to perform life tasks [93, 71]. As a person's visual acuity decreases there is a steady decrease in their ability to perform everyday tasks [92].

Measurements of spatial contrast sensitivity are also an important predictor of a person's ability to see visual detail. Individual differences in spatial contrast sensitivity have been shown to affect people's capabilities in everyday tasks, such as seeing faces, road signs and other commonplace objects [62, 93].

5.3 ETDRS, Pelli-Robson & Campbell-Robson Eye Charts

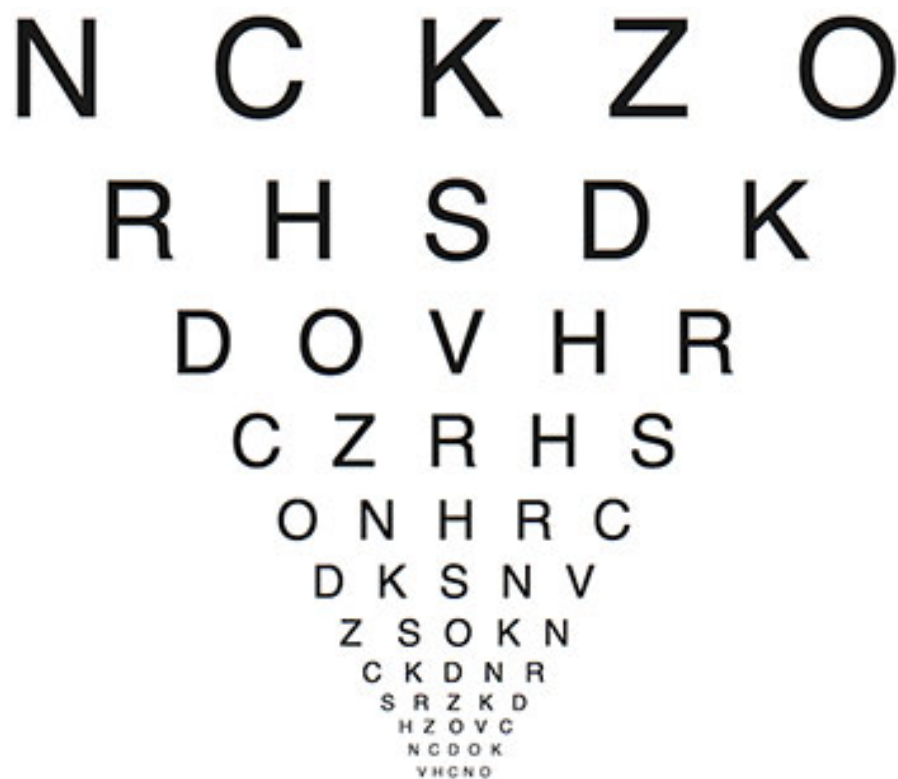


Figure 5.1: Example of ETDRS Eye Chart.

In our experiments we used three different kinds of eye charts. The ETDRS eye chart (see Figure 5.1), the Pelli-Robson eye chart (see Figure 5.3) and the

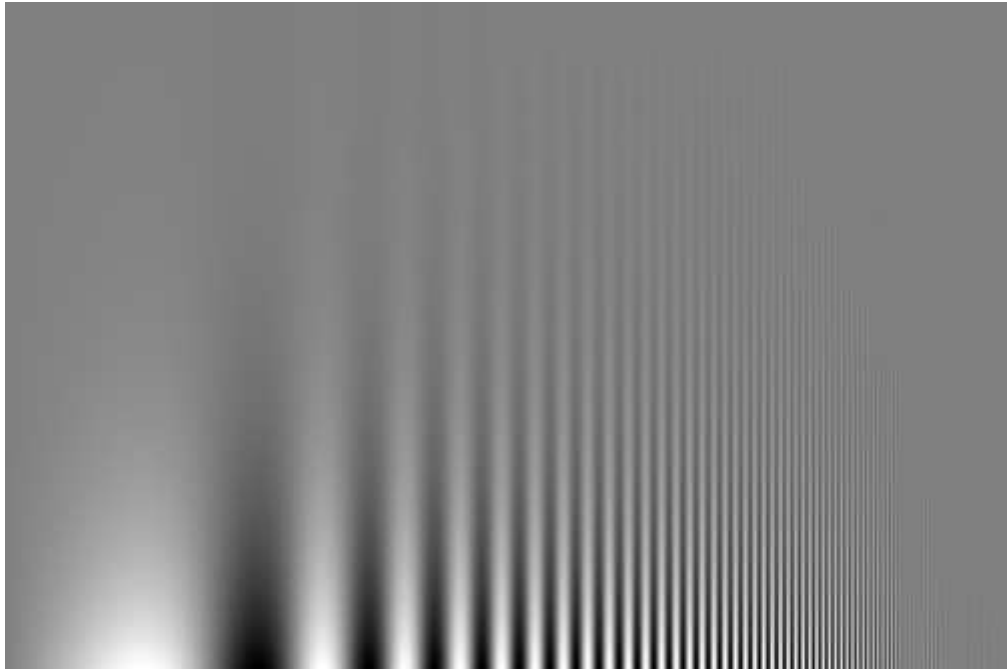


Figure 5.2: Example of Campbell-Robson Contrast Sensitivity Eye Chart.

Campbell-Robson Contrast Sensitivity eye chart (see Figure 5.2). All these charts are used to measure facets of human visual acuity and have been designed and tested by the optometry and vision science research communities [61, 63].

Each chart has very different visual characteristics. By using such visually different charts we are establishing whether the visual acuity predictor is robust. Robust in the sense that it is capable of making accurate predictions on different kinds of visual designs.

5.3.1 ETDRS Eye Chart

Optometrists commonly measure visual acuity by taking psychophysical measures of a person's ability to identify and discriminate optotypes (see Figure 5.1 and Figure 5.3). Optotypes, such as letters, are presented on eye charts, such as the Snellen, Landolt C, Bailey-Lovie [4] and ETDRS charts [34]. Optotype identification and mis-identification is used as a fundamental measure



Figure 5.3: Example of Pelli-Robson Contrast Sensitivity Eye Chart [66].

of a person's ability to visually perceive and recognize letters, shapes, patterns and much more complex sights in their visual environments.

The suitability of optotype eye charts for measuring and predicting the effects of visual acuity has been extensively studied with a large body of research literature dating back more than one hundred years. During that time period subtle methodological discrepancies with measures obtained using the charts have been experimentally proven [4] and rectified [34]. Discrepancies demonstrated include problems with repeatedly obtaining the same visual acuity measurement when a person's visual acuity is measured multiple times on different dates.

Other problems include unexpectedly strong interactions between optotypes due to optotype shapes and distances between the optotypes. Optotype interactions can make visual acuity judgement comparisons between optotypes and lines of optotypes non-linear.

With ETDRS eye charts people's ability to visually identify optotypes decreases from good at the top of the chart to bad at the bottom of the chart, in increments of rows.

5.3.2 Pelli-Robson Eye Chart

When eye charts are used to test a person's visual acuity the contrast between the optotypes and the eye chart background is extreme. Usually optotypes are black and the background is white. Over the years it has been demonstrated that contrast sensitivity plays a very important part in people's abilities to resolve visual detail [19, 36].

Contrast sensitivity can be tested and measured by using a Pelli-Robson chart (see Figure 5.3). In a Pelli-Robson chart [66] the contrast between the optotypes and the background decreases as a function of optotype distance from the top left of the chart. Optotypes are in contrast groups of 3, which means in Figure 5.3 the top left letters V R S share the same contrast, followed with a decreased contrast the next optotype group K D R.

People's ability to discern detail in the Pelli-Robson chart also functions as distance from the top left of the chart, with a granularity at the level of individual optotype groups.

5.3.3 Campbell-Robson Contrast Sensitivity (CRCS) Eye Chart

A CRCS chart (see Figure 5.2) consists of a grey scale sinusoidal grating where the sine wave frequency increases continuously from left to right and the contrast increases logarithmically from top to bottom. CRCS charts are used to measure a person's Contrast Sensitivity Function (CSF) [19].

In CRCS charts an individual's visual acuity is plotted as an inverted U. The area within the inverted U represents what a person can see when viewing the chart at a fixed distance. The CSF changes position as a person's visual acuity changes, that is the inverted U moves to the right as a person's ability to see the Campbell-Robson chart decreases and to the left as a person's ability to see the chart increases. The area under the CSF curve represents the area of the image a person can see reasonably clearly. For further details on CSFs consult Norton et al [61].

5.4 Virtual Eyes

To test STRIVE and the related $PERS_{va}$ the Zernike coefficients in Table 3.2 are used to generate 3700 random normal virtual eyes. Each of the virtual eyes is then applied in step 2 of the STRIVE algorithm.

For the virtual eyes the Zernike coefficients are kept within 2 standard deviations of the mean for a Gaussian distribution. There are indications that the values of certain Zernike coefficients correlate [84]. Inter-Zernike mode correlations are not maintained for the virtual eye generator as the underlying data used to generate the correlations was unavailable.

The Zernike coefficients listed in Table 3.2 are based on an analysis of 200 normal eyes. The virtual eyes generated are of pupils dilated to 6 mm, for

monochromatic light at 555 nm. As more research emerges on normal Zernikes coefficients for different pupil sizes over different age groups the Zernike coefficients in the virtual eye generator can be refined [73, 77, 17].

Between virtual eyes the RMS values differ based on the underlying Zernike coefficients. One virtual eye can have a low RMS, while another virtual eye can have a large RMS.

5.5 Experiments

The purpose of the experiments is to test whether $PERS_{va}$ with STRIVE generates results in agreement with human judgements for eye charts.

Here we describe four experiments broken into twelve sub-experiments where the STRIVE algorithm and $PERS_{va}$ are tested on a range of visually distinct and perceptually different eye charts. The experiments establish that STRIVE and $PERS_{va}$ successfully measures the degree of perceptual change (PERS), while also identifying the locations of perceptual change (LocPERS).

5.5.1 Overall Approach & Details

5.5.1.1 Approach

The overall structure of the experiments uses multiple virtual eyes to generate multiple simulated retinal images. Next the STRIVE algorithm and $PERS_{va}$ evaluate the perceptual stability of each simulated retinal image. Finally the summed results of $PERS_{va}$ for all virtual eyes are compared with the expected outcome for each eye chart.

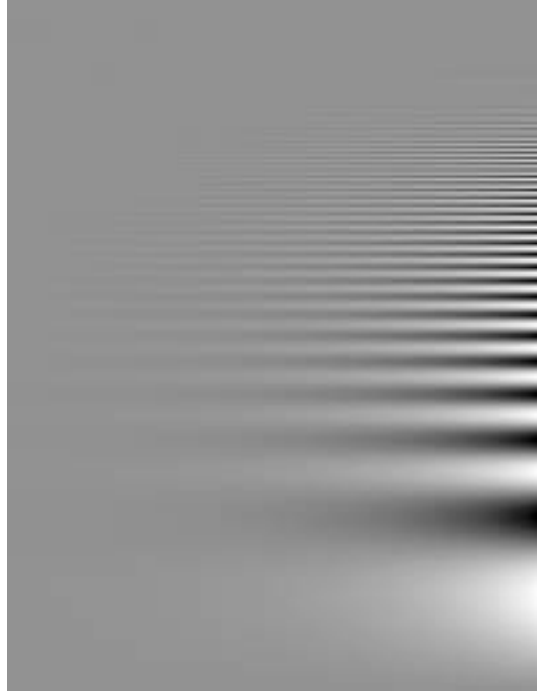


Figure 5.4: Orientation of Campbell-Robson eye chart during experiments. The chart is shrunk for the figure, with the result that detail at the top of the chart cannot be seen.

5.5.1.2 Details

Each of the three eye charts used are shown in Figure 5.1, Figure 5.2 and Figure 5.3, though Figure 5.2 had a different orientation. Eight bit TIFF grayscale images of the charts are tested. The ETDRS chart is 433 by 388 pixels, and the Pelli-Robson chart is 461 by 461 pixels. The Campbell-Robson chart is 612 by 792 orientated with the high frequency components at the top and intensity increasing from left to right, as shown in Figure 5.4.

All the eye charts are processed with the STRIVE algorithm, and the resulting STRIVE curves are measured with $PERS_{va}$. Multiple virtual eyes are generated and used in step 2 of the STRIVE algorithm. The starting value of RMS is set to 0.0. With each increment in RMS the underlying Zernike coefficients are increased in fixed increments until the maximum value of RMS for each virtual eye is reached. There are 20 increments between starting RMS value and the maximum RMS for each virtual eye.

Experiments are developed in and run with Matlab. Matlab code generates the Zernike polynomials with Zernike coefficients. Then we convert the Zernike wavefront to a PSF, after which the PSF is used as an image convolution kernel on the eye charts. The PSF is encoded in a 40 by 40 matrix, within which the minimum size of an optical aberration generated by a $0 \mu\text{m}$ Zernike coefficient is 6 by 6.

As is standard with image convolutions the border of the images getting convolved must be managed. For the experiments Figure 5.1 and Figure 5.3 have a white border half the width of the PSF matrix added, which stops the eye charts becoming cropped during the convolution. With the Figure 5.2 Campbell-Robson eye chart the area analysed in the experiments is the area that remains after the chart is cropped by a 20 pixel border.

5.5.2 Hypothesis

Both hypothesis are linked, the second hypothesis relies on the first. Identifying the location of perceptual stability requires that the degree of perceptual stability measure is accurate and reliable.

For Hypothesis 2 the accurate ranking of $PERS_{va}$ (Step 10 in Section 5.5.3) will be correct when Hypothesis 1 holds true.

5.5.2.1 Hypothesis 1: Degree of Perceptual Stability (PER)

The first hypothesis is that the STRIVE algorithm with $PERS_{va}$ successfully generates meaningful $PERS_{va}$ scores. Meaningful is defined as agreeing with human judgements. Human judgements are known for eye charts, therefore the $PERS_{va}$ scores should correctly score the perceptual stability of eye charts.

5.5.2.2 Hypothesis 2: Location of Perceptual Stability (LocPER)

The second hypothesis is linked to the first, in that the STRIVE algorithm with $PERS_{va}$ will successfully make predictions about the location of perceptual stability on eye charts. Success is where the predictions are consistent with human judgements, that is regions of eye charts which people find perceptually unstable would be the same as what STRIVE with $PERS_{va}$ finds perceptually unstable.

5.5.3 Experiment Design

To test Hypothesis 2, and indirectly validate Hypothesis 1, the experiments are structured as follows:

1. eye charts are used as stimuli
2. eye charts are divided into multiple regions
3. regions are equally sized within an eye chart
4. thousands of normal virtual eyes are randomly generated
5. each virtual eye is considered a subject of the experiment
6. every virtual eye is used to generate simulated retinal images
7. NNSE for each region in the simulated retinal images is measured
8. all the NNSE measures are used to generate STRIVE curves
9. each STRIVE curve is used to generate a perceptual stability score $PERS_{va}$
10. for each virtual eye the $PERS_{va}$ score for each region is ranked

11. ranking starts at 1 for the most stable region in a stimuli, and increases in units of 1 as instability increases
12. total rankings across all eyes is calculated for each stimuli, i.e. establish what are the dominant region rankings are for each visual design
13. dominant region rankings are compared with human rankings to establish whether they concur.

5.5.3.1 Details

Each eye chart is independently used as a stimulus, and each eye chart is divided into a range of regions depending on how the eye chart is used to test human vision. Table 5.1 lists how each chart is broken into regions, along with how human perceivers actually perceive each chart.

Three hundred independent virtual eyes are created for each part of each experiment, i.e. 300 virtual eyes are created for the 2 by 2 region analysis of the ETDRS chart and another 300 virtual eyes are created for the 2 by 1 region analysis of the ETDRS chart. Pupil size is fixed at 6 mm.

It is important to note that the same 300 eyes are not used for each experiment sub-part as using a wider range of virtual eyes means *STRIVE* and $PERS_{va}$ are more robustly tested. In total 3600 virtual eyes are tested across the four experiments.

Each virtual eye is exposed to the stimuli in multiple steps from 0 RMS to the maximum RMS, with RMS increasing to the maximum in 20 equal increments. At each exposure a simulated retinal image is generated, which represents how a real eye with equivalent Zernike coefficients to the virtual eye will see the eye chart.

Every simulated retinal image has its NNSE measured and all the NNSE meas-

Table 5.1: List of experiments arranged by eye charts and region divisions along with standard human visual acuity judgements. Human visual acuity column lists the established human judgements of the eye charts, where "Top to bottom" means the perceptual stability of the eye chart decreases from the top to the bottom of the chart.

Experiment	Eye Chart	Regions	Human visual acuity
Exp 1, part 1	ETDRS	1 by 2	Top to bottom
Exp 1, part 2	ETDRS	1 by 3	Top to bottom
Exp 1, part 3	ETDRS	1 by 4	Top to bottom
Exp 2, part 1	Pelli-Robson	2 by 1	Left to right
Exp 2, part 2	Pelli-Robson	2 by 2	Left to right downwards
Exp 2, part 3	Pelli-Robson	2 by 3	Left to right downwards
Exp 3, part 1	Campbell-Robson	1 by 2	Bottom to top
Exp 3, part 2	Campbell-Robson	1 by 3	Bottom to top
Exp 3, part 3	Campbell-Robson	1 by 4	Bottom to top
Exp 4, part 1	Campbell-Robson	2 by 2	CSF curve on right
Exp 4, part 2	Campbell-Robson	3 by 3	CSF curve on right
Exp 4, part 3	Campbell-Robson	4 by 4	CSF curve on right

ures are recorded as floating point numbers. The NNSE measures are used to generate STRIVE curves, then the STRIVE curves are used to generate $PERS_{va}$ measures.

Multiple STRIVE curves are generated for each virtual eye as each stimuli is divided into regions. Dividing a stimuli into regions is undertaken to enable comparison of the $PERS_{va}$ values of the regions. For example the ETDRS eye chart are divided into 2 equal sized regions, which divides the chart into a top half and a bottom half.

5.5.3.2 Expected Outcomes

If Hypothesis 2 and Hypothesis 1 prove true then the region rankings for each stimuli should agree with the region rankings of human observers (see Table 5.1, Human visual acuity). We are not asking human subjects to generate regions rankings, as this has already been extensively established in previous literature [66, 19, 34, 61].

5.5.4 Experiment 1: PER & LocPER of ETDRS Eye Chart

The ETDRS eye chart is tested by dividing it into 1 by 2 regions, 1 by 3 regions and 1 by 4 regions as shown in Figure 5.5.

Table 5.2 presents the results of the sub-parts of the experiment. $PERS_{va}$ accurately predicts that the top of the eye chart is more perceptually stable due to different visual acuities than the bottom, and correctly ranks the regions when the chart is divided into a finer number of regions.

Figure 5.6, Figure 5.7 and Figure 5.8 show the results in visual form using heatmaps. For illustrative purposes, the heatmaps overlay the eye charts - due to scaling and cropping the alignment between the heatmaps and eye charts does not exactly match that which occurred in the experiments. Equal col-

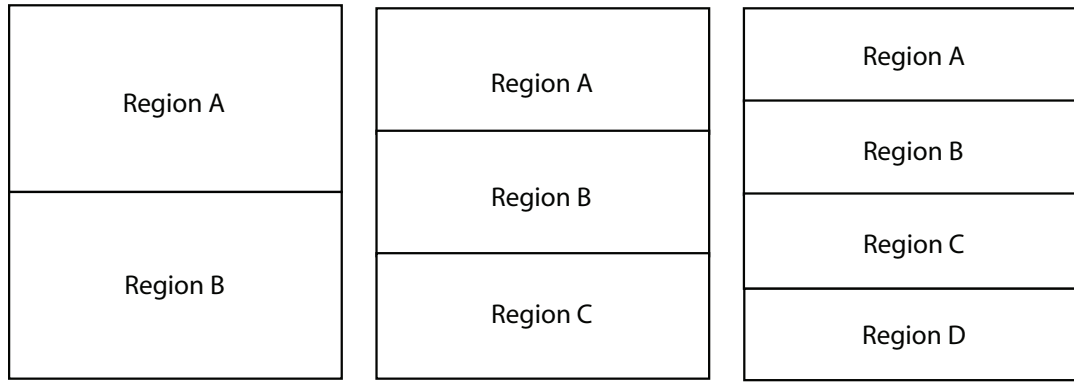


Figure 5.5: Depiction showing how the ETDRS and Campbell-Robson eye charts were divided into 1 by 2, 1 by 3 and 1 by 4 regions.

Table 5.2: Results of analysing the ETDRS eye chart for LocPER. A, B in Rank means Region A is more perceptually stable than Region B (see Figure 5.6). % Agreement is the percentage of simulated eyes per experiment that give the same ranking.

Experiment	Regions	Rank: Most to Least Stable	% Agreement
Exp 1, part 1	1 by 2	A, B	100
Exp 1, part 2	1 by 3	A, B, C	100
Exp 1, part 3	1 by 4	A, B, C, D	100

oured areas of the heatmap correspond to the regions the ETDRS eye chart are divided into. The red text is the $PERS_{va}$ score for each region, which successfully enables the correct ranking of the regions.

Region division is not by optotype line, rather the eye chart is divided into equal sized regions. When the eye chart is divided into 5 or more regions the ranking for the regions change. This is to be expected as the content of the regions can then include regions that are mostly white space and cropped optotypes.



Figure 5.6: Heatmap of ETDRS eye chart when divided into 1 by 2 regions. Red text shows normalised $PERS_{va}$ scores.



Figure 5.7: Heatmap of ETDRS eye chart when divided into 1 by 3 regions. Red text shows normalised $PERS_{va}$ scores.

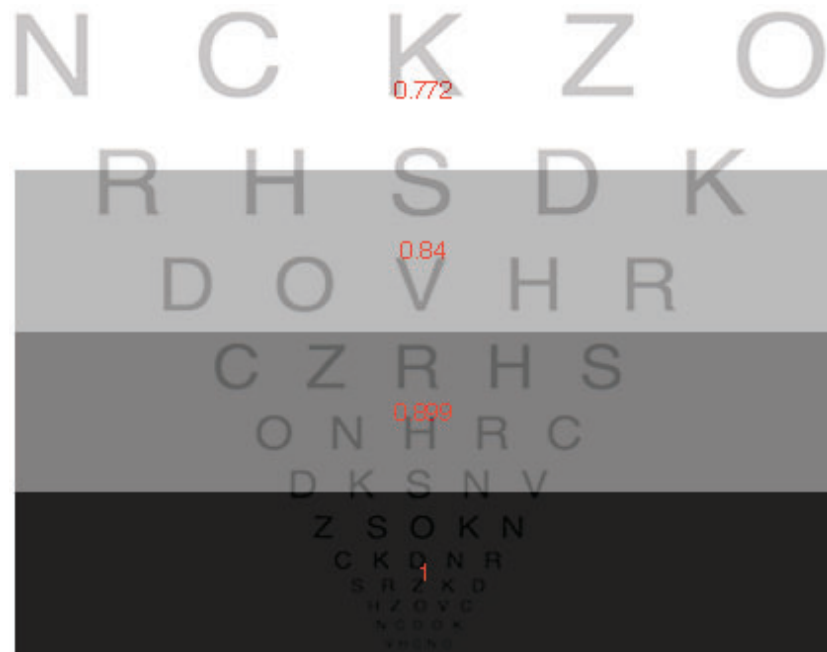


Figure 5.8: Heatmap of ETDRS eye chart when divided into 1 by 4 regions. Red text shows normalised $PERS_{va}$ scores.

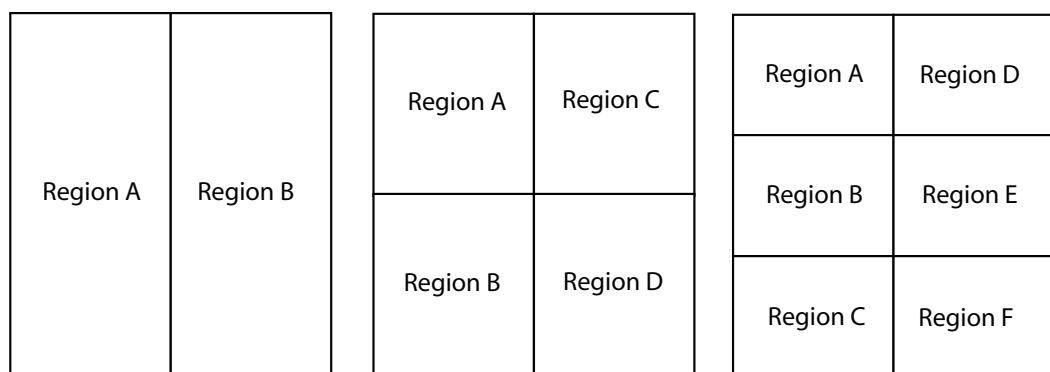


Figure 5.9: 2 by 1, 2 by 2 and 2 by 3 regions the Pelli-Robson chart is divided into to test for LocPER contrast sensitivity.

Table 5.3: Averaged result of analysing the Pelli-Robson eye chart for LocPER. When rankings are not separated by commas they are of equal rank. Ranking is carried out based on average $PERS_{va}$ scores for each region. Average $PERS_{va}$ scores are generated from $PERS_{va}$ region scores for each virtual eye. Ranking for 2 by 1 regions had 100% agreement. Percentage agreement breakdown for 2 by 2 regions in Table 5.4, and 2 by 3 regions in Table 5.5.

Experiment	Regions	Rank: Most to Least Stable
Exp 2, part 1	2 by 1	A, B
Exp 2, part 2	2 by 2	A, C, B, D
Exp 2, part 3	2 by 3	A, D, E, B, C, F

5.5.5 Experiment 2: PER & LocPER of Pelli-Robson Eye Chart

As shown in Figure 5.9 the Pelli-Robson chart is tested by dividing it into 2 by 1, 2 by 2 and 2 by 3 regions. 2 regions across are used as there are 6 optotypes per line, which are designed to test human vision by grouping optotypes into sets of 3 optotypes sharing the same contrast [66].

Table 5.3 shows the results, with the corresponding heatmaps shown in Figure 5.10, Figure 5.11 and Figure 5.12. When the Campbell-Robson eye chart is divided into 2 by 1 and 2 by 2 regions the results are as expected, i.e. perceptual stability decreases from left to right going downwards.

When the eye chart is divided into 2 by 3 regions the results are correct going from the top to the bottom of the chart, that is perceptual stability decreases from the top to the bottom of the columns. When going from the left to the right of the chart region A, D, C and F are correct, but regions E and B are ranked in reverse order to the expected order. From the breakdown in Table 5.5 we can see that the correct ordering occurred 13% of the time, even though the averaged result are reasonably accurate.

As the number of regions increases above 2 by 3 the rankings continue becoming less ordered and conform less to human judgements. For 2 by 1 and 2 by

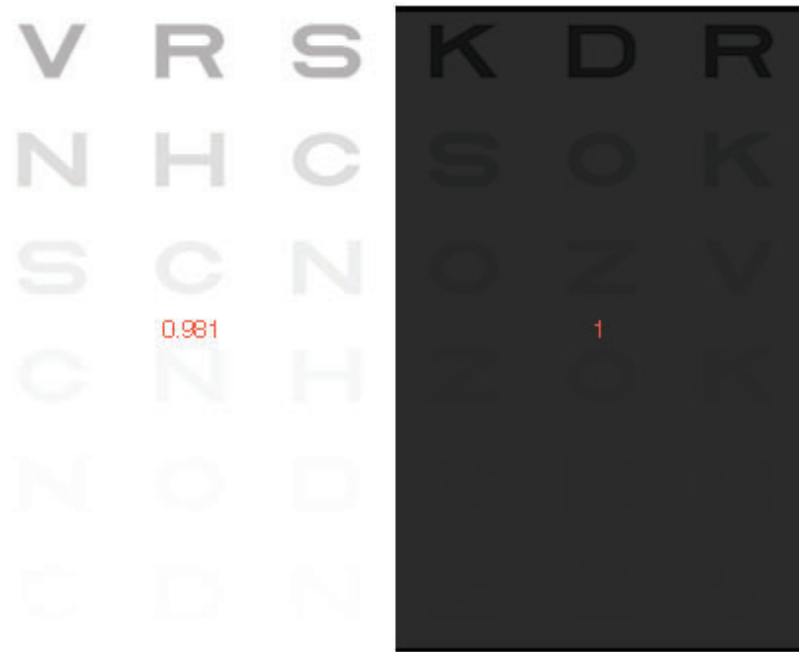


Figure 5.10: Heatmap of Pelli-Robson eye chart when divided into 2 by 1 regions. Red text shows normalised $PERS_{va}$ scores.



Figure 5.11: Heatmap of Pelli-Robson eye chart when divided into 2 by 2 regions. Red text shows normalised $PERS_{va}$ scores.



Figure 5.12: Heatmap of Pelli-Robson eye chart when divided into 2 by 3 regions. Red text shows normalised $PERS_{va}$ scores.

Table 5.4: Breakdown of results of analysing the Pelli-Robson eye chart for LocPER. Chart divided into 2 by 2 regions. 300 virtual eyes used. Rank number ranges from 1 to 4 and regions A to D. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C	D
61	1	3	2	4
22	1	4	2	3
8	1	2	3	4
4	1	4	3	2
4	2	1	3	4
1	3	1	4	2

Table 5.5: Breakdown of results of analysing the Pelli-Robson eye chart for LocPER. Chart divided into 2 by 3 regions. 300 virtual eyes used. Rank number ranges from 1 to 6 and regions A to F. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C	D	E	F
31	1	4	5	2	3	6
8	1	4	5	3	2	6
2	1	5	6	3	2	4
1	2	4	6	3	5	1
3	1	4	6	3	2	5
14	1	4	6	2	3	5
6	1	3	6	2	4	5
3	2	5	6	3	4	1
1	1	4	3	2	5	6
13	1	3	5	2	4	6
1	1	5	4	3	2	6
1	2	5	3	4	6	1
1	2	5	6	3	1	4
1	1	5	6	3	4	2
2	1	5	4	2	3	6
2	1	5	3	2	4	6
1	2	6	5	3	4	1
2	2	5	6	4	3	1
1	1	4	6	2	5	3
1	2	4	6	3	1	5
1	1	5	6	2	3	4
1	3	4	2	5	6	1
1	2	6	3	4	5	1
1	1	5	6	4	3	2
1	2	3	6	4	5	1

2 regions the results are correct and the column ordering for 2 by 3 is correct. These results, coupled with the bias found below, tell us that $PERS_{va}$ does take contrast into consideration but is more sensitive to optotype shape rather than optotype contrast.

5.5.5.1 Bias leading to 2 by 3 region E, B order

In order to establish why the results are not completely as expected for 2 by 3 regions let us examine the contribution of the optotype shape to the $PERS_{va}$ scores. With the Pelli-Robson chart the optotype shape and contrast is varied, though the design of the chart is meant to balance and thereby remove the variation between optotypes groups due to optotype shapes, i.e. optotype contrast should effect subject's visual acuity judgments rather than shape.

To examine the contribution of optotype shape we set the contrast of all the optotypes to be equal, then obtain a ranking for 100 virtual eyes of the first four lines of the Pelli-Robson chart. Once this is complete we compute the mean rank and standard deviation of the rank for each region.

Figure 5.13 shows the results, where we see region C and region G have ranks that are separated by a value of 3.44 with low standard deviations. Making them the widest separated regions.

These results tell us that independent of contrast the shape of the optotypes are biasing the ranking towards region C having a lower ranking than region G, which for the 2 by 3 region Pelli-Robson chart means region E is biased to have a better ranking than region B.

One very interesting implication of the shape bias is whether the bias is due to the nature of how STRIVE and $PERS_{va}$ work (possible), or whether there is a bias in the Pelli-Robson chart itself? Unfortunately establishing whether there is a bias in the chart is well outside the scope of this thesis.

Region A m: 6.83 std: 1.694	Region E m: 4.66 std: 1.1825
Region B m: 3.34 std: 2.1092	Region F m: 5.86 std: 1.907
Region C m: 5.87 std: 1.3077	Region G m: 2.43 std: 1.3428
Region D m: 3.1 std: 2.5839	Region H m: 3.91 std: 1.6087

Figure 5.13: Results of first four lines of Pelli-Robson contrast chart, when divided into 2 by 4 regions to test effect of optotype shape without contrast. M is mean rank and std is the standard deviation.

5.5.6 Experiment 3: Row Regions PER & LocPER of Campbell-Robson Eye Chart

For Experiment 3 the Campbell-Robson eye chart is orientated as shown in Figure 5.4. Slight gradients are present in the Campbell-Robson eye chart so the Noise Threshold is set to 10 (see Section 4.2.5). The chart is divided into row regions as shown in Figure 5.5.

Table 5.7 lists the results, with heatmaps for the results in shown in Figure 5.14, Figure 5.15 and Figure 5.16.

As can be seen in the heatmaps for regions 1 by 2 and 1 by 3 the results are as expected, with the perceptual stability decreasing from the bottom to the top of the chart. For regions 1 by 4 the results are correct for the top three rows, but the fourth row (region D) has a higher $PERS_{va}$ value 0.553 than the third row (region C) 0.503. This is in reverse order to the expected order.

When more than 4 rows are used the results do continue as expected for the upper part of the chart, with CSF curves forming. As the number of rows increase the size of each region decreases, with the effect that regions towards

the bottom of the chart become composed of continuous gradients. The result is that the region ranking for the bottom of the chart is dependent on the NT for accurate ranking.

These results tell us that $PERS_{va}$ does make predictions as expected when there are visual details, but it is open to improvement for continuous gradients.

5.5.6.1 Noise Threshold sensitivity of Region D gradient

Upon investigating the reason for the reverse ordering we found that the Noise Threshold is very important for the correct classification of the bottom regions in the Campbell-Robson eye chart.

This is because region D is predominately composed of a continuous gradient, rather than multiple visual details. As the Noise Threshold value is increased in increments the results converge on the expected, until they are ordered correctly when the Noise Threshold is set to 20. Listed in Table 5.6 are the changes in $PERS_{va}$ values with increases in the Noise Threshold.

Region D exhibits the most sensitivity to NT changes due to the continuous gradient. A continuous gradient leads to sudden changes in the underlying $PERS_{va}$ NNSE values when the NNSE minimum and maximum cover a small range. More details on NT are in Section 4.2.5.

One possible solution to this issue is to improve the STRIVE algorithm, such that it detects when the rate of change of the STRIVE curve is due to a small NNSE minimum and maximum range. If it is due to a small NNSE range then automatically set an appropriate NT value.

Table 5.6: Effect of increases in Noise Threshold on $PERS_{va}$ values for Campbell-Robson eye chart divided into 1 by 4 regions. Once NT hits 20 the results are as expected.

Noise Threshold	Region A	Region B	Region C	Region D
0	0.9992	0.746	0.5031	0.5804
10	0.9996	0.7463	0.5033	0.5535
20	0.9996	0.7463	0.5033	0.4998
30	0.9996	0.7463	0.4999	0.4535
40	0.9996	0.7463	0.4944	0.4145
50	0.9996	0.7463	0.4944	0.36

Table 5.7: Averaged result of analysing the Campbell-Robson eye chart for LocPER when the eye chart is divided into rows. When rankings are not separated by commas they are of equal rank. Ranking is carried out based on average $PERS_{va}$ scores for each region. Average $PERS_{va}$ scores are generated from $PERS_{va}$ region scores for each virtual eye.

Experiment	Regions	Rank: Most to Least Stable	Agreement
Exp 3, part 1	1 by 2	B, A	100%
Exp 3, part 2	1 by 3	C, B, A	in Table 5.8
Exp 3, part 3	1 by 4	C, D, B, A	in Table 5.9

Table 5.8: Breakdown of results of analysing the Campbell-Robson eye chart for LocPER. Chart divided into 1 by 3 regions. 300 virtual eyes used. Rank number ranges from 1 to 3 and regions A to C. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C
18.333	3	1	2
81.000	3	2	1
0.667	2	1	3

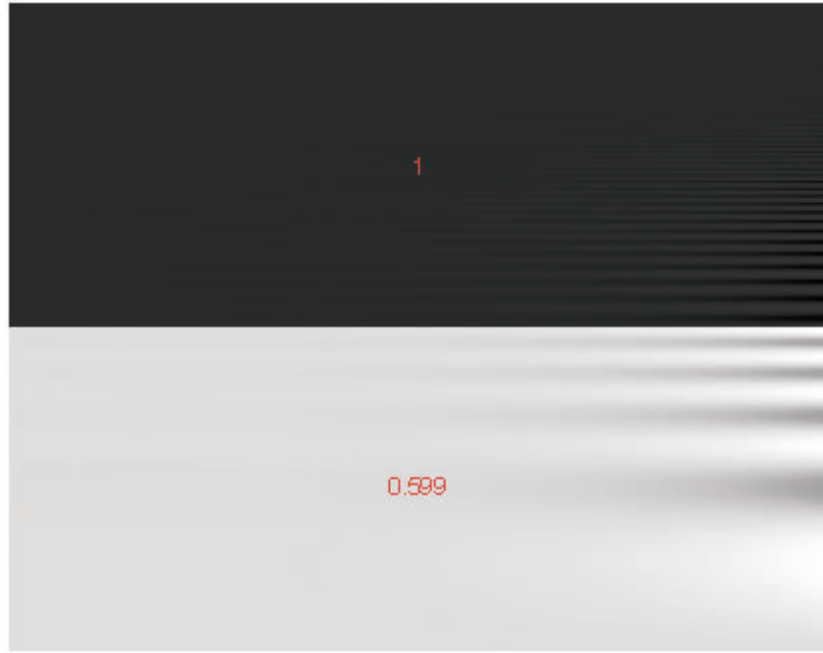


Figure 5.14: Heatmap of Campbell-Robson eye chart when divided into 1 by 2 regions. Red text shows normalised $PERS_{va}$ scores.

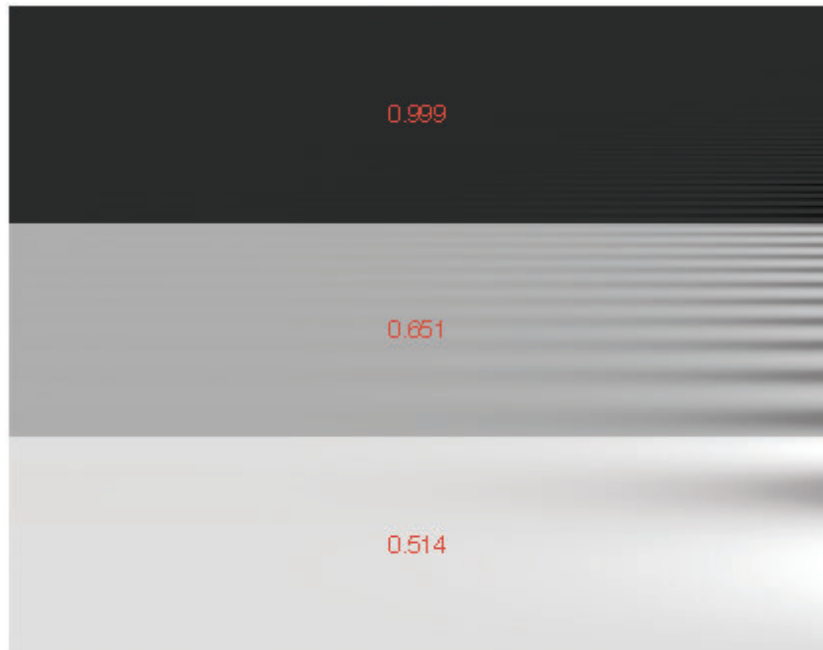


Figure 5.15: Heatmap of Campbell-Robson eye chart when divided into 1 by 3 regions. Red text shows normalised $PERS_{va}$ scores.

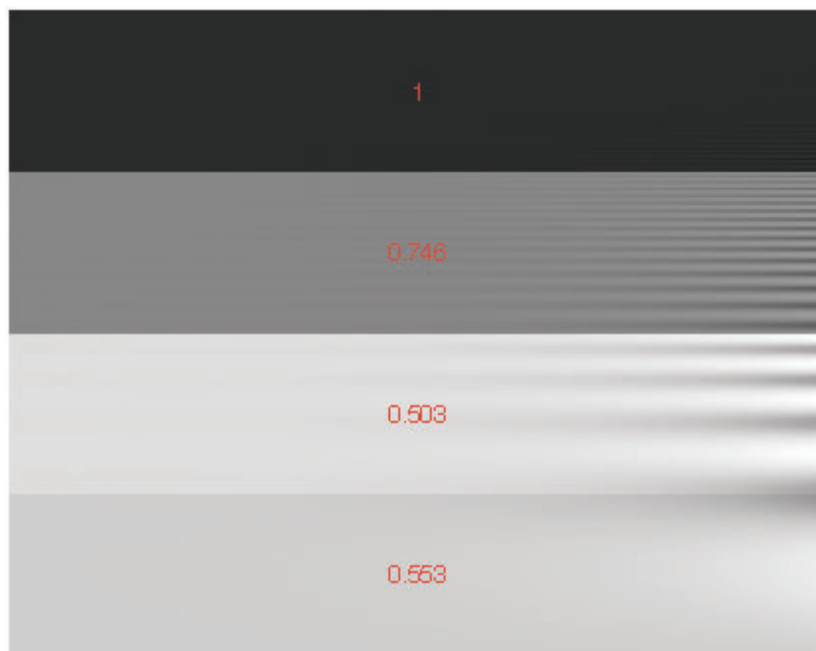


Figure 5.16: Heatmap of Campbell-Robson eye chart when divided into 1 by 4 regions. Red text shows normalised $PERS_{va}$ scores.

Table 5.9: Breakdown of results of analysing the Campbell-Robson eye chart for LocPER. Chart divided into 1 by 4 regions. 300 virtual eyes used. Rank number ranges from 1 to 4 and regions A to D. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C	D
34.000	4	3	2	1
46.000	4	3	1	2
19.333	4	2	1	3
0.667	3	2	1	4

Region A	Region D	Region G	Region A	Region E	Region I	Region M
Region B	Region E	Region H	Region B	Region F	Region J	Region N
Region C	Region F	Region I	Region C	Region G	Region K	Region O
			Region D	Region H	Region L	Region P

Figure 5.17: 3 by 3 and 4 by 4 regions the Campbell-Robson chart is divided into to test for the CSF curve.

5.5.7 Experiment 4: Grid Regions PER & LocPER of Campbell-Robson Eye Chart

By dividing the chart into 2 by 2, 3 by 3 and 4 by 4 regions as shown in Figure 5.9 and Figure 5.17 we seek to establish whether $PERS_{va}$ enables the ranking of regions into the area under a CSF curve.

Table 5.10 has the resultant ranking when the $PERS_{va}$ scored for each region is averaged over 300 virtual eyes. The average rankings do agree with human vision by producing areas under the CSF curves where the least stable regions are at the top and to the right of the chart.

The effects of individual variability in eye function are more obvious with the Campbell-Robson eye chart. In Table 5.11 the breakdown of the ranks is shown for 2 by 2 regions, with the related heatmap in Figure 5.18. Not all virtual eyes gave the same region ranking with the Campbell-Robson chart. Clusters of ranks emerge, with 94.333% of the virtual eyes in agreement about the ranking of the regions.

For the 3 by 3 grid the area under the CSF curve is a little easier to discern in Figure 5.19, with the corresponding breakdown in Table 5.12. As the number of regions increases in the grid the diversity of rankings increases. Even though

Table 5.10: Averaged result of analysing the Campbell-Robson eye chart for LocPER, when chart divided into grids. When rankings are not separated by commas they are of equal rank. Ranking is carried out based on average $PERS_{va}$ scores for each region. Average $PERS_{va}$ scores are generated from $PERS_{va}$ region scores for each virtual eye.

Experiment	Regions	Rank: Most to Least Stable	Agreement
Exp 4, part 1	2 by 2	B, D, A, C	in Table 5.11
Exp 4, part 2	3 by 3	A B C F, I, E, H, D, G	in Table 5.12
Exp 4, part 3	4 by 4	A B C D E G H, L, K, F, O, P, J, N, M, I	115 groups

76.33% of the virtual eyes produce the same ranking, clusters of rankings begin to emerge with 8% and 7% of eyes in agreement.

When the number of regions is increased to 4 by 4 the resulting area forming the CSF curve becomes clearer in the heatmap shown in Figure 5.20. Furthermore 115 different clusters of ranking occur, with no clear dominant ranking emerging - though the averaged result does correctly form a CSF curve. The variety and diversity of individual rankings is not unexpected, as the Campbell-Robson eye chart is commonly used for research purposes due to its heightened sensitivity to individual eye function. With the effect that individual variability in eye function is more readily detected with the Campbell-Robson eye chart, leading to more varied rankings.

Region P does have an unexpectedly high value 0.537, which leads to it having an incorrect ranking. As with region D in Exp 3 Part 3 this is an artifact of the low value of the Noise Threshold, since region P is predominately a gradient. When the NT value is increased the region ranking for region P is correct.

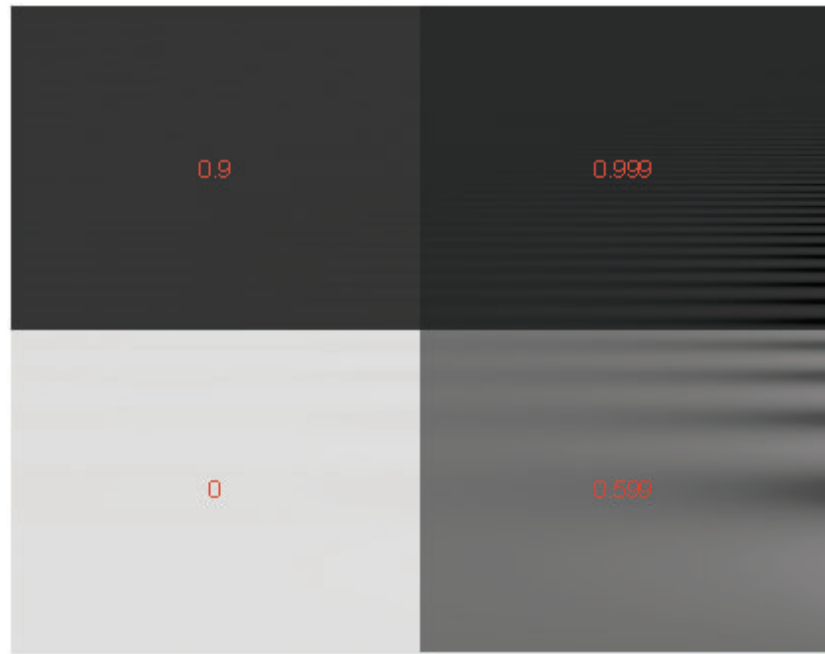


Figure 5.18: Heatmap of Campbell-Robson eye chart when divided into 2 by 2 regions. Red text shows normalised $PERS_{va}$ scores.

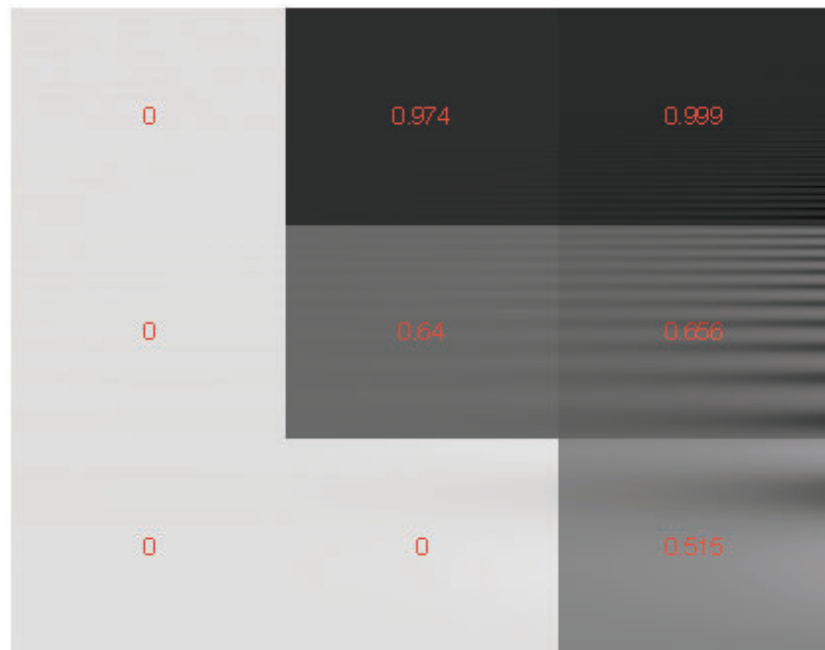


Figure 5.19: Heatmap of Campbell-Robson eye chart when divided into 3 by 3 regions. Red text shows normalised $PERS_{va}$ scores.

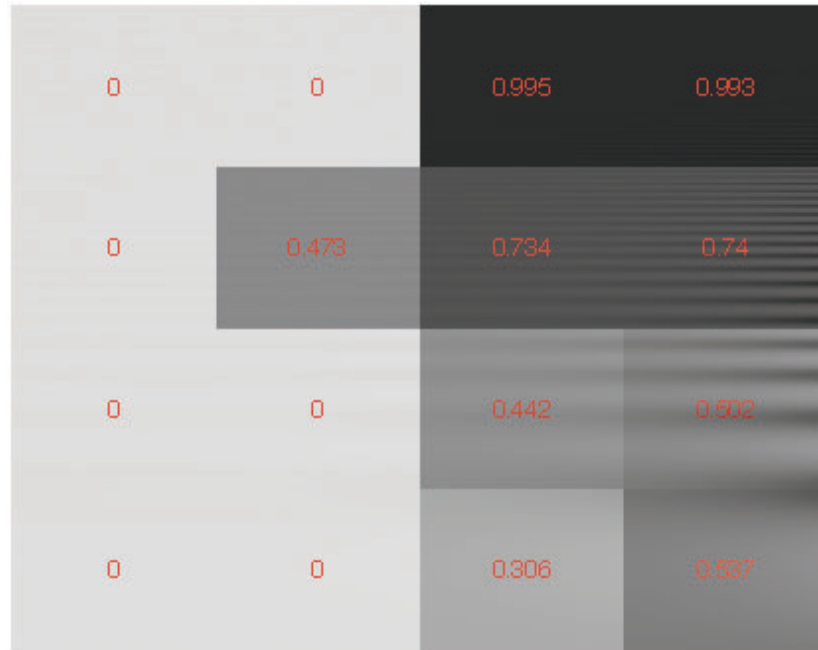


Figure 5.20: Heatmap of Campbell-Robson eye chart when divided into 4 by 4 regions. Red text shows normalised $PERS_{va}$ scores.

Table 5.11: Breakdown of results of analysing the Campbell-Robson eye chart for LocPER. Chart divided into 2 by 2 regions. 300 virtual eyes used. Rank number ranges from 1 to 4 and regions A to D. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C	D
94.333	3	1	4	2
4.333	4	1	3	2
0.334	1	1	3	4
1.000	1	1	4	3

Table 5.12: Breakdown of results of analysing the Campbell-Robson eye chart for LocPER. Chart divided into 3 by 3 regions. 300 virtual eyes used. Rank number ranges from 1 to 9 and regions A to I. When multiple regions share the same rank they have the same rank value. A higher rank means a region is more perceptually stable.

% Agreement	A	B	C	D	E	F	G	H	I
76.330	1	1	1	8	6	1	9	7	5
8.000	1	1	1	8	5	1	9	6	7
7.000	1	1	1	8	6	1	9	5	7
1.000	1	1	1	8	7	1	9	5	6
2.330	1	1	1	8	7	1	9	6	1
0.670	1	1	1	7	6	1	8	5	9
0.670	1	1	1	9	5	1	8	6	7
2.000	1	1	1	8	7	1	9	6	5
0.033	1	1	1	9	6	1	8	5	7
0.033	1	1	1	8	6	1	9	7	1
1.000	1	1	1	9	6	1	8	7	5
0.033	1	1	1	8	5	1	9	7	6

5.6 Conclusion

When STRIVE with $PERS_{va}$ are used to evaluate eye charts they perform well at measuring the perceptual stability of a range of different visual stimuli (Hypothesis 1). In most cases the averaged results are in agreement with human judgements about eye charts (Hypothesis 2).

Areas for further improvement have been shown. Sensitivity to contrast versus shape can be improved, and the approach to Noise Thresholding can be improved for gradients. Region size is also found to have an effect on the accuracy of ranking, with increasingly smaller regions leading to less accurate rankings.

An unexpected optometry contribution from this work was presented in Section 5.5.5.1, where we experimental show how $PERS_{va}$ may have found a bias in the Pelli-Robson eye chart. Within optometry a considerable amount of research is focused on detecting and removing biases in eye charts.

Visual Stability Predictor: Colour Perception

In this chapter STRICODI (SimulaTed Retinal Images with COlour DIstances) for measuring the effects of differences in colour perception on the perception of visual designs is presented. Also introduced is $PERS_{cp}$, which generates a score of the perceptual stability due to differences in colour perception.

With $PERS_{cp}$ information visualisations and interfaces can be automatically evaluated to test how differently they appear to individuals with variants of colour deficient vision.

6.1 Predicting Colour Stability

In Section 4.1 of Chapter 4 our approach to predicting perceptual stability for visual acuity is introduced. As with that predictor, here a predictor of colour stability is required to make predictions about:

- Degree of perceptual change ($PERS$)
- Location of perceptual change ($LocPERS$)

A predictor of colour stability should help identify how visually unstable a design may be when viewed by an individual with colour deficient vision (CDV), i.e. what areas of a visual design exhibit weakness' due to differences in colour perception? Weakness' arise because a visual design may appear unexpectedly different when viewed with reduced colour spaces. Such as how the red shoes in Figure 6.2 are less obvious in the simulated retinal image.

By measuring the perceptual stability due to colour perception the perceptual effects of individual differences in visual function contribute to establishing the upper bound on display "information" density (see Section 4.1).

As outlined in Section 4.1.2 any predictor needs to correlate with human performance. To establish whether the predictor of colour stability achieves this here it is tested on Ishihara eye charts [61], which are used to detect and classify the extent to which an individual does or does not have colour blindness.

6.2 Measuring Colour Perception Visual Stability

Colour stability is measured by establishing whether the colours in a visual design change significantly when viewed by eyes with different colour spaces (see Section 2.3.2 and Section 3.6).

In this thesis the colour difference measure ΔE_{00}^* , which is discussed in Section 3.6 of Chapter 3 is built upon to create the perceptual stability measure.

6.2.1 Perceptual Measure

In manufacturing, art and other areas various ΔE measures have been used to measure and maintain colour accuracy. Often these measures are used to detect whether a colour stays within particular a range. That is, the ΔE tolerance is set to an acceptable upper bound, and anything which is not within the

bound is an unacceptable colour match.

For example, during the manufacturing of many household goods and clothes, the manufactured household goods are regularly automatically checked to make sure colour flaws are not occurring, e.g. make sure white dinner plates are actually white.

6.2.2 Perceptual Stability Algorithm: STRICODI

As individual differences in colour perception occur we examine whether colour differences, as measured with ΔE_{00}^* , can be used to predict which sections of a visual design appear perceptually different.

Shown in Figure 6.1 is our STRICODI (SimulaTed Retinal Images with COlour DIstances) algorithm for quantifying the effect of colour differences on visual designs. The steps in the algorithm are as follows:

1. Visual design is captured as an RGB image, e.g. Figure 6.2(a).
2. A reference image is created by converting an RGB image to CIELAB 1976 colour space, with a 2 degree standard observer and D60 reference illuminant (see Section 3.6.4).
3. A particular type of colour deficit vision to simulate is set, i.e. protanopia or deuteranopia.
4. A simulated retinal image is generated, based on the algorithm presented in Section 3.6.7 and depicted in Figure 3.11 and Figure 6.2(b).
5. For each pixel, colour differences between reference image and simulated image are measured with ΔE_{00}^* (see Section 3.6.5).
6. ΔE_{00}^* measures for all pixels are stored in a perceptual stability change map C , i.e. 2D matrix where each matrix coordinate corresponds to a

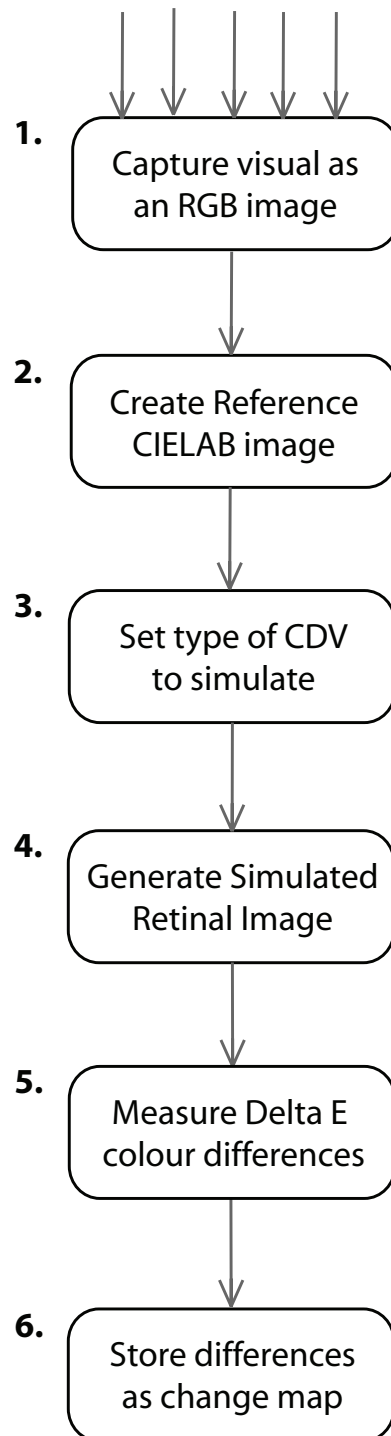


Figure 6.1: Steps in STRICODI (SimulaTEd Retinal Images with COlour Differences) algorithm to generate measure of change due to different colour perceptions of a visual design.

pixel location in the reference image, as shown in Figure 6.2(c).

Shown in Figure 6.2 is an example of the algorithm analysing a photograph of red shoes on green grass.

6.2.3 Calculating Perceptual Stability: Colour Perception

STRICODI generates a 2D matrix C containing the ΔE_{00}^* colour difference score for each pixel. The raw colour difference measures are not sufficient for accurately calculating the perceptual stability (see Section 6.2.3.1).

Equation 6.1 shows the calculation for $PERS_{cp}$. Matrix C generated by STRICODI is used to generate the $PERS_{cp}$ score. As part of calculating $PERS_{cp}$ minimum and maximum ΔE_{00}^* thresholding is applied to the perceptual stability change map values. More details on why thresholding is applied is available in Section 6.2.3.1.

As with the $PERS_{va}$ visual acuity predictor, the perceptual stability predictor for colour $PERS_{cp}$ can be applied to regions in a visual design, as well as whole designs.

Figure 6.2 shows an example of STRICODI and $PERS_{cp}$ analysing a photograph of a red shoe on green grass. Image 6.2(a) is the photograph which is analysed for perceptual colour stability. Image 6.2(b) is the simulated retinal image of Image 6.2(a), where CDV protanopia is simulated. Image 6.2(a) and Image 6.2(b) are then used to generate the colour differences perceptual stability change map, which is shown in Image 6.2(c). Image 6.2(c) is a representation of the 2D change matrix C , when the raw values in C have been scaled to range between 0 and 255 with a grayscale palette. Image 6.2(d) is the matrix C after it has been thresholded where TMIN is 10 and TMAX is 20, after which it is adjusted to range between 0 and 255 with a grayscale palette. In Image 6.2(d) it is clear that the red shoes that undergo the most perceptual change

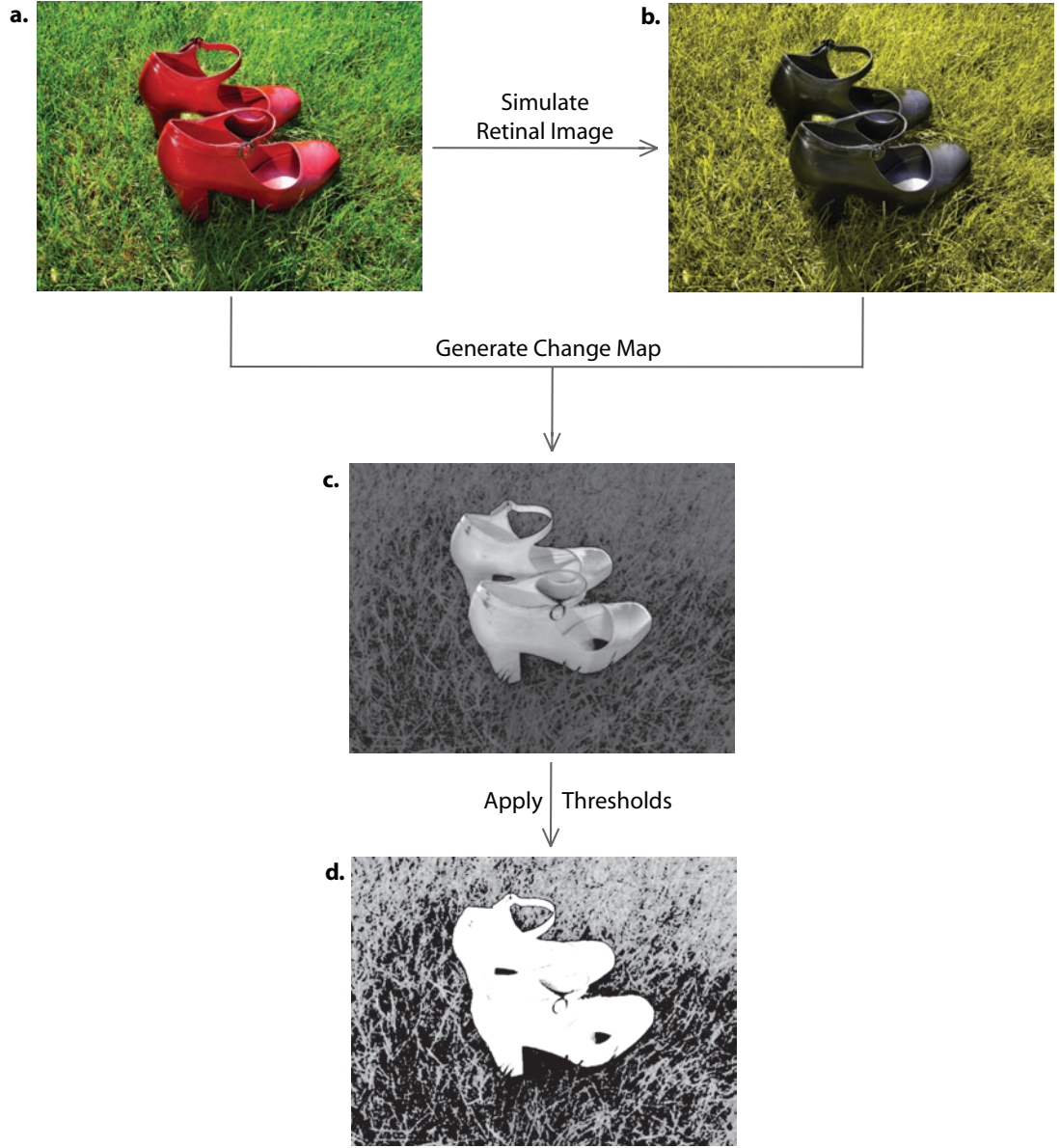


Figure 6.2: Example of STRICODI with $PERS_{cp}$ thresholding an image. Image 6.2(a) is the original image, and shows red shoes on green and yellow grass. Image 6.2(b) is the protanope version of Image A. Image 6.2(c) is the perceptual stability change map C generated by STRICODI. Image 6.2(d) is the thresholded version of Image 6.2(c) - where white means more change and black means no change. The values in matrix C were scaled to make Image 6.2(c) and Image 6.2(d) easier to see. Thresholding is applied where TMIN is 10 and TMAX is 20 (see Equation 6.1).

Equation 6.1: Generate $PERS_{cp}$ (perceptual stability for colour perception) score by measuring the thresholded area of the colour difference perceptual stability change map C .

$$PERS_{cp} = \frac{\sum_{i,j}^{n,m} C_{i,j}}{n * m} \quad (6.1)$$

where

$$C_{i,j} = \begin{cases} C_{i,j} & TMIN \leq C_{i,j} \leq TMAX \\ TMAX & C_{i,j} > TMAX \\ 0 & C_{i,j} < TMIN \end{cases}$$

and

n is the count of the pixels across
 m is the count of the pixels down

$TMAX$ is maximum ΔE_{00}^* threshold (see Section 6.2.3.1)
 $TMIN$ is minimum ΔE_{00}^* threshold (see Section 6.2.3.1)

have the highest change value, i.e. white is most change, black is no change.

The $PERS_{cp}$ score for Image 6.2(a) is 9.4105, while the heatmap in Figure 6.3 shows the $PERS_{cp}$ scores when Image 6.2(a) is broken into 3 by 3 equal sized regions. As expected, the regions in the middle of the heatmap are more perceptually unstable, i.e. the location of the red shoes in Image 6.2(a). The regions at the top of the heatmap are more perceptually unstable than regions at the bottom of the heatmap, which corresponds with the colour of the grass in Image 6.2(a), i.e. at the bottom of Image 6.2(a) the grass is green with yellow, while towards the top of the image the grass is more consistently green.

6.2.3.1 Thresholding

Raw ΔE_{00}^* measures are not sufficient for calculating the perceptual stability because the ΔE_{00}^* measure signifies whether two colours differ, but large values of ΔE_{00}^* do not meaningfully indicate how different the colours are.

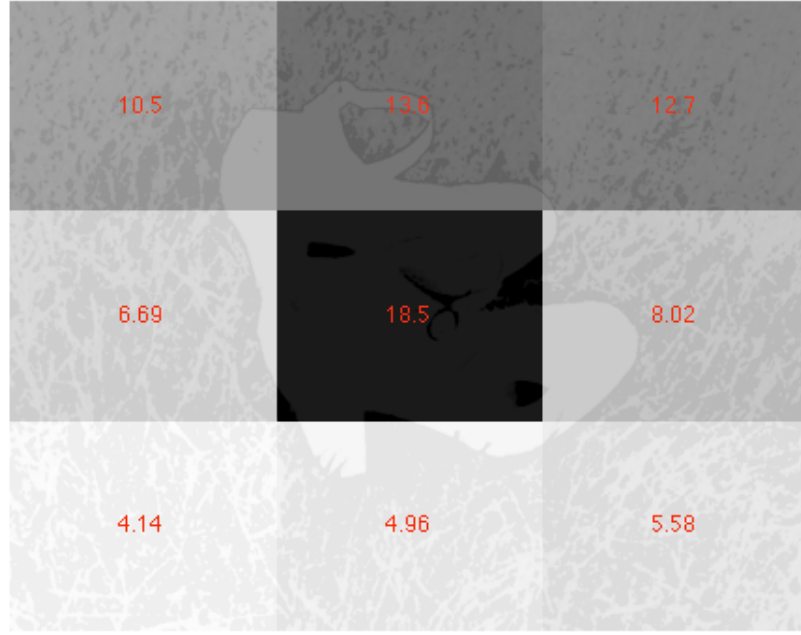


Figure 6.3: Heatmap of Image 6.2(a) (see Figure 6.2) when it is divided into 3 by 3 equal sized regions. Heatmap values are $PERS_{cp}$ scores. Higher score means less perceptually stable, while lower score means more perceptually stable.

For example if there are four colours A, B, C and D - where A (RGB: 255, 0, 0) and B (RGB: 255, 53, 48) are different shades of red, C (RGB: 27, 81, 255) is blue and D (RGB: 62, 192, 41) is green. The ΔE_{00}^* measure between A and B is 5.5038, which indicates that the colours are perceptually close but not exactly the same. For A and C the ΔE_{00}^* measure is higher at 53.2992, while for A and D the ΔE_{00}^* score is higher again at 74.3559.

A measure of 74.3550 ΔE_{00}^* between A and D does not mean that A and D are more perceptually different than A and C. Large ΔE_{00}^* scores do not clearly indicate the perceptual relationship between colours, while small ΔE_{00}^* scores do indicate how perceptually close two colours are. A large value does mean the colours are significantly different, but the degree of difference between the colours is not captured in a large ΔE_{00}^* score.

For this work the implications are that large ΔE_{00}^* scores in the STRICODI perceptual stability change map need to be accounted for indirectly in the $PERS_{cp}$ measure. Thresholding is used to account for the effects of large ΔE_{00}^* measures. In Equation 6.1 the thresholding ΔE_{00}^* range is set with the TMIN and TMAX values. TMIN should be set to a value which captures the required minimum ΔE_{00}^* difference between two colours, while TMAX should be set to the maximum meaningful ΔE_{00}^* difference. ΔE_{00}^* values above TMAX are no longer meaningful in the sense that higher values do not provide more information about the degree of colour difference.

6.3 Conclusion

In this Chapter STRICODI with $PERS_{cp}$ are introduced for measuring the perceptual stability of visual designs due to differences in perceiver's colour perception. An example of the techniques analysing a photograph is given, along with justifications for thresholding.

In the following Chapter $PERS_{cp}$ is evaluated on a range of eye charts, to test whether it agrees with human judgements.

Experimental Evaluation of Colour Stability Predictor

STRICODI and $PERS_{cp}$ are experimentally tested and validated on Ishihara eye charts, which shows they are effective predictors of perceptual stability for colour perception.

Ishihara eye chart plates are used as a diagnostic tool to establish whether an individual has colour deficient vision and if so what kind of colour deficient vision they experience.

By establishing that $PERS_{cp}$ make predictions consistent with human vision, we know that it can be applied to evaluating the effects of colour blindness on the perception of information visualisations and user interface designs (see Chapter 8).

7.1 Experiments: Colour Perception Stability

In Chapter 5 a number of criteria are identified, which the visual acuity predictor has to fulfill. Many of these criteria also apply to the colour perception stability predictor, that is the predictor should consistently correlate with hu-

man performance while enabling measures of:

- Degree of perceptual stability (*PERS*)
- Location of perceptual stability (*LocPERS*)

As with the visual acuity predictor (see Section 5.2), eye charts are used as the gold standard for testing whether the colour perceptual metric performs as hypothesised. That is, will the colour perceptual measure make colour predictions consistent with human judgements, i.e. parts of an eye chart that are perceptually different due to protanopia should be scored as perceptually different by the perceptual metric.

7.2 Ishihara Eye Charts

There is a large body of research into modelling and measuring colour perception differences [96, 14, 78, 58, 89, 87, 42]. One early research outcome were Ishihara plates, which are used to detect and identify differences in colour perception. Ishihara plates are used as a diagnostic tool to establish whether an individual has colour deficient vision and if so what kind of colour deficient vision they experience.

Figure 7.1 shows two Ishihara plates, along with versions of the plates as seen by protanopes. Each plate consists of a large circle made up of many small coloured dots. The dots are varied in colour and are surrounded by varying amounts of white space. Dots are coloured to form identifiable visual objects, such as numbers and lines. In most cases the visual objects are optotypes, while a few of the plates have visual objects that are curving lines formed by equivalently coloured dots.

Identification, mis-identification and alternative identification of optotypes

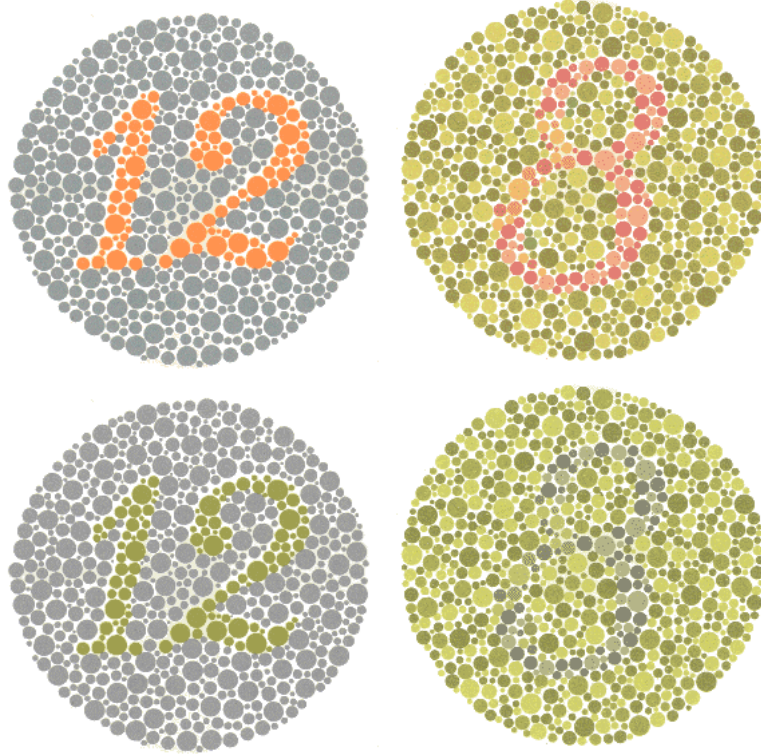


Figure 7.1: Examples of Ishihara colour plates. Number 12 in plate on the top left can be seen no matter what kind of colour blindness, while the number 8 on the plate on the top right may not be seen by individuals with protanopia and deuteranopia. Bottom left shows how the number 12 appears to individuals with protanopia, while the bottom right plate shows how the number 8 appears as number 3 to individuals with protanopia.

forms the diagnostic basis for the Ishihara plates. Optotype identification occurs when an individual either can or cannot see an optotype. Alternative identification occurs when an individual with colour deficit vision sees a visually different optotype than a typical observer. For example alternative identification occurs in Figure 7.1 when the number 8 optotype is seen as a number 3 by a protanope.

7.3 Experiments

What follows are two experiments where the STRICODI with $PERS_{cp}$ is tested on 24 standard Ishihara eye chart plates. The experiments establish that

STRICODI with $PERS_{cp}$ do measure the degree of perceptual stability, along with location of perceptual stability.

7.3.1 Overall Approach & Details

7.3.1.1 Approach

Virtual eyes simulating different types of colour deficient vision are used to generate simulated retinal images of the Ishihara plates, then STRICODI with $PERS_{cp}$ is used to evaluate the perceptual stability of the eye charts. After which the expected outcome for each eye chart is compared with the outcome from the perceptual stability measure.

7.3.1.2 Details

Shown in Figure 7.2 are the 24 Ishihara plates used in each experiment. Each plate is a TIFF RGB 8 bit per channel image. Plate sizes ranged about 280 by 275 pixels.

Each eye chart is processed with STRICODI, then each $PERS_{cp}$ is measured. Two types of colour deficient vision are simulated, protanopia and deuteranopia (see Section 2.3.2 and Section 3.6 for more details).

The experiments are developed for and run in Matlab. Matlab code converts the Ishihara plates to the CIELAB colour space. We then convert the Ishihara plates to colour deficient versions. Then we measure and store the ΔE_{00}^* colour differences in the perceptual stability change map.

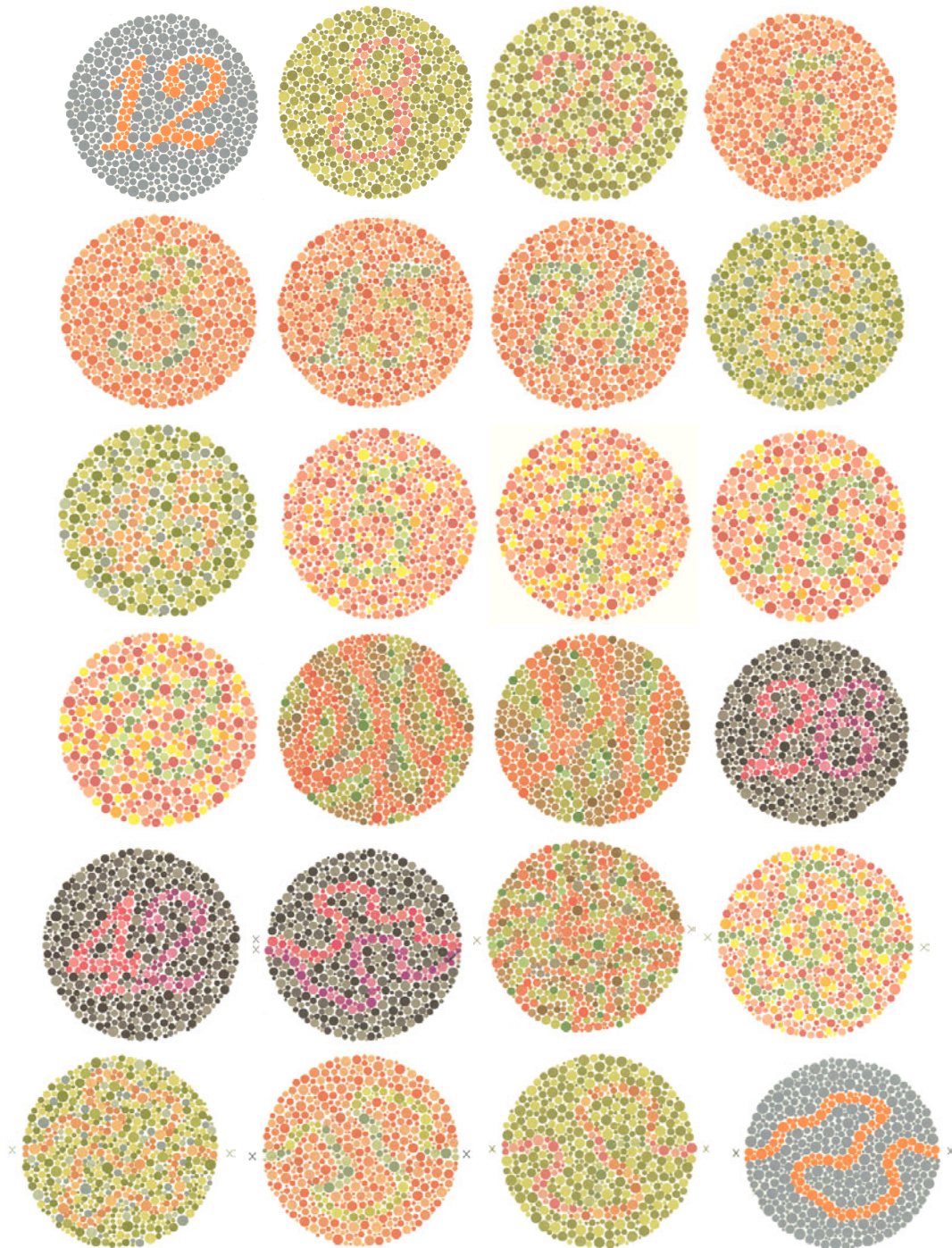


Figure 7.2: Ishihara 24 colour plates. x indicates start and end points of lines that observers may or may not trace depending on their visual function.

7.3.2 Hypothesis

As with the hypothesis for the visual acuity predictor (see Section 5.5.2) both colour stability hypothesis are linked, such that proving the second hypothesis proves the first. Identifying the location of perceptual stability requires that the degree of perceptual stability measure is effective.

7.3.2.1 Hypothesis 1: Degree of Perceptual Stability (PER)

The first hypothesis is that the STRICODI algorithm with $PERS_{cp}$ successfully generates meaningful $PERS_{cp}$ scores. Meaningful is defined as agreeing with human judgements. Human judgements are known for Ishihara eye charts, therefore the $PERS_{cp}$ scores should correctly score the perceptual stability of eye charts.

7.3.2.2 Hypothesis 2: Location of Perceptual Stability (LocPER)

The second hypothesis is linked to the first, in that the STRICODI algorithm with $PERS_{cp}$ will successfully make predictions about the location of perceptual stability on eye charts. Success is where the predictions are consistent with human judgements, that is regions of eye charts which people find perceptually unstable are the same as what STRICODI with $PERS_{cp}$ finds perceptually unstable.

7.3.3 Experiment Design

To test Hypothesis 2, and indirectly validate Hypothesis 1, the experiments are structured as follows:

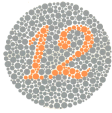
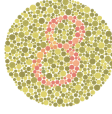
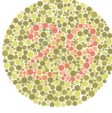
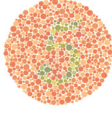
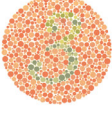
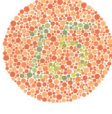
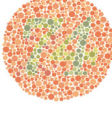
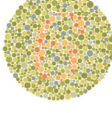
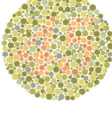
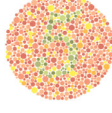
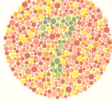
1. eye charts are used as stimuli

2. eye charts are divided into multiple regions
3. regions are equally sized within an eye chart
4. two types of colour deficient vision are simulated, i.e. protanopia and deuteranopia
5. each virtual eye is a subject in the experiment
6. every virtual eye generates simulated retinal images
7. ΔE_{00}^* for each region in the simulated retinal images is measured
8. all the ΔE_{00}^* measures are used to determine perceptual stability change maps
9. each perceptual stability change map is used to generate a perceptual stability score $PERS_{cp}$
10. for each virtual eye the $PERS_{cp}$ score for each region is calculated
11. region $PERS_{cp}$ scores are compared with human rankings to establish whether they concur.

7.3.3.1 Details

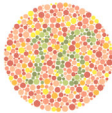
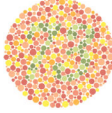
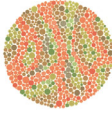
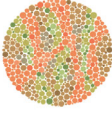
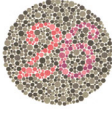
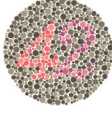
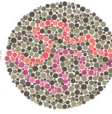
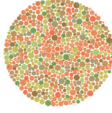
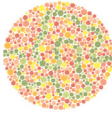
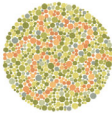
Each eye chart is independently used as a stimulus, and each eye chart is divided into an equal range of regions. Table 7.1 lists how the optotypes and lines in each chart are seen by human perceivers with normal, protanope and deuteranope colour perception.

Table 7.1: How optotypes and lines on Ishihara eye charts are seen by protanopes and deuteranopes.

Ishihara Plate	Eye Chart	Normal	Protanope	Deuteranope
Plate 1		12	12	12
Plate 2		8	3	3
Plate 3		29	70	70
Plate 4		5	2	2
Plate 5		3	5	5
Plate 6		15	17	17
Plate 7		74	21	21
Plate 8		6	Nothing	Nothing
Plate 9		45	Nothing	Nothing
Plate 10		5	Nothing	Nothing
Plate 11		7	Nothing	Nothing

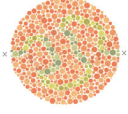
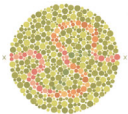
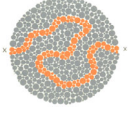
Continued on Next Page...

Table 7.1 – Continued

Experiment	Eye Chart	Normal	Protanope	Deuternope
Plate 12		16	Nothing	Nothing
Plate 13		73	Nothing	Nothing
Plate 14		Nothing	5	5
Plate 15		Nothing	45	45
Plate 16		26	6	2
Plate 17		42	2	4
Plate 18		Both Lines	Bottom Line	Top Line
Plate 19		Nothing	Line	Line
Plate 20		Line	Nothing	Nothing
Plate 21		Line	Nothing	Nothing

Continued on Next Page...

Table 7.1 – Continued

Experiment	Eye Chart	Normal	Protanope	Deuternope
Plate 22		Yellow- Green / Blue- Green Line	Blue- Green	Blue-Green
Plate 23		Purple / Orange Line	Purple / Blue- Green Line	Purple / Blue-Green Line
Plate 24		Line	Line	Line

7.3.3.2 Expected Outcome

If Hypothesis 2 and Hypothesis 1 are true then the regions that undergo the most perceptual change correspond to areas occupied by the optotypes and lines in the eye charts. That is, regions which appear perceptually different between normal and colour deficient perceivers should have the highest $PERS_{cp}$ scores.

7.3.4 Experiment 5: Protanopia

In the first experiment protanopia is simulated, and the Ishihara plates are divided into 20 by 20 regions. Figure 7.3 shows the 24 Ishihara plates as seen by protanopes, and Figure 7.5 shows the resulting visual heatmaps generated

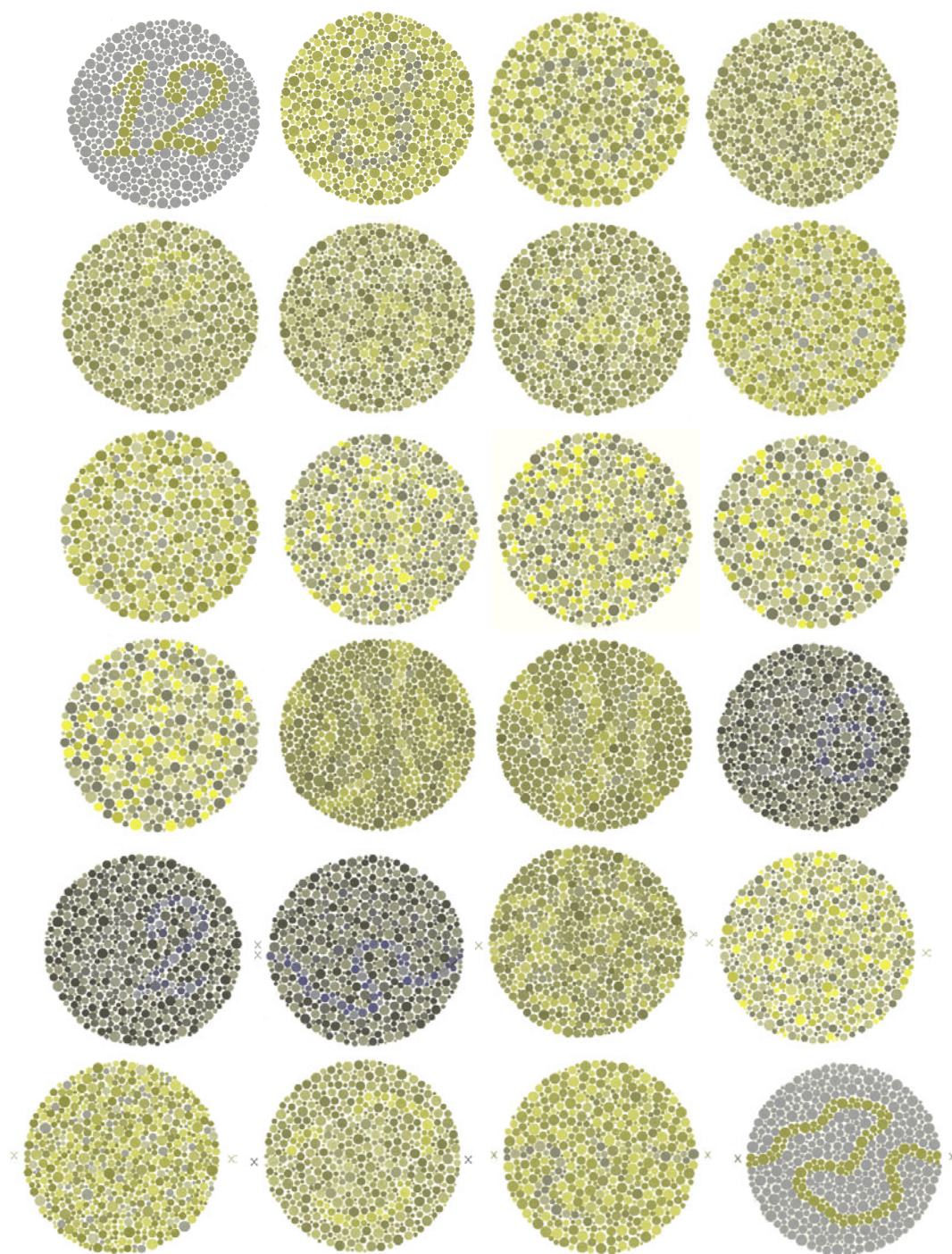


Figure 7.3: Ishihara 24 colour plates as seen with protanopia.

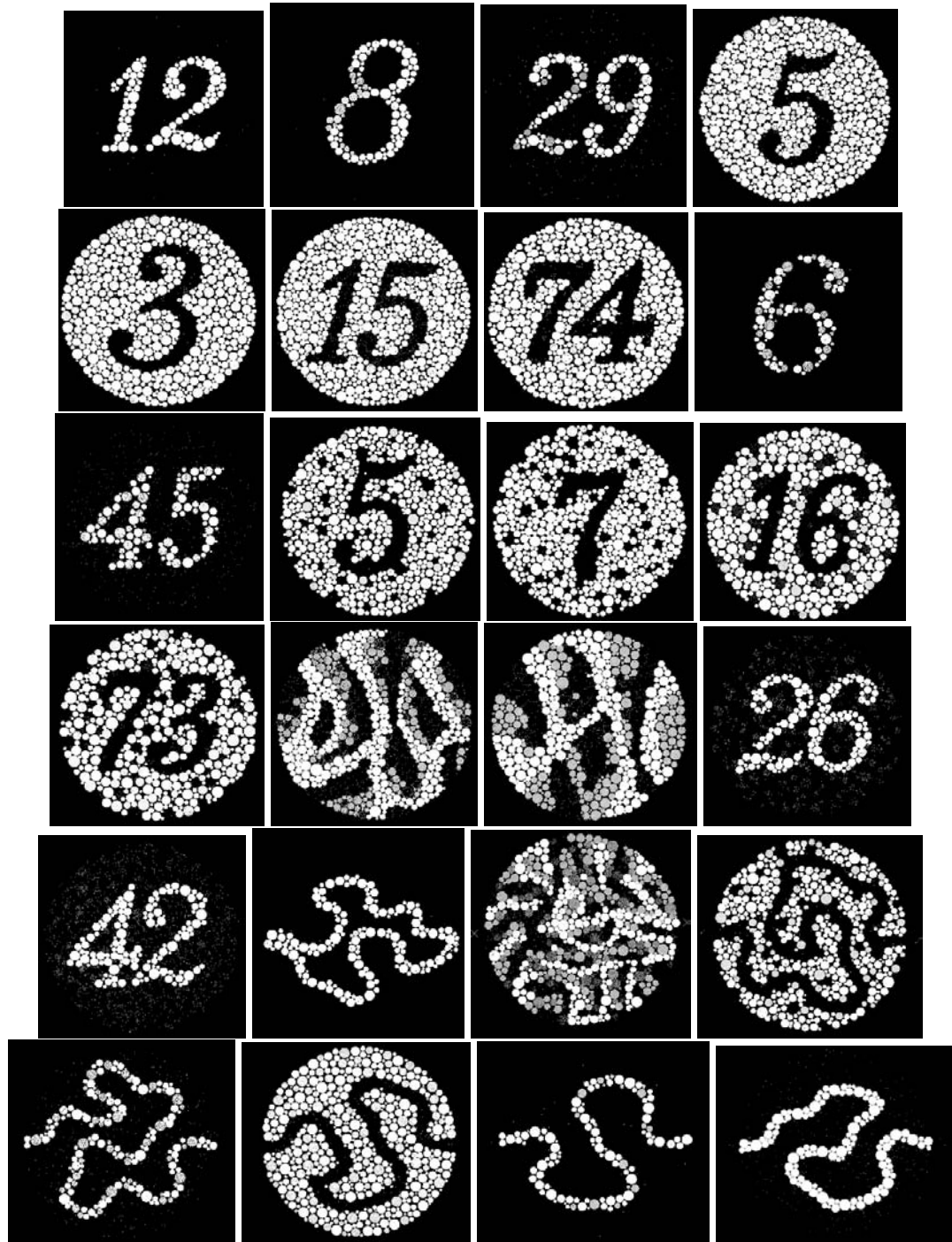


Figure 7.4: Ishihara 24 colour plates with perceptually unstable regions for protanopia highlighted. Highlighted regions have ΔE range between 10 and 20. Change map values were adjusted to range between 0 and 255 for visual clarity.



Figure 7.5: Heatmaps showing perceptual stability of Ishihara 24 colour plates when viewed by protanopes. Plates are divided into 20 by 20 regions. White means region has lowest $PERS_{cp}$ score, while black means region has highest $PERS_{cp}$ score.

by the $PERS_{cp}$ scores for each Ishihara plate.

As can be seen by the Figure 7.3 plates not all visual features expected for protanopes appear, i.e. the optotypes on plates 14, 15 and 19 are not appearing as expected for protanopia. This tells us that the simulation of colour deficit vision is imperfect, though it functions adequately as the optotypes on many of the other plates appear as expected.

The next step in the experiment compares the colour differences between the original Ishihara plates and the protanopia versions. Figure 7.4 shows the perceptual change maps generated by this step.

Next the $PERS_{cp}$ score for each region of each Ishihara plate is used to generate heatmaps, which are shown in Figure 7.5. A black region is perceptually unstable, while a white region is perceptually stable.

From Figure 7.5 we see that colour perceptual stability occurs in three classes for the Ishihara plates. The first class consists of plates 1, 2, 3, 8, 9, 16, 17, 18, 21, 23 and 24. The second class consists of plates 4, 5, 6, 7, 10, 11, 12, 13, 20, 22 and the third class is comprised of plates 14, 15, 19.

Plates in the third class rely on visual features appearing to colour deficient observers which normal observers cannot see (see Table 7.1). As can be seen in both Figure 7.4 and Figure 7.5 visual features which can only be seen by those with colour deficient vision are not detected. This suggests that a limitation of our STRICODI with $PERS_{cp}$ measures is that they cannot detect when visual features will be seen only by colour deficient observers versus normal observers.

Many of the plates in Table 7.1 have alternative choice optotypes (see Section 7.1) where for example optotypes appear as a 3 for a normal observer and as a 5 for a colour deficient observer. As with the third class of plates only perceptual changes from what a normal observer would see are detected in

alternative choice optotypes. Features which appear only to colour deficient observers are not detected with STRICODI and $PERS_{cp}$.

For the first class of plates the optotypes are perceptually unstable as they change colour when viewed with colour deficient vision, while for the second class of plates that which surrounds the optotypes changes colour but the optotypes themselves do not significantly change colour. Therefore due to a foreground colour change the first class is perceptually unstable, while the second class is perceptually unstable due to a background colour change.

The heatmaps in Figure 7.5 show that STRICODI with $PERS_{cp}$ successfully identifies the regions that are perceptually unstable in the first and second class of Ishihara plates. As with the third class of plates and the alternative choice optotypes we find that STRICODI with $PERS_{cp}$ does not identify perceptual instabilities caused by visual features only appearing to colour deficient observers.

7.3.4.1 Foreground versus background optotypes

The foreground versus background colour changes in the first and second class of plates raise interesting questions and possibilities for future work.

Currently STRICODI with $PERS_{cp}$ ranks the plates which experience background change as more perceptually unstable, as more of the colours in the plate change for colour deficient vision.

For example, consider the experience of comparing plate 4 of the normal (see Figure 7.2) and colour deficient plates (see Figure 7.3). In plate 4 the background colour changes and the 5 optotype disappears because it now shares the same colour as the background. From a human perspective the 5 optotype seems to disappear, that is the visually salient feature is the 5 optotype rather than the background colour.

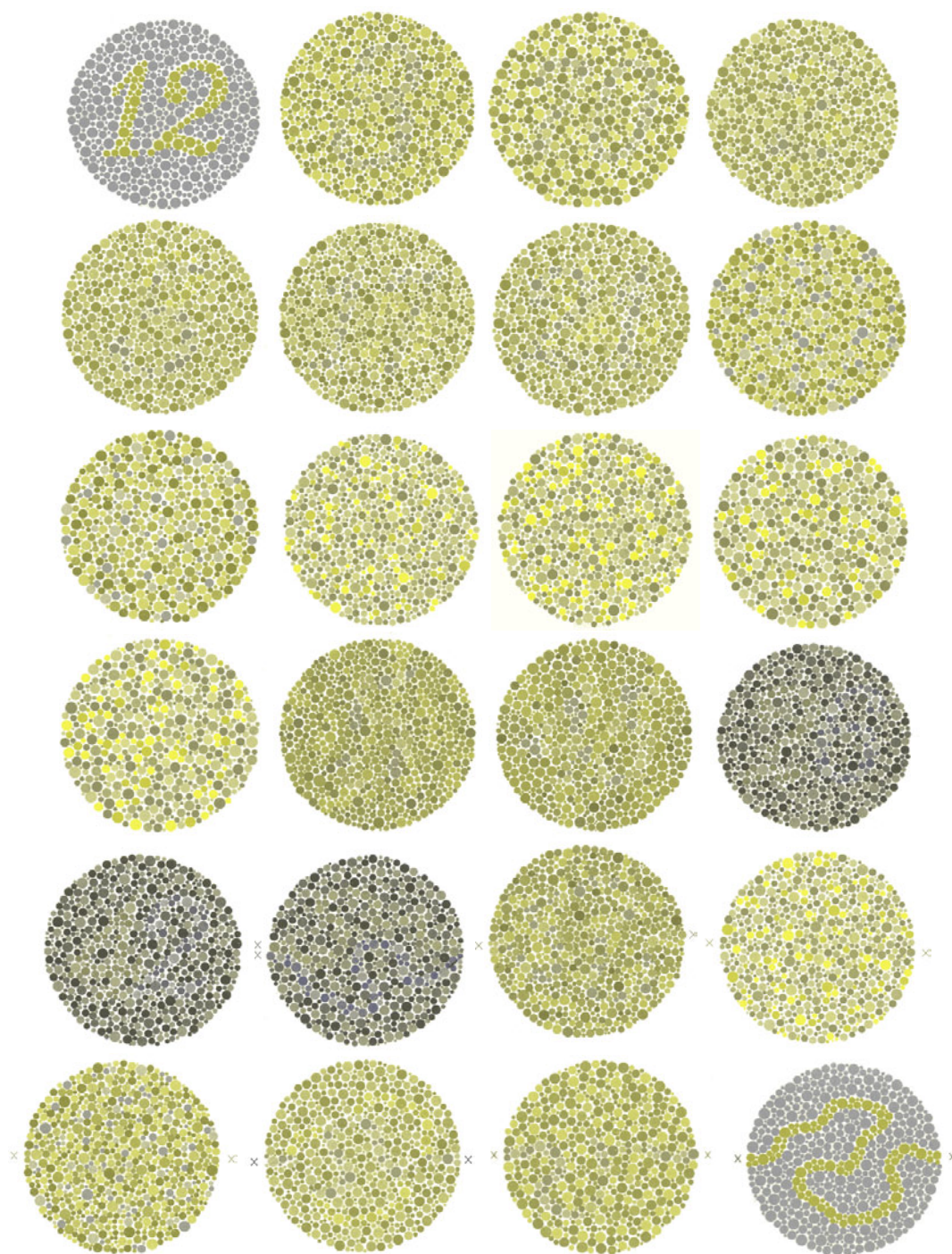


Figure 7.6: Ishihara 24 colour plates as seen with deuteranopia.

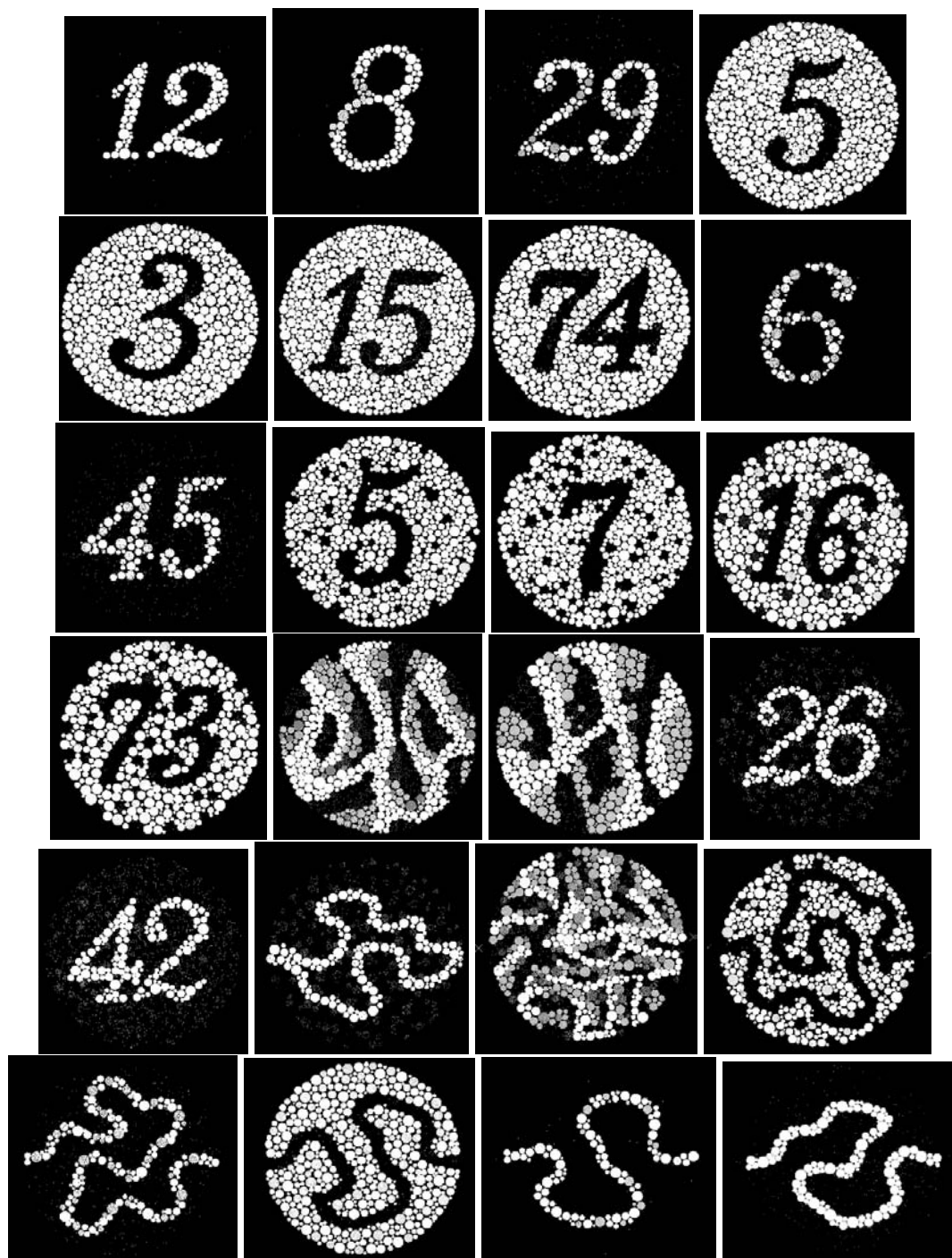


Figure 7.7: Ishihara 24 colour plates with perceptually unstable regions for deuteranopia highlighted. Highlighted regions have ΔE range between 10 and 20. Change maps values were adjusted to range between 0 and 255 for visual clarity.



Figure 7.8: Heatmaps showing perceptual stability of Ishihara 24 colour plates when viewed by deuteranopes. Plates are divided into 20 by 20 regions. White means region has lowest $PERS_{cp}$ score, while black means region has highest $PERS_{cp}$ score.

7.3.5 Experiment 6: Deuteranopia

In the second experiment deuteranopia is simulated, and the Ishihara plates are divided into 20 by 20 regions.

Figure 7.6 shows the 24 Ishihara plates as seen by deuteranopes, and Figure 7.8 shows the resulting heatmaps generated by the $PERS_{cp}$ scores for each Ishihara plate. Figure 7.7 shows the perceptual change maps.

The analysis and results for Experiment 6 are exactly the same as Experiment 5 (see Section 7.3.4). That is, STRICODI with $PERS_{cp}$ does detect perceptual stability changes caused by colour deficient reductions in colour spaces.

An additional finding is that for all the plates we found that STRICODI with $PERS_{cp}$ showed there are few perceptual differences between the plates when compared between protanopes and deuteranopes, i.e. the protanope heatmaps in Figure 7.5 closely resemble the deuteranope heatmaps in Figure 7.8. This result agrees with the diagnostic use of the plates. Many of the optotypes and lines on the plates are perceived equivalently across protanopes and deuteranopes (see Table 7.1).

As with Experiment 5 we find that STRICODI with $PERS_{cp}$ does not detect perceptual instabilities caused by visual features that can be seen by colour deficient observers but not by normal observers. An implication of this is that $PERS_{cp}$ cannot distinguish between protanopia and deuteranopia on the 24 Ishihara plate test. This is because plates 16, 17 and 18 which diagnostically distinguish between protanopia and deuteranopia, rely on features that only colour deficient observers can see.

7.4 Conclusion

The experiments in this Chapter have shown that STRICODI with $PERS_{cp}$ successfully agrees with human judgements for most of the Ishihara eye charts. $PERS_{cp}$ scores are as expected (Hypothesis 1) across the range of varied coloured Ishihara eye charts. In most cases the region $PERS_{cp}$ are in agreement with human judgments (Hypothesis 2).

Areas that emerge as weak are the inability to detect perceptual stabilities due to features only seen by colour deficient observers.

Further improvements may be possible by attempting to take visual saliency into consideration, to detect perceptual instabilities caused by foreground versus background colour changes.

Adapting Visual Designs To Individual Differences

This chapter demonstrates how our visual acuity measure and colour perception measure can be used to evaluate and adapt designs to suit individual's eye function.

These adaptations show the strengths and weaknesses of the predictors, along with how they can be used in visual design tools and automated layout techniques for interfaces and information visualisations.

Demonstrated in this Chapter are automatic evaluations of text and font styles, network graph designs and layouts, and the pseudocolouring of scientific visualisations.

8.1 Adaptions Demonstrated

A range of adaptations are presented that demonstrate the diversity of uses for the perceptual stability metrics. By focusing on a diverse range of interface and information visualisation adaptations the effectiveness and strengths of the metrics become clear.

The following adaptations are demonstrated:

- Font text adaptations for visual acuity (see Section 8.3.1)
- Network graph adaptations for visual acuity (see Section 8.3.2)
- Map colour scale adaptations for colour perception (see Section 8.4.1)

Text blocks and font styles are an important basic building block in nearly all interfaces and many visual designs. By demonstrating that $PERS_{va}$ is applicable to analysing text styles we show it can be widely used to analyse the text displayed in interfaces and information visualisations.

Legibility and aesthetics are an important criteria when designing effective network graph visualisations. Many of the existing techniques for laying out network graphs rely on counting and minimizing edge crossing, while maintaining distances between network graph components. In this thesis we demonstrate a complementary approach for analysing and improving network graph layouts, which is based on physiological models of visual function (see Chapter 3).

Maps and many information visualisations rely on colour to create visually distinct groups and categories. A continuous colour gradient scale is commonly used to pseudocolour maps. In this Chapter we demonstrate how the colour stability measure can aid in selecting colour sequences that are effective for those with colour deficient vision.

8.2 Physiological Versus Semantic Adaptions

Particular relevant for the effective use of the metrics is that they evaluate designs from a physiological rather than a semantic perspective. For example consider two vertical side by side black lines, as shown in Figure 8.1. An auto-



Figure 8.1: Example of how the $PERS_{va}$ score is a physiological rather than semantic metric. Each Group of vertical bars is separately analysed to calculate $PERS_{va}$. $PERS_{va}$ scores from Group 1 to Group 4 are 0.9225, 0.9251, 0.9903 and 0.9337; averaged over 20 randomly generated virtual eyes. The two vertical bars on the left hand side, Group 1, are the most perceptually stable with the lowest $PERS_{va}$ score, while Group 4 on the right has the third lowest $PERS_{va}$ score. Making Group 4 more perceptually stable than Group 3.

mated layout technique which uses $PERS_{va}$ as a metric would find that when the two lines move very close together, as in Group 4 Figure 8.1, the $PERS_{va}$ score improves (lower = more perceptually stable).

As the two lines move closer together the $PERS_{va}$ perceptual stability score changes, with the perceptual stability decreasing from Group 1 ($PERS_{va}$ 0.9225) to Group 2 ($PERS_{va}$ 0.9251) to Group 3 ($PERS_{va}$ 0.9903). In Group 4 ($PERS_{va}$ 0.9337) the lines are the closest of all the Groups yet the $PERS_{va}$ score is considerably better than for Group 3. Since the lines in Group 4 are so close it is reasonable to assume the lines are the least perceptually stable of all Groups, yet that is not the result $PERS_{va}$ gives.

The reason is that $PERS_{va}$ evaluates whether a visual feature is more perceptually stable than another. In the case of Group 4, when the two lines undergo a small amount of optical aberrations they form a single visual feature, while in the other Groups the two lines maintain their existence as two separate visual features (see Figure 8.2) when subject to more optical aberration.

Figure 8.2 visually demonstrates how quickly the two lines in Group 4 form a single larger line, when subject to a small amount of optical aberrations. Optical aberrations were simulated with Zernike mode Z_2^0 , and the Zernike coef-

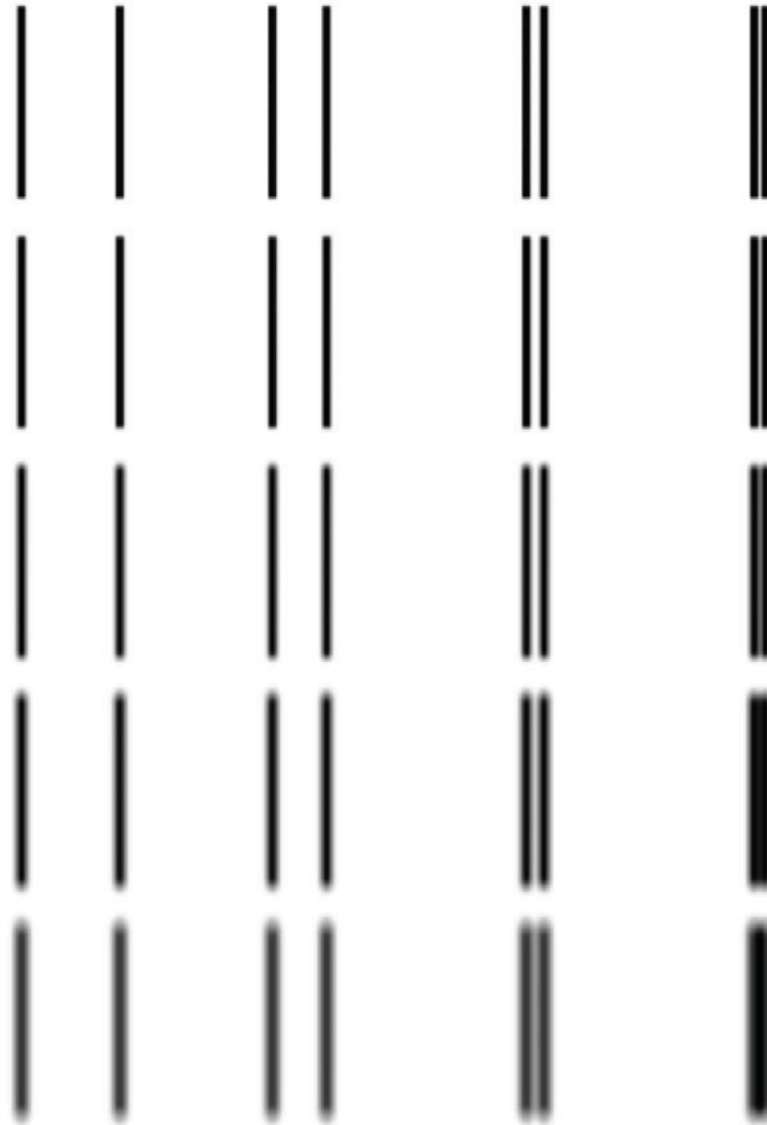


Figure 8.2: Effect of Z_2^0 optical aberration on vertical bars set at different distances from each other. Vertical bars in Group 4 (see Figure 8.1) quickly form a single visual feature when subject to a low amount of optical aberrations. Each row corresponds to a 0.02 increase in the Zernike coefficient. Initial Zernike coefficient is 0.

ficient ranged between 0.0 and 0.10. Each row in Figure 8.2 represents a 0.02 increase in the Zernike coefficient.

From a semantic perspective we know there are two lines in Group 4, yet from an aberrant optical system perspective the two lines become one when undergoing small amounts of optical aberrations. Therefore for Group 4 $PERS_{va}$ is effectively evaluating one line, which is seen as double the width of the other single lines. While this behaviour of $PERS_{va}$ is non-obvious and somewhat counter-intuitive it is to be optically expected [51], i.e. when visual features in close proximity undergo optical aberrations they form larger less detailed visual features.

An awareness of how visual features behave when subject to optical aberrations is important for realising the appropriate use and interpretation of $PERS_{va}$, especially due to $PERS_{va}$ functioning at a physiological low-level in vision rather than at a semantic level of visual feature recognition.

8.3 Visual Acuity Adaptions

When a designer is laying out interfaces and text they implicitly make tradeoffs about font size, layout and colouring to optimise for clarity, legibility and audience attention.

In the following Sections we demonstrate how the perceptually stability measures are used to automate design decisions. Designs are adapted for perceptually stability, with the benefit of improving the designs for differences in visual acuity and colour perception, i.e. should font A or font B be used for a block of text, or what graph design node is easier to see?

8.3.1 Text & Font Style

Below are two demonstrations where two different fonts at three sizes are tested to establish which font is more perceptually stable for visual acuity. The Times-Roman and Courier-Bold fonts are compared. Times-Roman is a proportional font, while Courier-Bold is a fixed width font.

Elsewhere [56] the fonts have been empirically measured to test which of the fonts is more legible, for both normal and low-vision observers. These findings indicate that Courier-Bold is easier to see for both normal and low-vision observers, though in certain cases reading speed of normal observers is higher for Times-Roman. $PERS_{va}$ cannot make predictions about reading speed, but its predictions do agree with ranking Courier-Bold as more legible than Times-Roman.

Shown in Figure 8.3 and Figure 8.7 are the two blocks of text, which $PERS_{va}$ evaluates to establish which font style is more perceptually stable.

8.3.1.1 English Lower-Case Alphabet

The first text block consists of the lower-case letters from the English alphabet ordered in a row, with each letter separated by a single space (see Figure 8.3). Rows 1 to 3 are in Courier-Bold font, and Rows 4 to 6 are in Times-Roman font. Rows 1 and 4 have a font size of 11 pt, Rows 2 and 5 are 16 pt and Rows 3 and 6 are 21 pt.

To analyse the rows 20 normal virtual eyes are randomly created, and each virtual eye is used in STRICODI for examining each row (see Section 5.4). For each virtual eye the Zernike coefficients are incremented in 20 steps from 0 to maximum RMS. Results of the analysis are plotted and graphed in Figure 8.4, Figure 8.5 and Figure 8.6.



Figure 8.3: Image showing the English alphabet lower-case letters that are compared with $PERS_{va}$. In the analysed image there is more white space between rows, to prevent interaction between rows undergoing optical aberrations. Rows 1 to 3 are in the Courier-Bold font, and Rows 4 to 6 are in the Times-Roman font. Rows 1 and 4 have a font size of 11 pt, Rows 2 and 5 are 16 pt and Rows 3 and 6 are 21 pt.

Plotted in Figure 8.4 is the mean and standard deviation of the normalised $PERS_{va}$ scores for all 20 virtual eyes. For the Courier-Bold font the mean $PERS_{va}$ is consistently lower across all three font sizes than for Times-Roman. This tells us that the Courier-Bold font is rated as more perceptually stable than the Times-Roman font, which agrees with the findings in [56].

Shown in the scatter plot in Figure 8.5 is the distribution of the Courier-Bold versus Times-Roman $PERS_{va}$ scores. From the scatter plot it can be seen that Courier-Bold is consistently ranked as more perceptually stable than Times-Roman across all 20 virtual eyes.

When the standard deviation in Figure 8.4 is coupled with the scatter plot distribution in Figure 8.5 we find that individual differences in optical aberrations lead to a wider range of $PERS_{va}$ scores for the Courier-Bold font, than for the Times-Roman font.

As is expected, the scatter plot shows that due to individual differences in eye function there is a cross-over between the perceptual stability of font sizes, i.e. some eyes are worse at seeing font size 21 pt than other eyes are at seeing font

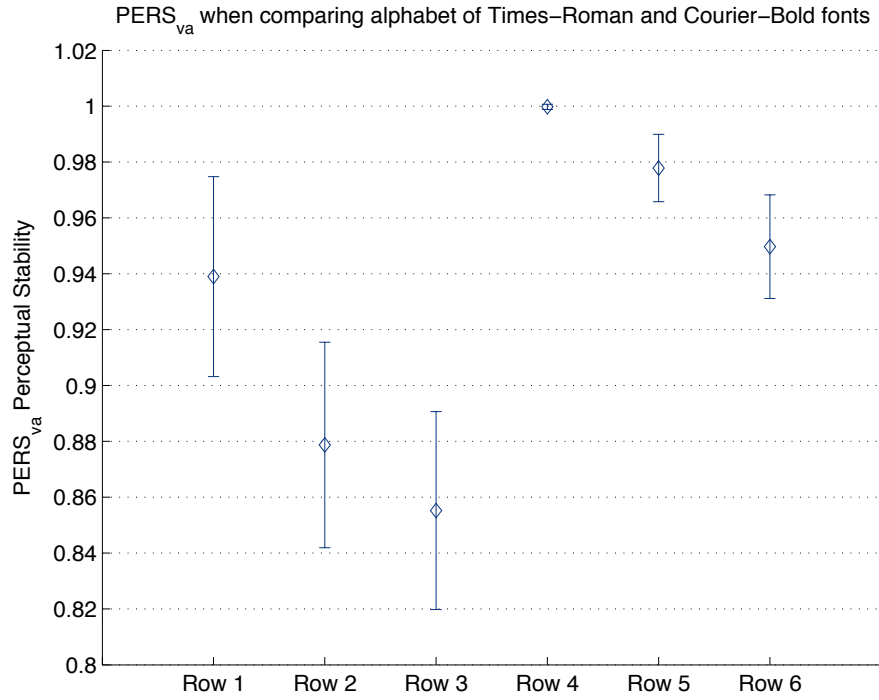


Figure 8.4: Mean and standard deviation of $PERS_{va}$ scores, when analysing the rows of Courier-Bold and Times-Roman text as shown in Figure 8.3.

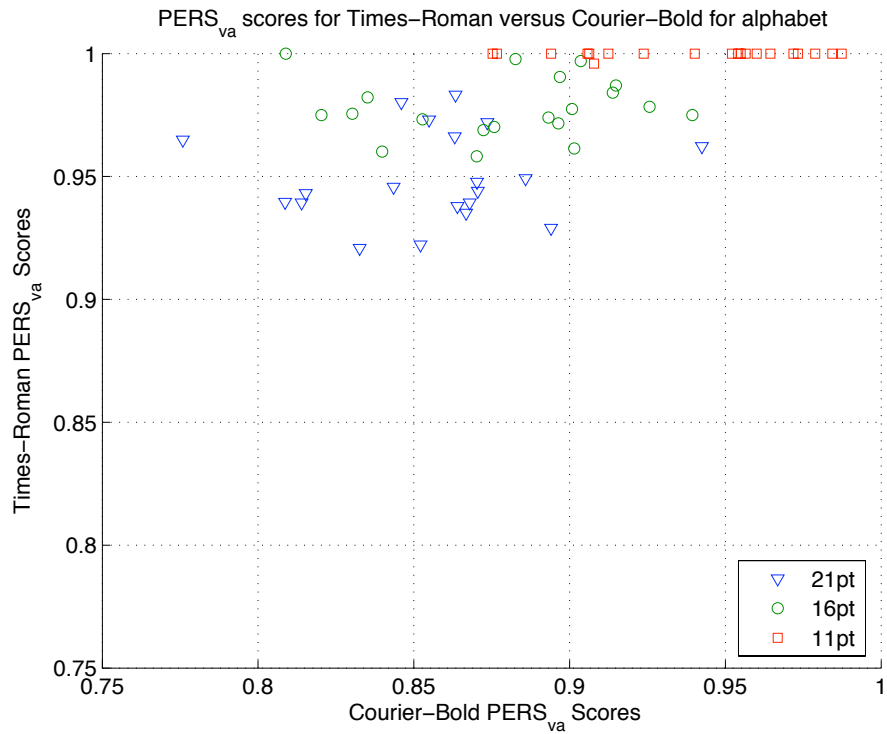


Figure 8.5: Scatter plot distribution of $PERS_{va}$ scores for Times-Roman versus Courier-Bold, when analysing alphabet text shown in Figure 8.3. Each data point is an eye, each data group contains the same set of 20 virtual eyes.

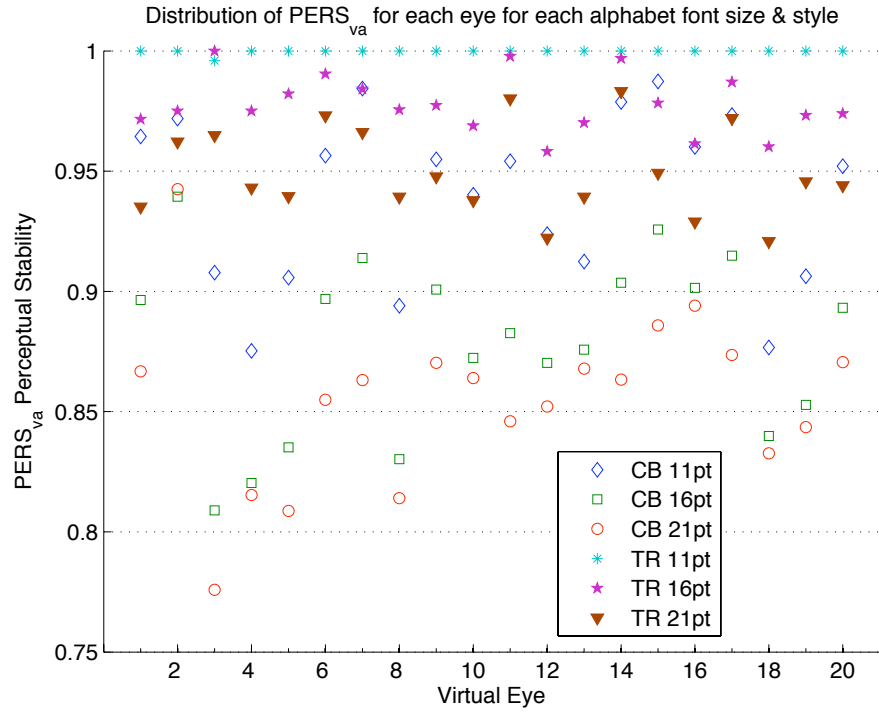


Figure 8.6: Distribution of $PERS_{va}$ score for each virtual eye for each font style and font size, when analysing alphabet text. CB is Courier-Bold and TR is Times-Roman.

size 16 pt.

This demonstration illustrates that $PERS_{va}$ is a capable predictive measure when adapting rows of equal sized text, such that the text is optimised for legibility.

8.3.1.2 Text Paragraph

Analysed in Figure 8.7 is the text paragraph from an earlier work on the intersection of graphical design and perception [24], which reads as follows *"The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process*

of graphical perception”.

Each paragraph is treated as a separate Region. The distances between the Regions are greater in the analysed image - to prevent optical interactions between the Region contents. Regions 1 to 3 are Courier-Bold font, and Regions 4 to 6 are Times-Roman. Regions 1 and 3 are 11 pt, Regions 2 and 5 are 16 pt and Regions 3 and 6 are 21 pt.

For the analysis 20 normal virtual eyes are randomly created, which are different to the 20 virtual eyes created in the previous demonstration. Zernike coefficients are incremented in 20 steps to the maximum RMS for each eye. Results from the analysis are graphed and plotted in Figure 8.8, Figure 8.9 and Figure 8.10.

The results of this demonstration concur with the results of the previous demonstration, that is Courier-Bold is more perceptually stable than Times-Roman across a range of font sizes.

Differences in the results are that the standard deviations in Figure 8.8 are smaller for Courier-Bold and larger for Times-Roman. From the scatter plot in Figure 8.9 it can be seen that Courier-Bold is consistently ranked as more perceptually stable than Times-Roman across all 20 virtual eyes, across all three font sizes.

Interestingly it can be seen that the effects of individual differences in eye function are more pronounced for the Times-Roman text paragraph (see Figure 8.9 and standard deviations of Region 4 to Region 6 in Figure 8.8) than for the Times-Roman alphabet (see Figure 8.5 and the standard deviations of Row 4 to Row 6 in Figure 8.4). This suggests that using Times-Roman to display text, which will be read by an audience with diverse age range and in a diverse range of viewing conditions, may result in widely varied viewing experiences due a wide range of age related optical aberrations.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on graphical perception - the visual decoding of information encoded on graphs - and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception.

Figure 8.7: Image showing the paragraphs of text that were compared with $PERS_{va}$. In the analysed image there is more white space separating Regions. Rows 1 to 3 are in the Courier-Bold font, and Rows 4 to 6 are in the Times-Roman font. Rows 1 and 4 have a font size of 11 pt, Rows 2 and 5 are 16 pt and Rows 3 and 6 are 21 pt.

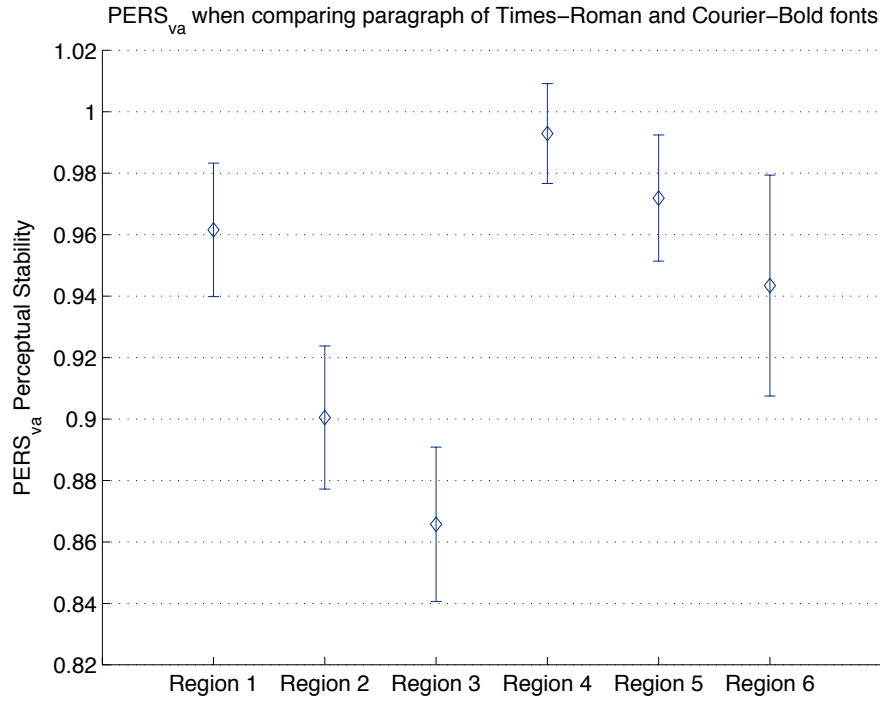


Figure 8.8: Mean and standard deviation of $PERS_{va}$ scores, when analysing the Regions of Courier-Bold and Times-Roman text as shown in Figure 8.7.

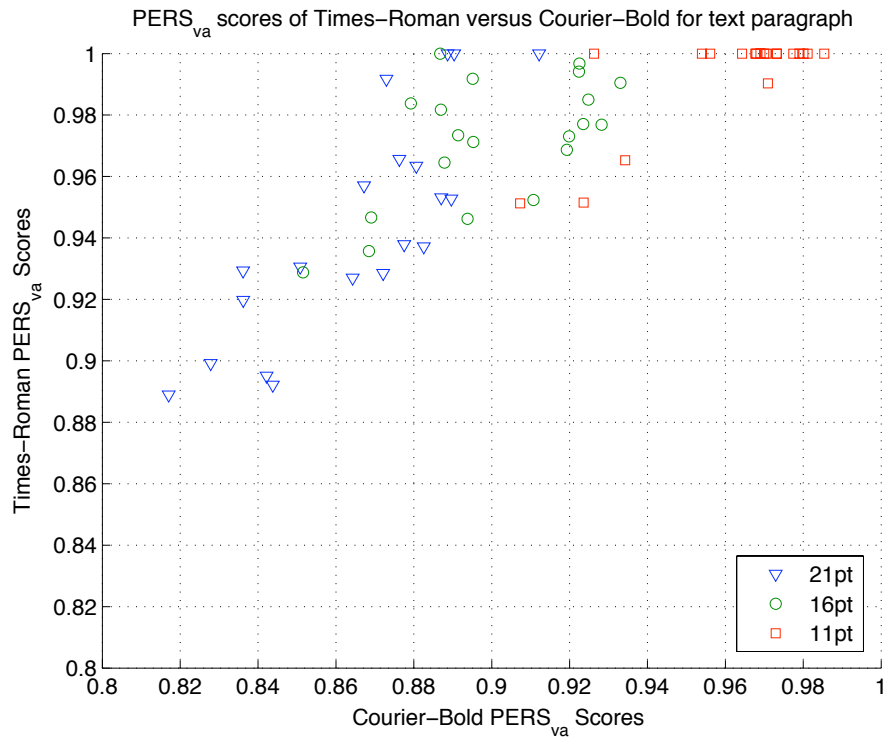


Figure 8.9: Scatter plot distribution of $PERS_{va}$ scores for Times-Roman versus Courier-Bold, when analysing paragraph text shown in Figure 8.7. Each data point is an eye, each data group contains the same set of 20 virtual eyes.

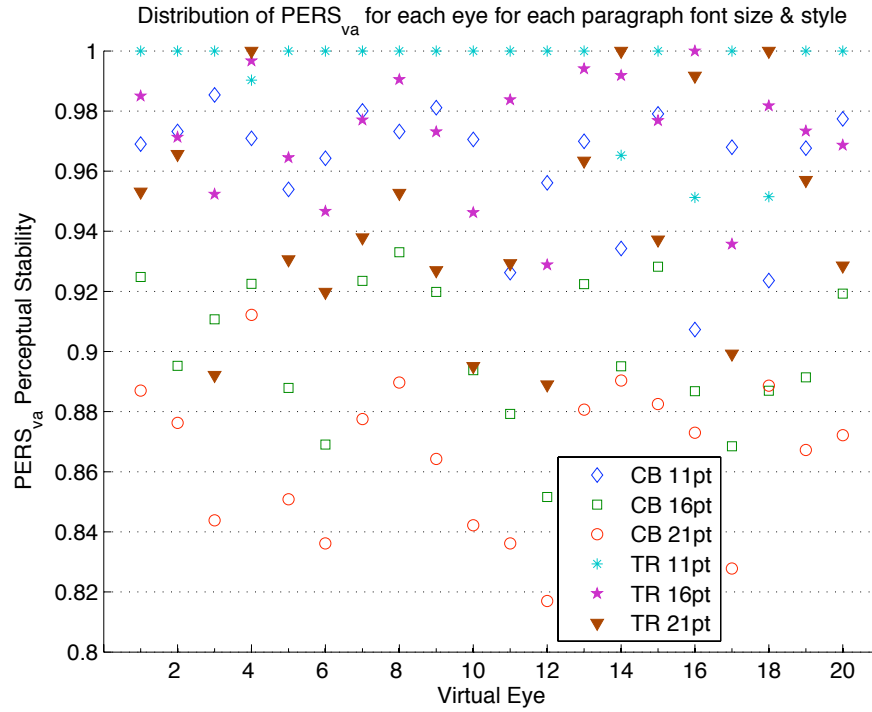


Figure 8.10: Distribution of $PERS_{va}$ score for each virtual eye for each font style and font size, when analysing paragraph text. CB is Courier-Bold and TR is Times-Roman.

As with the previous demonstration, this demonstration shows that $PERS_{va}$ is usable as a design aid when deciding what adaptations to make to font styles.

8.3.1.3 Comparing Across Font Sizes And Font Styles

One issue that does arise from these demonstrations is whether $PERS_{va}$ is suitable for making comparisons when font size and font style vary. For example when displaying a paragraph of text which is better, a Times-Roman 21 pt font or a Courier-Bold 11 pt font?

Base on Figure 8.9 the ranking from best to worst for displaying a paragraph of text is Courier-Bold 21 pt, Courier-Bold 16 pt, Times-Roman 21 pt, Courier-Bold 11 pt, Times-Roman 16 pt and then Times-Roman 11 pt.

The differences in score between Courier-Bold 11 pt and Times-Roman 16 pt are tiny at $\approx 0.01 PERS_{va}$, making it unclear whether Courier-Bold 11 pt is in

Table 8.1: List of network graph design adaptations evaluated with $PERS_{va}$.

Name	Graph	Adaption
Graph Set 1	Figure 8.11	Black node size increases
Graph Set 2	Figure 8.12	Edge width increases
Graph Set 3	Figure 8.13	Hollow node size increases
Graph Set 4	Figure 8.14	Hollow node border width increases
Graph Set 5	Figure 8.23	Maximum variations of Graph Sets 1 to 4

any way significantly more perceptually stable than Times-Roman 16 pt.

As with the two vertical bars in Figure 8.2 (Group 4) the difference in scores between Courier-Bold 11 pt and Times-Roman 16 pt may be due to Courier-Bold 11 pt having lost more detail but forming a large less detailed visual feature. Whether this is the case or not remains an open question for future work.

8.3.2 Network Graphs

For the network graph demonstration five sets of nodes are analysed, the layouts of which are shown in Figure 8.11, Figure 8.12, Figure 8.13, Figure 8.14 and Figure 8.23.

During the multiple analyses a five node circular network graph is measured, to establish which design permutations of the graphs are the most perceptually stable. Listed in Table 8.1 are the set of adaptations of the circular five node network.

For every Graph Set twenty random normal virtual eyes were independently generated. Zernike coefficients went from 0 to maximum RMS, in forty increments.

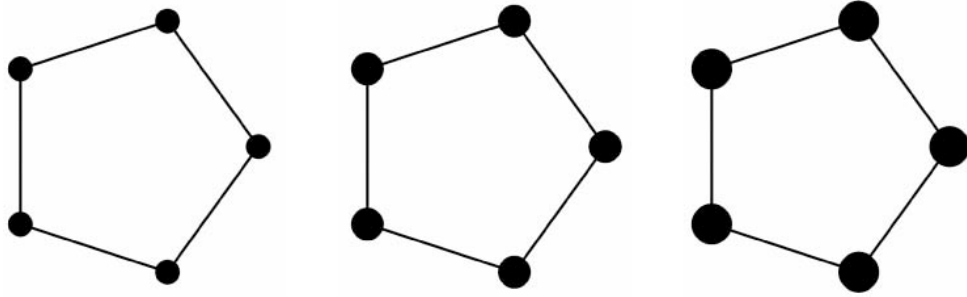


Figure 8.11: Graph Set 1 - Node size increases from left (Region 1) to right (Region 3).

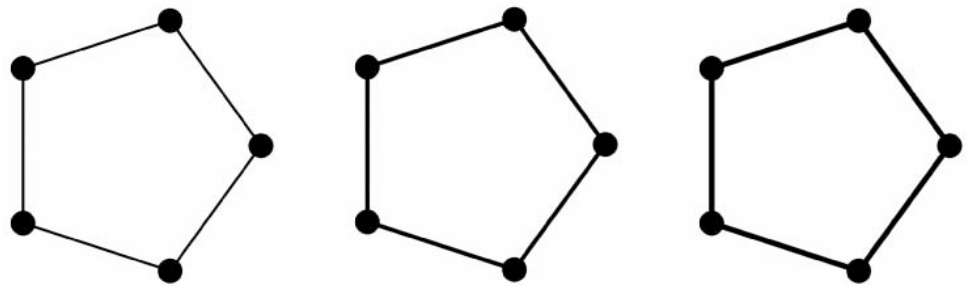


Figure 8.12: Graph Set 2 - Edge width increases from left (Region 1) to right (Region 3).

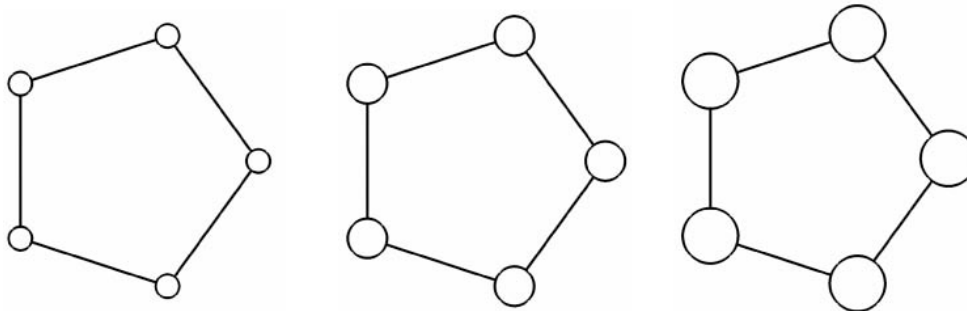


Figure 8.13: Graph Set 3 - Hollow node size increases from left (Region 1) to right (Region 3).

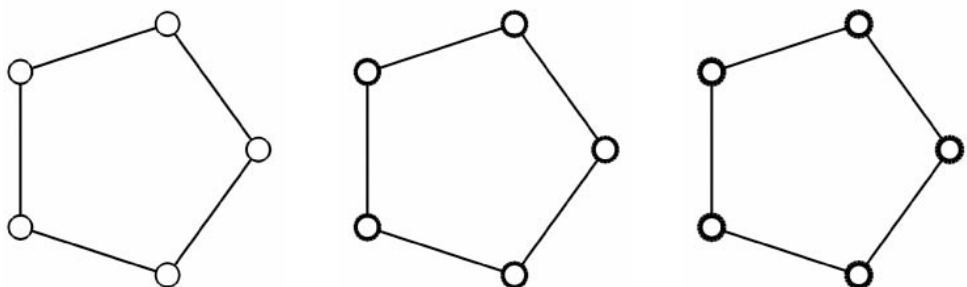


Figure 8.14: Graph Set 4 - Hollow node border width increases from left (Region 1) to right (Region 3).

8.3.2.1 One Dimensional Adaptions

The first four Graph Sets vary by changing one visual property within each set. Each of the Graph Sets 1 to 4 is divided in three regions. Regions are equal in size, independent, and each region covers the area occupied by a single graph network. These regions are numbered from 1 to 3, going from the left to the right.

In Graph Set 1 (see Figure 8.11) the node size increases, the node colour and edge widths stay the same. For Graph Set 2 (see Figure 8.12) the width of the edges increase and the node sizes are constant. With Graph Set 3 (see Figure 8.13) the nodes are made hollow and increase in size, with a border width equal to the edge width. In Graph Set 4 the nodes are hollow while the node border widths increase.

The results for Graph Set 1 are plotted in Figure 8.15, which shows the mean and standard deviation of $PERS_{va}$ for each region. Based on these results Region 3 is the most perceptually stable, with Region 2 ranked next and Region 1 the least perceptually stable. Differences in $PERS_{va}$ scores for each of the 20 virtual eyes for each of the Regions is in Figure 8.16. These results indicate that bigger nodes sizes are better.

For Graph Set 2 the mean and standard deviation is plotted in Figure 8.17. Yet again Region 3 is the most perceptually stable, followed by Region 2 and then Region 1. Figure 8.18 shows the individual variability of the $PERS_{va}$ scores for each virtual eye. From these results we learn that a wider edge is better than a thinner edge.

The $PERS_{va}$ mean and standard deviation for Graph Set 3 is shown in Figure 8.19. There is a tiny 0.0031 difference in mean $PERS_{va}$ scores between the Regions, along with very small standard deviations. Figure 8.20 shows the individual variability of $PERS_{va}$ for each virtual eye. From these results it

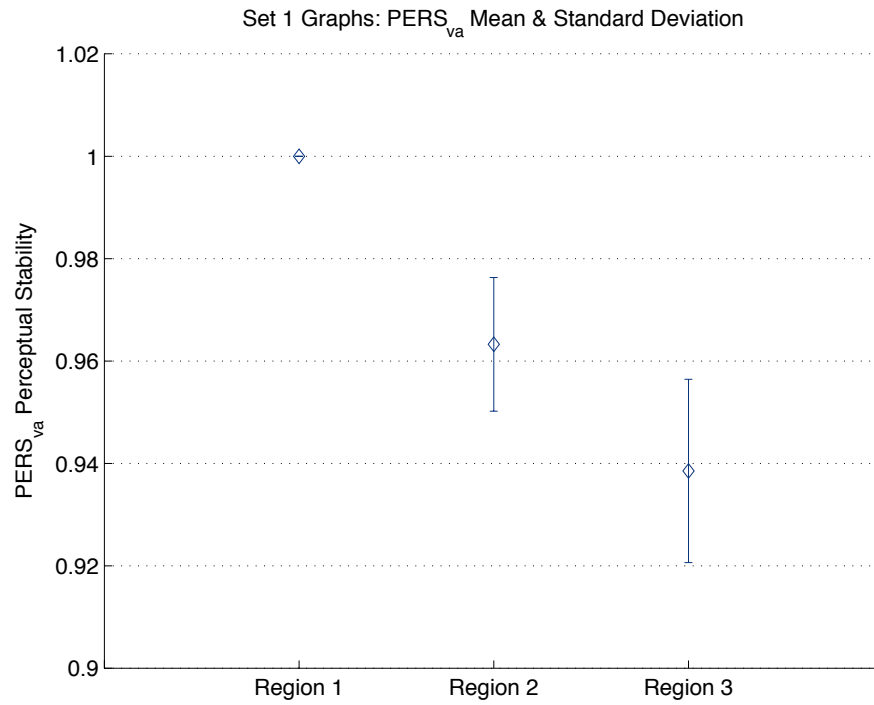


Figure 8.15: Graph Set 1 - $PERS_{va}$ mean and standard deviation.

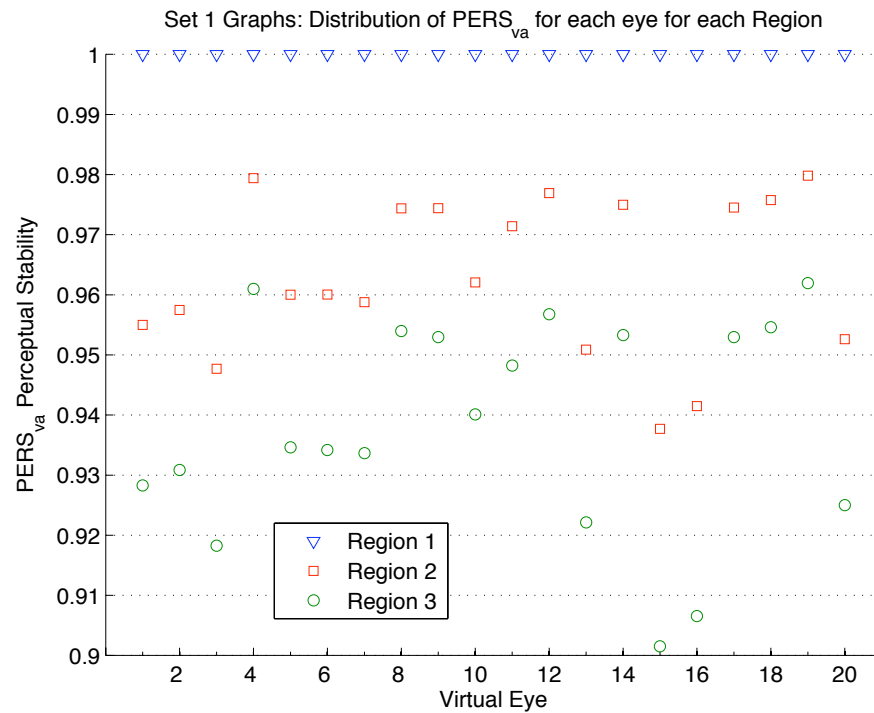


Figure 8.16: Graph Set 1 - Distribution of $PERS_{va}$ for each virtual eye for each region.

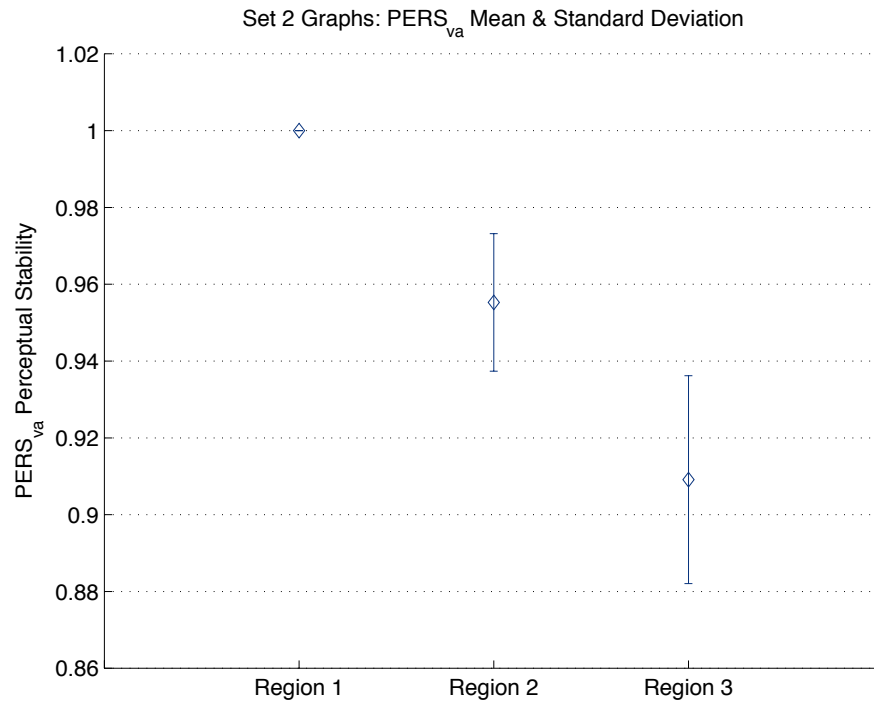


Figure 8.17: Graph Set 2 - $PERS_{va}$ mean and standard deviation.

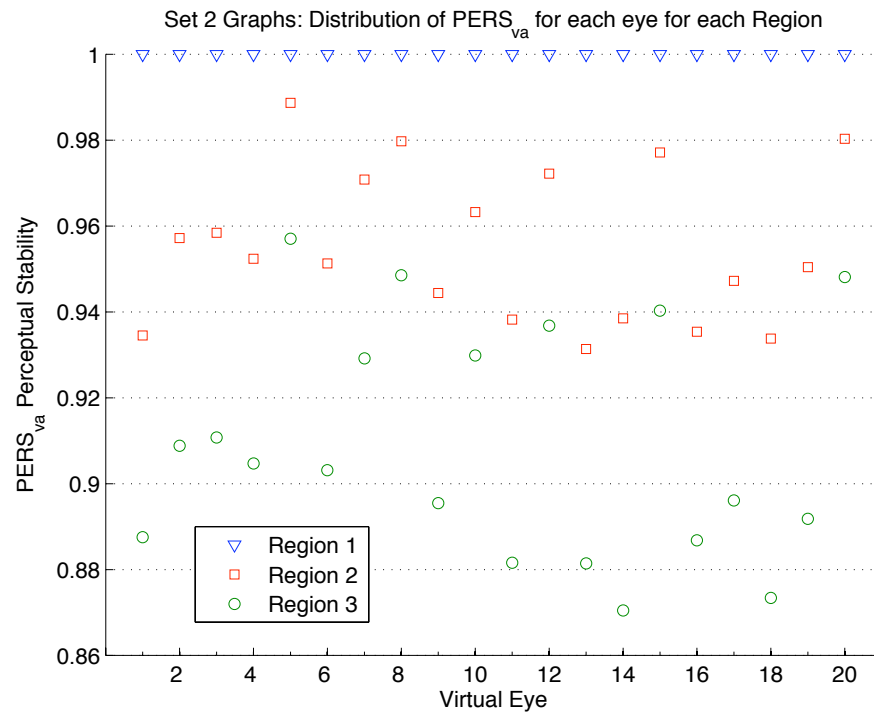


Figure 8.18: Graph Set 2 - Distribution of $PERS_{va}$ for each virtual eye for each region.

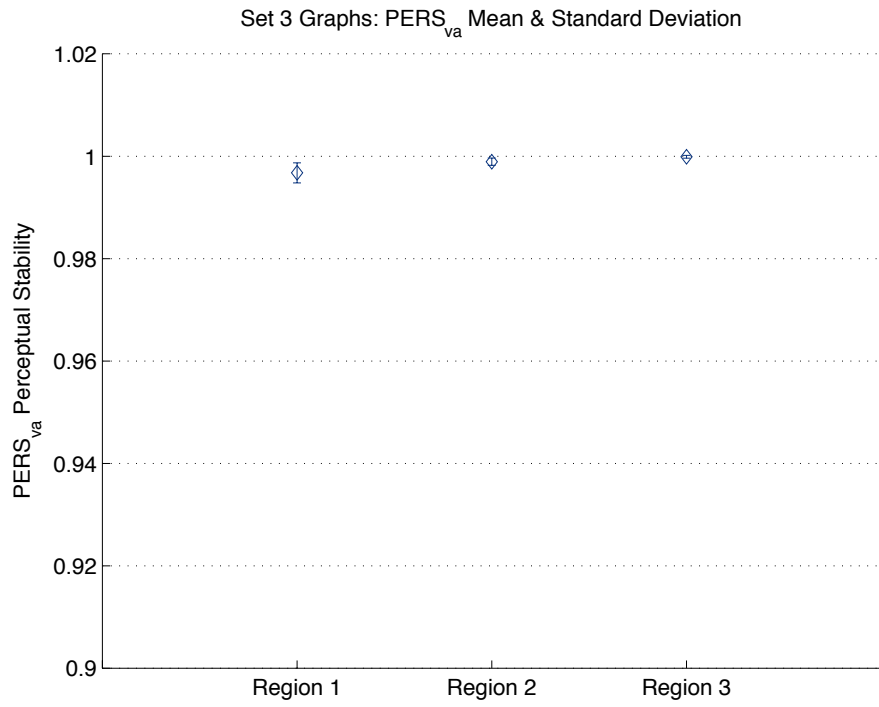


Figure 8.19: Graph Set 3 - $PERS_{va}$ mean and standard deviation.

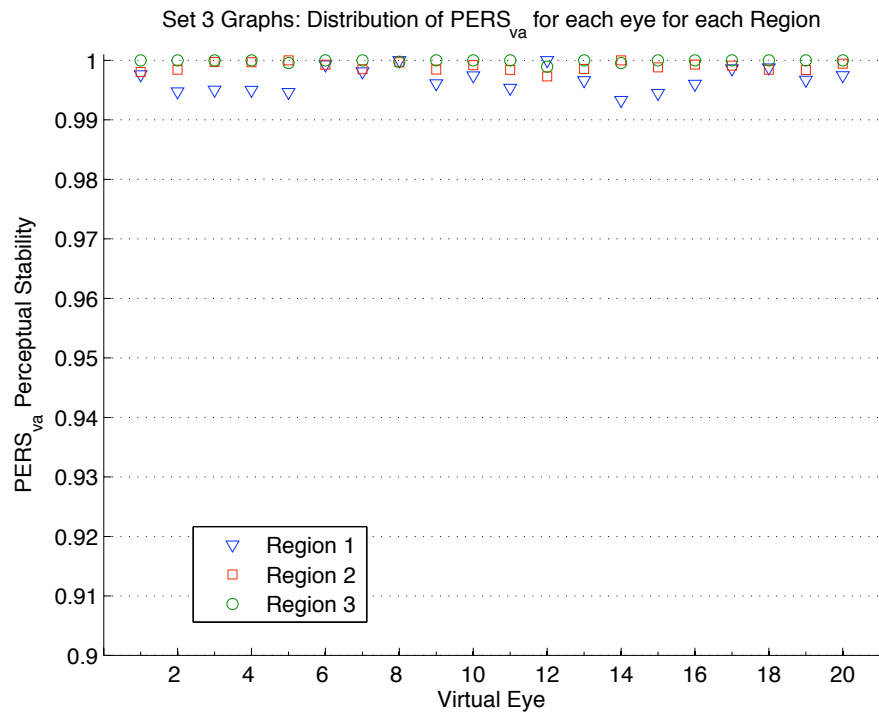


Figure 8.20: Graph Set 3 - Distribution of $PERS_{va}$ for each virtual eye for each region.

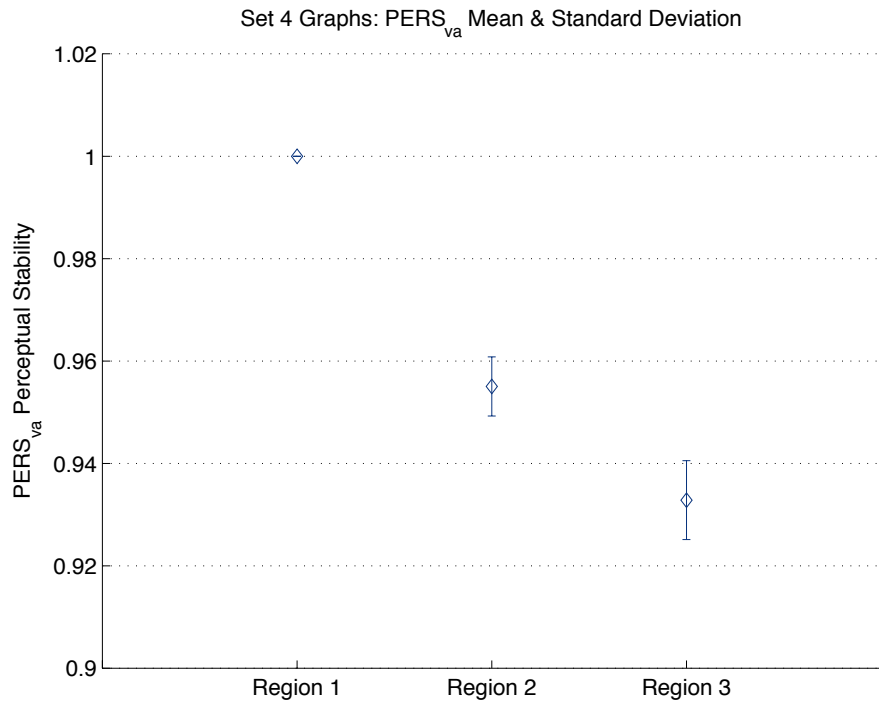


Figure 8.21: Graph Set 4 - $PERS_{va}$ mean and standard deviation.

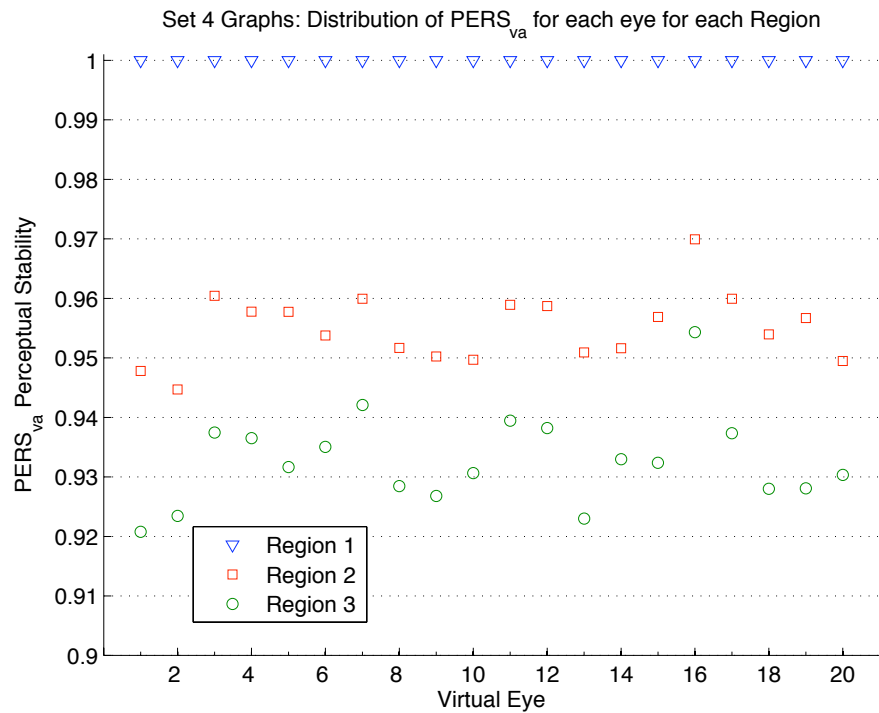


Figure 8.22: Graph Set 4 - Distribution of $PERS_{va}$ for each virtual eye for each region.

is unclear whether a larger hollow node is more perceptually stable than a smaller hollow node.

The result in Graph Set 3 is not inaccurate, as the graphs in Graph Set 3 are too small for the differences in node size to have an effect on perceptual stability. That is, the virtual eyes are effectively looking at the graphs from a long distance. From a long distance small increases in optical aberrations lead to large changes in how the graphs are perceived. This has the effect that when a small amount of optical aberrations occur the node and edge line widths are so thin that they quickly visually disappear. Then as optical aberration continue increasing the image of the graph is effectively blank. Therefore Graph Set 3 is initially very perceptually unstable, but over a full range of normal optical aberrations it is very stable.

Another perspective is if a person with normal optical aberrations looks at Graph Set 3 from far enough away, then the Graph Set cannot be seen. They will only see Graph Set 3 when much closer, or when the Graph Set is made much bigger.

The mean and standard deviation results for Graph Set 4 are shown in Figure 8.21. Region 3 is the most perceptually stable, followed by Region 2 and then Region 1. Individual variability in $PERS_{va}$ scores is plotted in Figure 8.22. These results establish that a thicker node border is better than a thinner node border.

The demonstrations and results for one dimensional adaption show how $PERS_{va}$ can be successfully used to adapt a design, or choose from a set of designs, to improve perceptual stability. While the results for Graph Set 3 further indicate the importance of recognising that $PERS_{va}$ is physiological based rather than semantic based measure (see Section 8.2).

8.3.2.2 Multi-Dimensional Adaptions

Unlike the first four Graph Sets the four graphs in Graph Set 5 are visually distinct from each other along a range of visual dimensions. Each of the graphs in Graph Set 5 is from the previous four Graph Sets, where the right most graph in Graph Set 1 to Graph Set 4 is used.

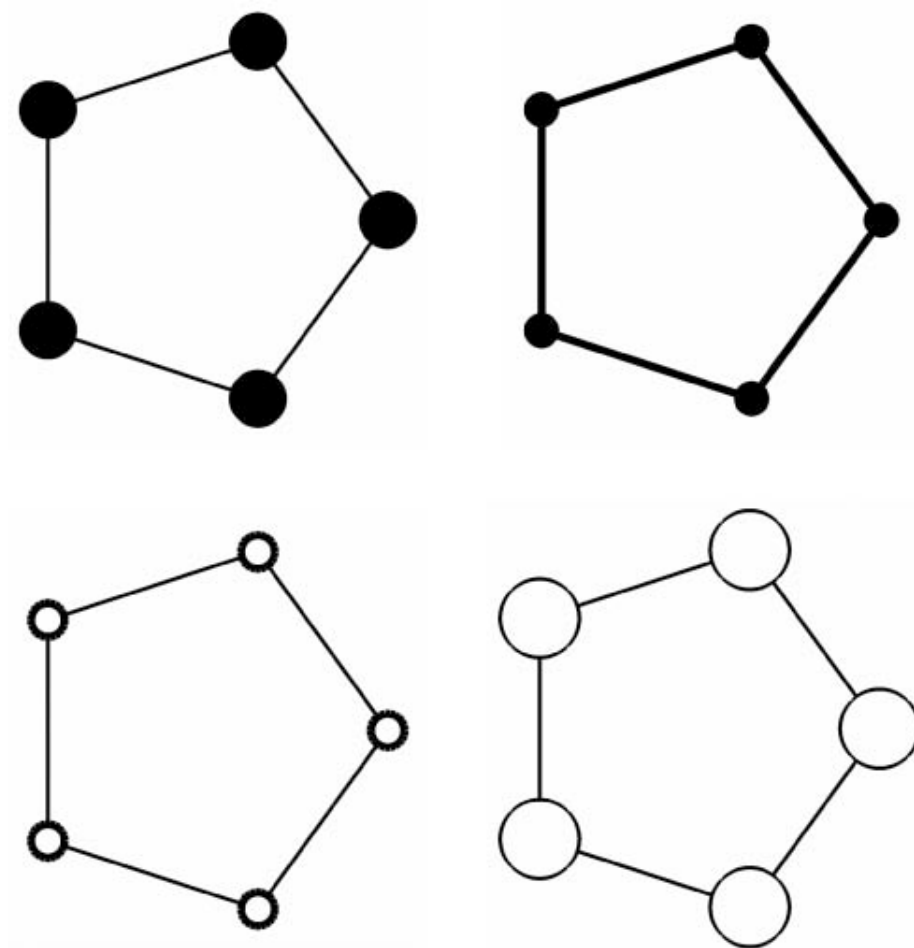
Graph Set 5 is divided into four equal sized Regions, the layout of which is shown at the bottom of Figure 8.23. Each Region covers the area occupied by a single graph, including when the graph undergoes optical aberrations.

Results of using $PERS_{va}$ to analyse the perceptual stability of Graph Set 5 are shown in Figure 8.24 and Figure 8.25. Based on the mean $PERS_{va}$ Region 3 is ranked as the most perceptually stable, very closed followed by Region 1, then Region 2 and finally Region 4.

These results tell us that the wider edge width (Region 3) is the best design adaption choice. Though the difference between the $PERS_{va}$ scores for node size (Region 1) and the winning Region 3 is a small 0.0235 $PERS_{va}$.

Furthermore, based on the standard deviations we note that the perceptual stability of Region 1 fluctuates considerably more than any other Region, making Region 1 a potentially bad choice when perceived by multiple viewers simultaneously. It is not consistently perceived between multiple perceivers.

Interestingly the variability of individual $PERS_{va}$ scores for each eye (see Figure 8.24) shows that six of the twenty eyes ranked Region 1 as more perceptually stable than Region 3. This result informs us that Graph Set 5 can be beneficially displayed in either of two ways - where the choice of which design to display is based on individual physiological visual function, i.e. if there is a measurement of a perceiver's individual aberration wavefront we can choose which design is better to display.



Region 1	Region 3
Region 2	Region 4

Figure 8.23: Top shows four network graphs that vary along multiple visual dimensions, which are compared with each other to establish which is most perceptually stable. Bottom shows how top is broken into four Regions.

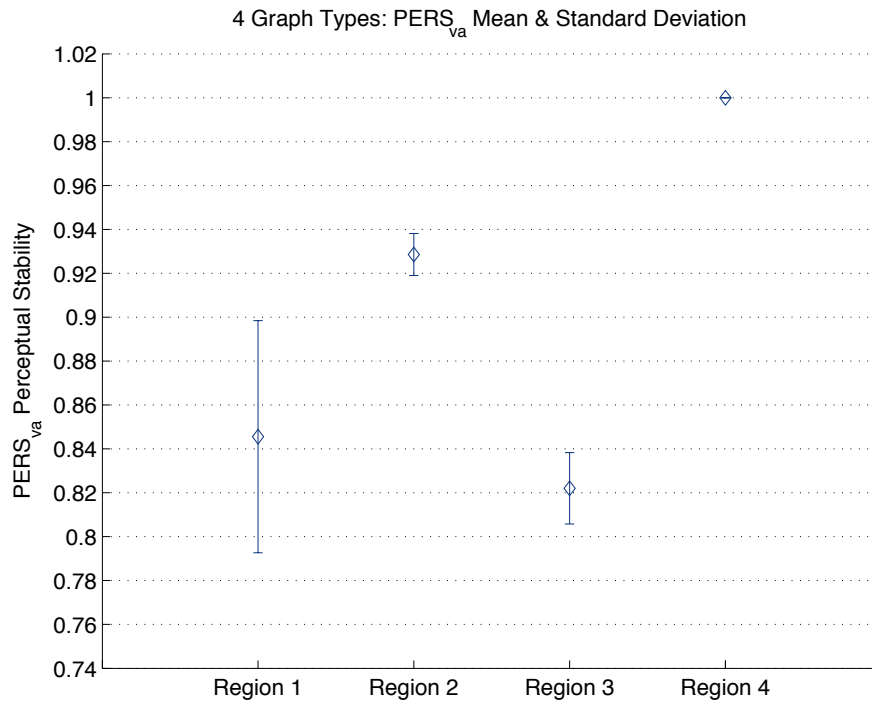


Figure 8.24: Graph Set 5 - $PERS_{va}$ mean and standard deviation.

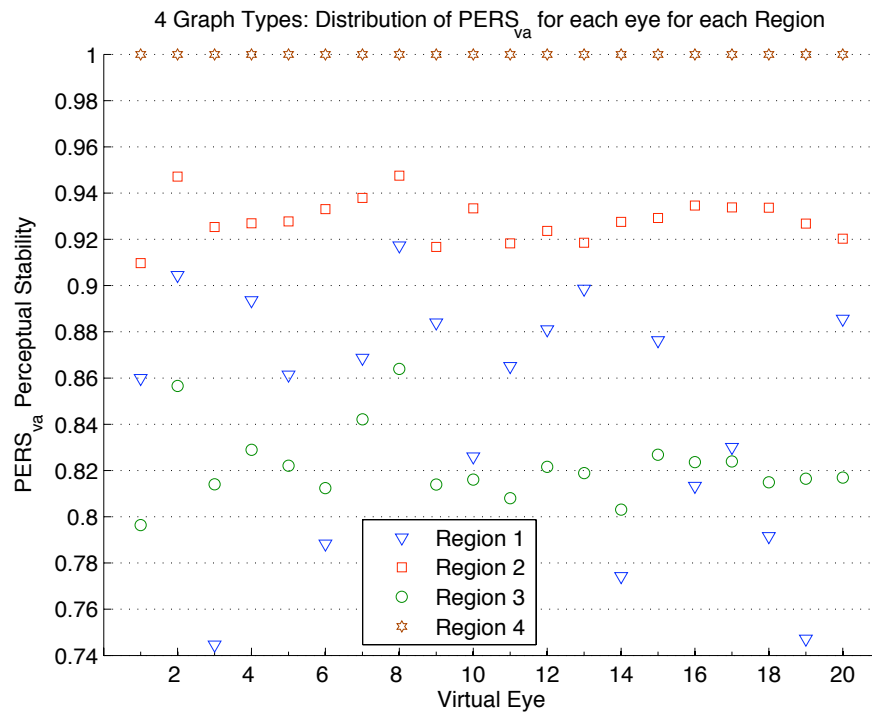


Figure 8.25: Graph Set 5 - Distribution of $PERS_{va}$ for each virtual eye for each region.

8.3.2.3 Graph Layout Adaptions: Future Directions

A core focus of research around graphs is how to lay them out to make them less cluttered, and enhance the graph aesthetics. Many approaches measure line crossing explicitly and attempt to maximum the distance between nodes, such as with force directed layout algorithms.

Figure 8.26 shows a potential new approach to laying out graphs, which is based on low-level vision rather than counting edge crossings or explicitly maintaining distances between edges and nodes. Section 9.3.2.2 provides more examples and detail.

Depicted in Figure 8.26 are multiple seven node seven edge graphs laid out based on $PERS_{va}$, where no edge crossings are explicitly counted. From left to right and top to bottom the fitness of the graph layout improves, when a variation of $PERS_{va}$ is part of the fitness function in a genetic algorithm.

Whether this perceptual graph layout technique generalises is an interesting future research direction. If so it has the potential such that $PERS_{va}$ could be enhanced, to make more complex predictions without requiring semantic knowledge about the analysed designs and interfaces.

8.4 Colour Perception Adaptions

Demonstrated in the following Section is the perceptual stability measure of colour, where it is used to decide what colouring to apply to a design to optimise for perceptually stability.

As with the visual acuity predictor the colour predictor serves as an aid in helping a designer, or automated layout technique, establish how differently designs are perceived due to differences in colour perception.

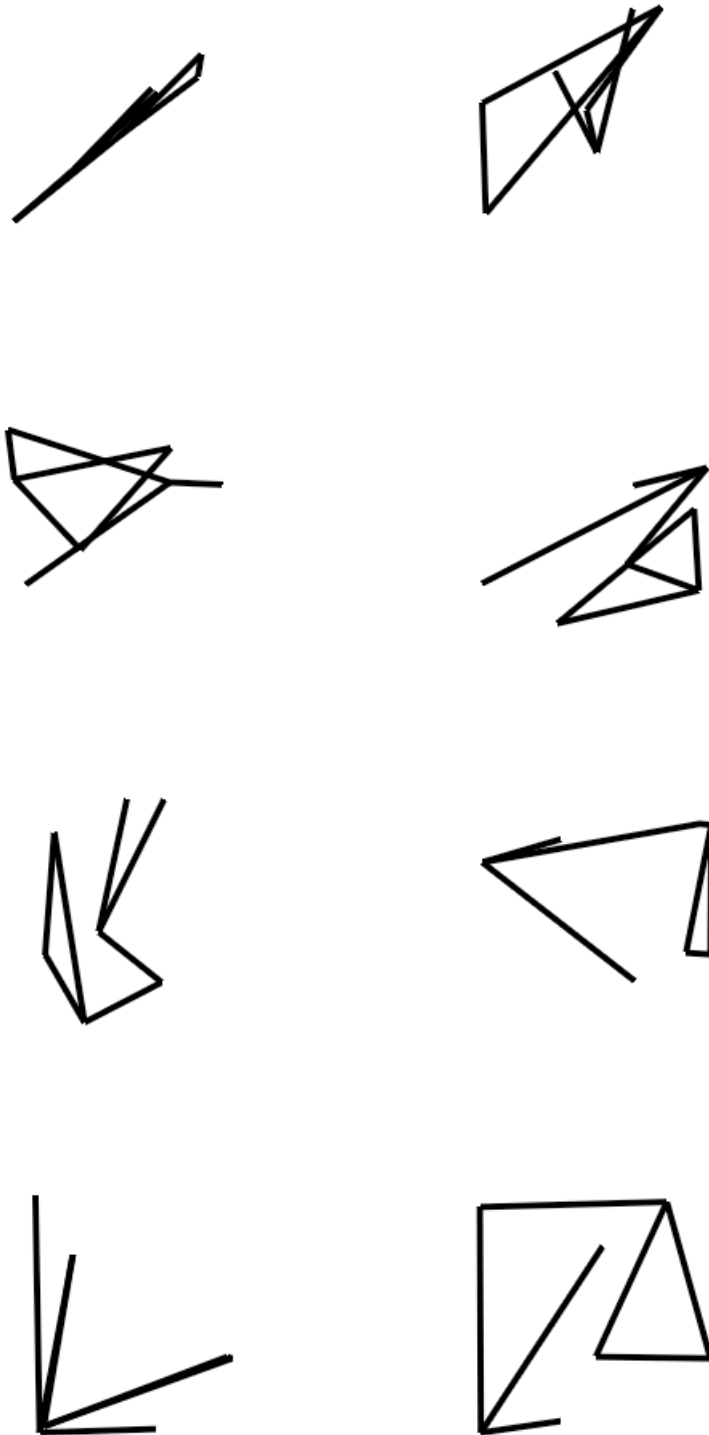


Figure 8.26: Perceptual graph layout - Example of a seven node seven edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.

8.4.1 Maps

What follows are demonstrations showing $PERS_{cp}$ evaluating the effects of colour blindness on the pseudocolouring of maps. Pseudocolouring is widely used in many scientific visualisations [90] to highlight and associate values with locations and objects in visualisations, e.g. a low area in the map is coloured blue, or an area of abnormal weather activity is coloured red.

Pseudocolouring can be confusing for individuals with colour blindness, as they will not see the full range of colours used in the pseudocolouring, which means they can mistakenly interpret the colour encoding.

In the demonstrations $PERS_{cp}$ establishes which pseudocolouring is the most stable between normal and colour deficient observers, i.e. $PERS_{cp}$ evaluates which choice of pseudocolouring differs least between different observers.

Shown in Figure 8.27 are twelve variations of the same map, where the maps differ only by the pseudocolouring used. Going from left to right and top to bottom the pseudocolourings are named in Matlab as Jet, HSV, Hot, Cool, Spring, Summer, Autumn, Winter, Gray, Bone, Copper and Pink.

Two different types of colour blindness are simulated, deuteranopia and protanopia. Figure 8.28 shows the maps as seen by individuals with deuteranopia, and Figure 8.29 shows the maps as seen by individuals with protanopia. Both the deuteranope and protanope versions of the maps bear a close resemblance to each other.

Shown in Figure 8.30 are the thresholded perceptual stability change maps for the deuteranope versions of the maps, while Figure 8.31 shows the protanope thresholded perceptual stability change maps. Highlighted pixels in the perceptual stability change maps have a ΔE range between 10 and 20. The perceptual change map values were normalised to range between 0 and 255 for visual clarity.

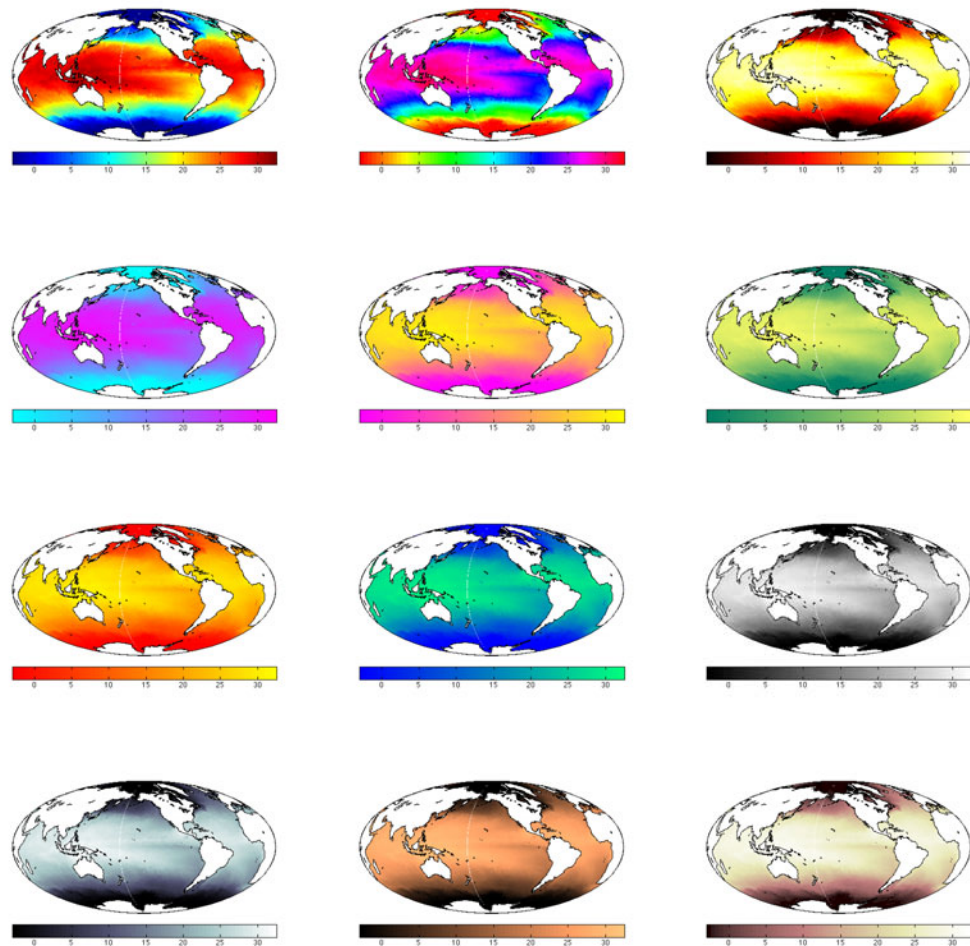


Figure 8.27: Pseudocoloured satellite maps showing NOAA/NASA AVHRR SST for November 1999. Maps were generated with M_Map [64] and publicly available data from NASA. Going from left to right and top to bottom the pseudocolourings are Jet, HSV, Hot, Cool, Spring, Autumn, Gray, Bone, Copper and Pink.

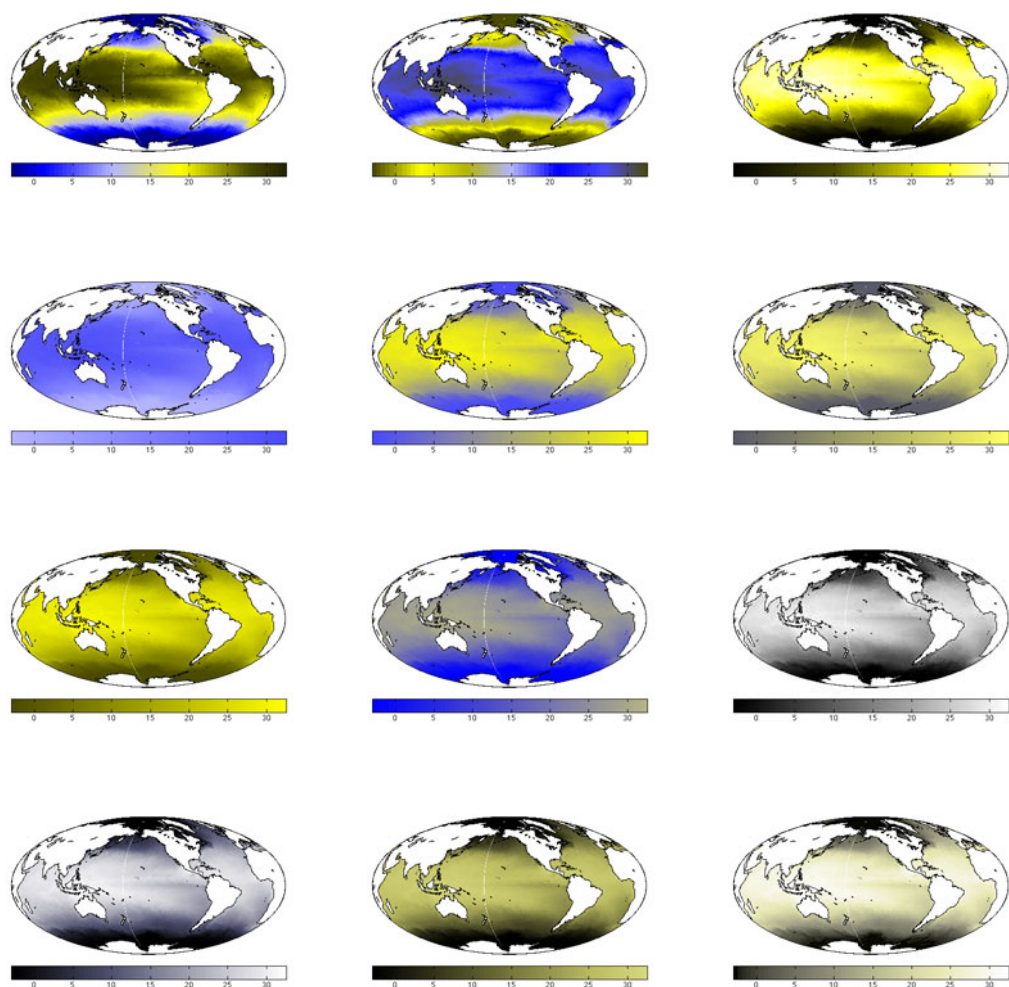


Figure 8.28: Deuteranope versions of maps.

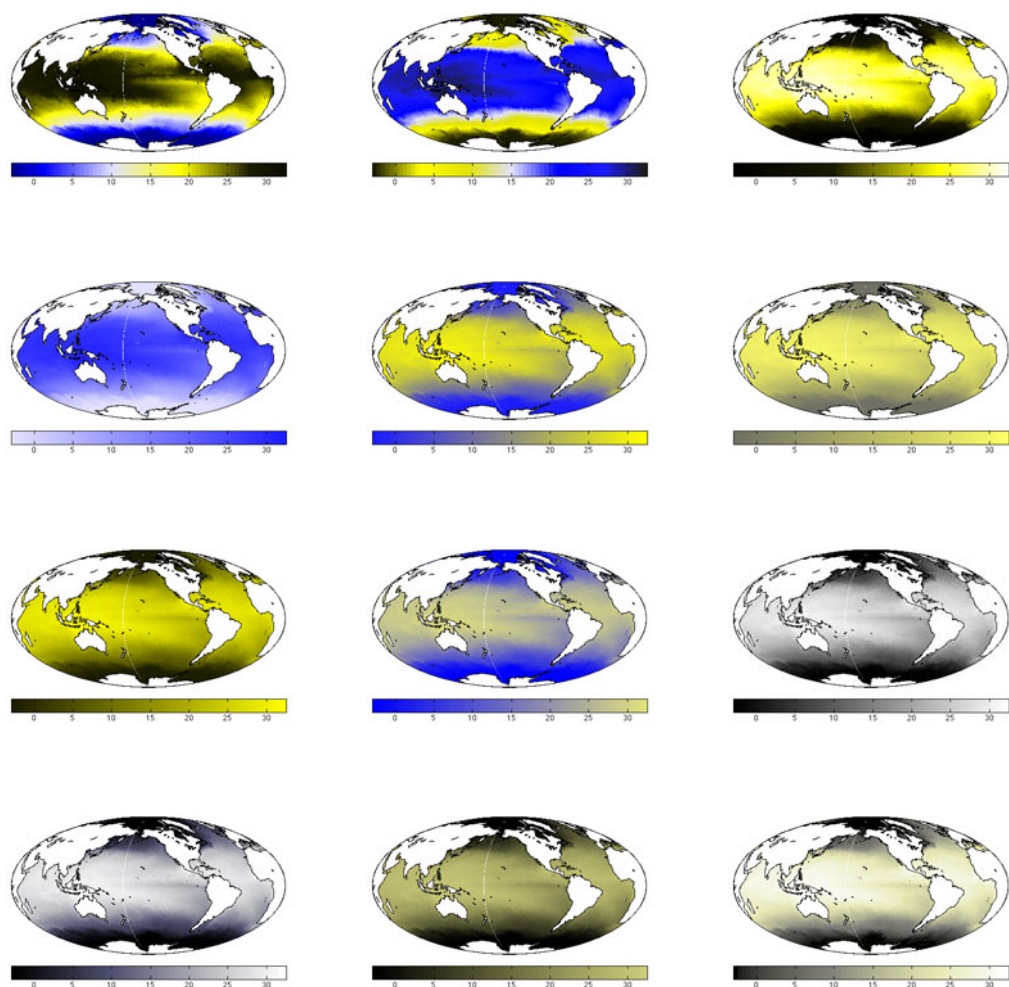


Figure 8.29: Protanope versions of maps.

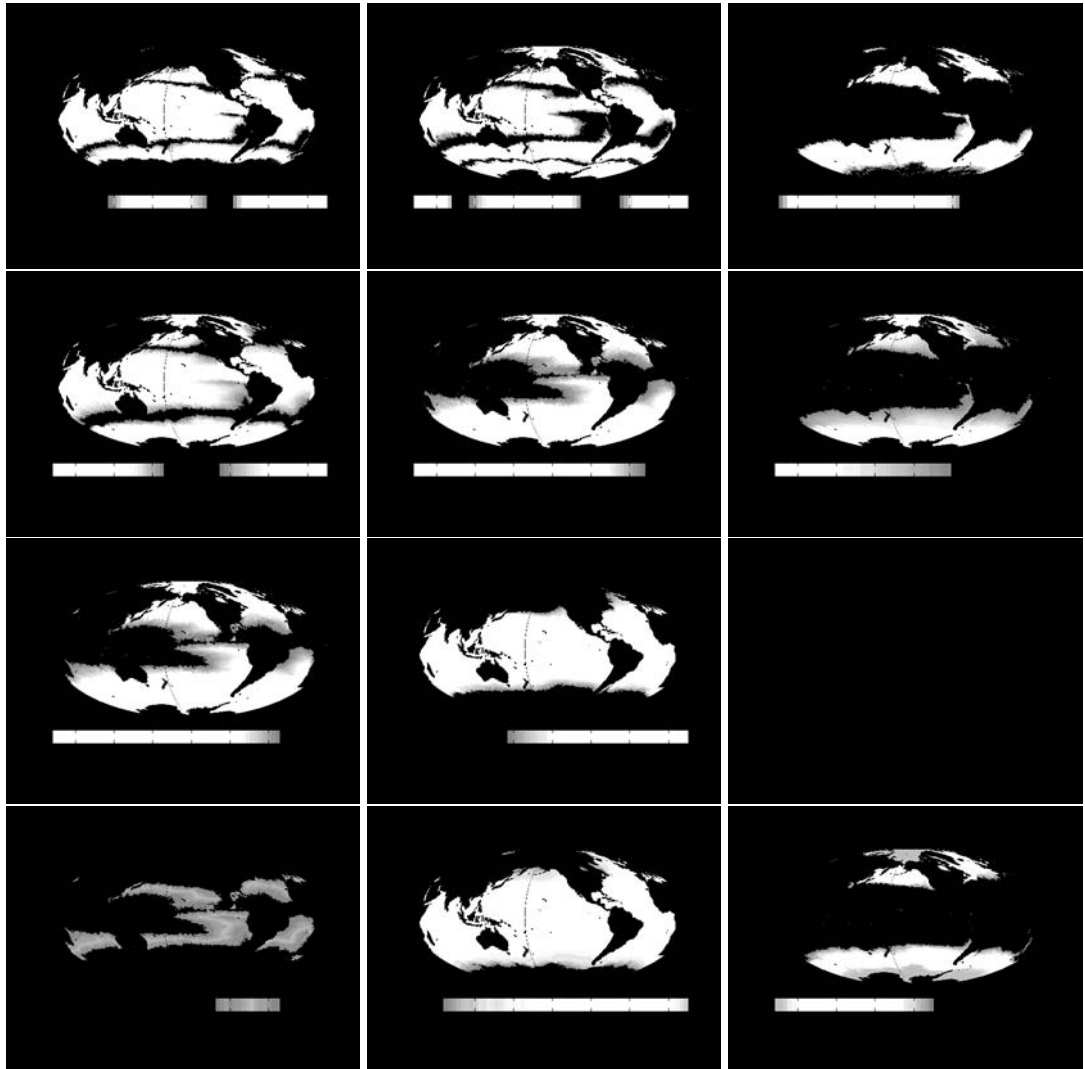


Figure 8.30: Thresholded perceptual change map for deuteranope versions of maps. ΔE ranged between 10 and 20.

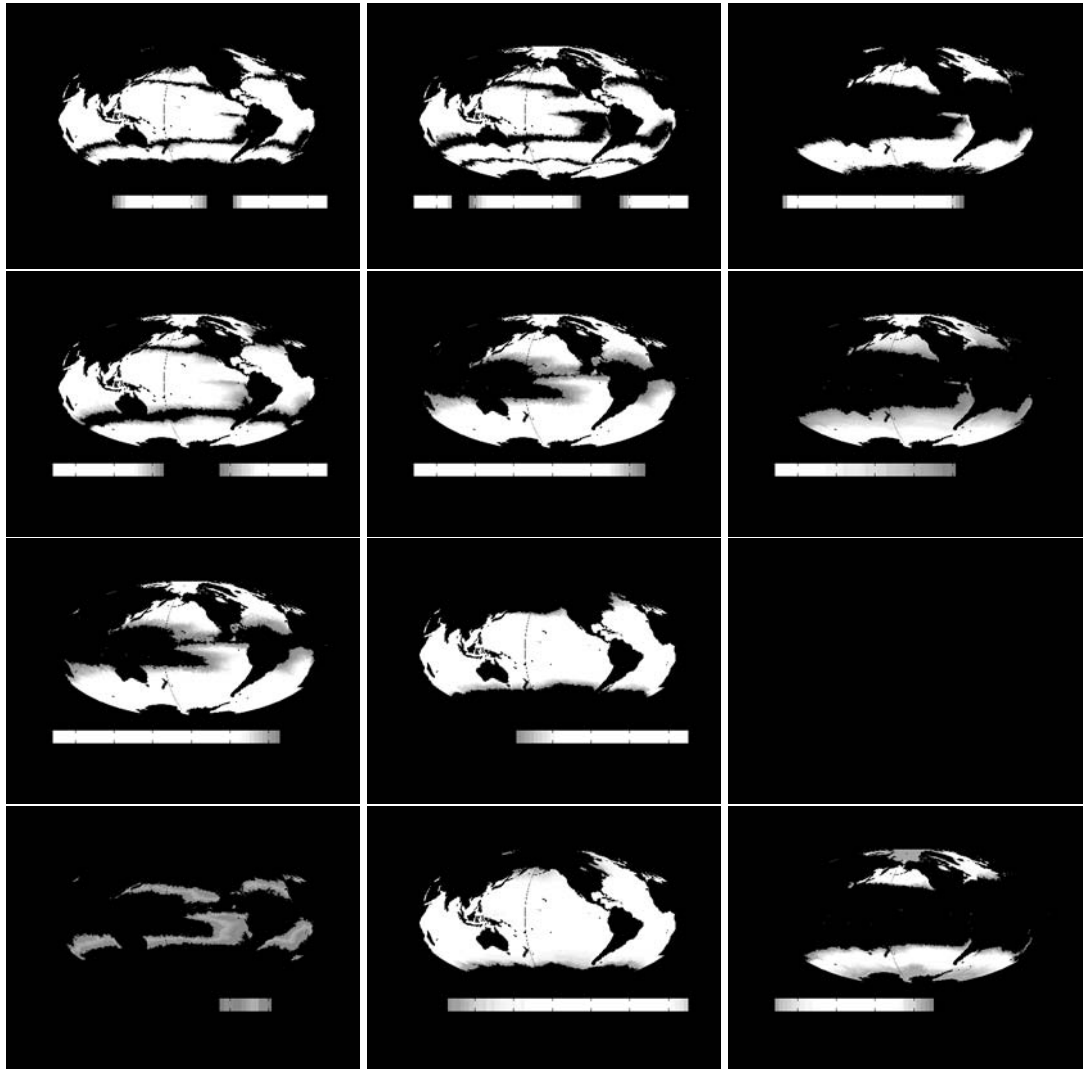


Figure 8.31: Thresholded perceptual change map for protanope versions of maps. ΔE ranged between 10 and 20.

Table 8.2: $PERS_{cp}$ scores for each version of each deuteranope map. Location in table corresponds to locations of maps in Figure 8.30.

3.5782 (Jet)	3.8161 (HSV)	2.1619 (Hot)
3.9318 (Cool)	3.5021 (Spring)	1.9739 (Summer)
3.3832 (Autumn)	3.5985 (Winter)	0 (Gray)
0.9626 (Bone)	4.2713 (Copper)	1.8503 (Pink)

Table 8.3: $PERS_{cp}$ scores for each version of each protanope map. Location in table corresponds to locations of maps in Figure 8.31.

3.55 (Jet)	3.8124 (HSV)	2.1536 (Hot)
3.9678 (Cool)	3.5026 (Spring)	2.0165 (Summer)
3.3862 (Autumn)	3.5152 (Winter)	0 (Gray)
0.6652 (Bone)	4.2343 (Copper)	1.7968 (Pink)

To calculate the $PERS_{cp}$ score each map is treated as a single region, i.e. n and m in Equation 6.1 are image width and height. Table 8.2 and Table 8.3 are the resulting $PERS_{cp}$ scores for each map variation.

For deuteranopes the ranking of the pseudocolouring from most perceptually stable to least perceptually stable is Gray, Bone, Pink, Summer, Hot, Autumn, Spring, Winter, Jet, HSV, Cool, Copper. Interestingly the ranking is the same for protanopes. There is a slight differences in the $PERS_{cp}$ scores between protanopes and deuteranopes but in general the results were in close agreement.

As is expected the Gray pseudocolouring is the most perceptually stable, which makes sense considering most types of colour deficient vision do not affect the perception of pure grays.

8.5 Conclusions

In this Chapter the effectiveness and application of $PERS_{va}$ and $PERS_{cp}$ has been demonstrated; for helping to decide between a range of potential designs. Successfully demonstrated in this Chapter are the automatic evaluations of text and font styles, network graph designs and layouts, and the pseudocolouring of scientific visualisations.

Interestingly we found that $PERS_{va}$ can help evaluate the impact of individual differences in eye function, such that it highlights that certain designs lead to more variability in how people see a design, e.g. the greater variability in how people will see the Times-Roman text paragraph (see Section 8.3.1.1).

Of note is that $PERS_{va}$ is a measure of low-level vision, rather than a semantic measure. Some initial insights on future work show that $PERS_{va}$ may be applied to laying out graphs.

Conclusions & Future Work

At the core this thesis addressed the question “What do people see in a design?”, which is motivated by considering at what point in the experience of designs do individual differences first matter?

This Chapter outlines the contributions of this thesis, while examining future research possibilities.

9.1 When Individual Differences First Matter

Our approach examines the earliest point at which individual differences in vision matter for the experience of seeing the designs (see Table 9.1). The earliest intersection between a design and the experience of a design is the point at which the design is communicated to our bodies and minds. This act of communication from an external occurrence to an internal occurrence alters the design experience (see Section 1.2.1).

Effectively we sought to examine the effect of low-level physiology on design experience, which has many applications when designing artefacts and experiences for everyday users.

Rich models and predictors of the effects of individual differences in low-level

Table 9.1: Refinements of the question “What do people see in a design?”

Facet	Refinement
What	Computationally driven visual designs.
People	Individual differences in low-level vision.
See	Spatial Perception (primarily spatial visual acuity), Colour Perception.
Design	HCI focused on Information Visualisations.

physiology, enable the creation of rewarding user experiences that are capable of adapting to individual physiological functions. Especially in a future where malleability is a standard property of materials and artefacts (see Section 1.2.2).

Furthermore, examining differences in low-level physiology contributes to providing an empirical basis for current design questions, i.e. How do people see and experience a design at different distances, on different sized displays, in different lighting conditions, with different kinds of eye sight?

9.2 Contributions

Multiple inter-related contributions are present in this work, with the main contributions being:

- simplified models of physiological eye function for HCI
- degree and location of perceptual stability due to spatial perception
- degree and location of perceptual stability due to colour perception
- optometry models of individual eye function integrated with perception
- demonstrations of predictions used to make design adaption decisions

9.2.1 Simplified Models Of Physiological Eye Function For HCI

9.2.1.1 STRIVE Algorithm

A contribution for modelling eye function is the STRIVE algorithm, introduced in Section 4.2.2. STRIVE generates simulated retinal images, which model the effects of differences in eye function on the visual entropy of a design.

Zernike Polynomials, which model the wavefront in individual eyes, are an integral aspect of the spatial acuity model (see Section 3.5.3.1). Building upon Zernike Polynomials enables the modelling of individual differences in physiological function, making it possible to measure the effects of individual differences in eye function upon the perception of designs, e.g. larger or smaller standard deviations of $PERS_{va}$ scores for graph adaptations in Chapter 8.

9.2.1.2 STRICODI Algorithm

STRICODI, introduced in Section 6.2.2, is the contribution that enables modelling the effects of differences in colour perception.

By building on ΔE_{00}^* (see Section 3.6.5) and previous work for simulating colour deficient vision STRICODI allows the modelling of differences of colour perception due to differences in eye receptor function.

9.2.2 Degree And Location Of Perceptual Stability Due To Spatial Perception

9.2.2.1 $PERS_{va}$ Perceptual Stability For Spatial Vision

$PERS_{va}$ is the contribution that creates the relationship between STRIVE and the effect differences in spatial acuity have on seeing designs (see Section 4.2.4).

With $PERS_{va}$ the visual change due to optical aberrations is given a meaningful and usable score. The output of STRIVE is a series of data point, as a function of increasing optical aberrations. Analysing the data points is achieved with $PERS_{va}$.

9.2.2.2 Experimental Groundtruth

In Chapter 5 $PERS_{va}$ is validated by testing it on a range of eye charts; where human function is well established. This established that $PERS_{va}$ performances as expected in most cases, and is capable of making degree and location of perceptual stability measures.

9.2.3 Degree And Location Of Perceptual Stability Due To Colour Perception

9.2.3.1 $PERS_{cp}$ Perceptual Stability For Colour Perception

For colour perception $PERS_{cp}$ provides a measure of the impact different types of colour defective vision have on the perception of designs (see Section 6.2.3).

$PERS_{cp}$ builds upon STRICODI to create a relationship between different models of colour perception, and how the resulting colour spaces impact upon

the experience of seeing designs (see Section 3.6).

9.2.3.2 Experimental Groundtruth

Chapter 7 presents the experimental testing and validation of $PERS_{cp}$ on eye charts, which shows that $PERS_{cp}$ concurs with many human judgements about the degree and location of perceptual stabilities due to differences in colour perception.

9.2.4 Optometry Models Of Individual Eye Function Integrated With Perception

$PERS_{va}$ and $PERS_{cp}$, along with STRIVE and STRICODI, create relationships between low-level models of human vision and the experience of seeing the designs.

A resulting contribution is that both models cater to individual differences in physiological eye function, along with the impact differences have on perception.

With $PERS_{va}$ different Zernike coefficients can be used to create individual wavefronts (see Section 2.4.4), which encode optical aberrations found in individual eyes. While $PERS_{cp}$ can use different models of colour defective vision, and is designed such that other colour spaces may be used to model other types of colour receptor function.

9.2.5 Demonstrations Of Predictions Used To Make Design Adaption Decisions

The final significant contributions are the design adaption demonstrations in Chapter 8. These adaptations show the strengths and weaknesses of the predictors, along with how they can be used in visual design tools and automated layout techniques for interfaces and information visualisations.

Demonstrated in Section 8.3 is $PERS_{va}$ deciding between the spatial perception perceptual stability of font styles and graph designs. While in Section 8.4 $PERS_{cp}$ helps automatically choose between different pseudocolouring of maps, to make them clear for individuals with colour blindness.

9.3 Future Directions

There are many possible directions for further research, some in the direction of HCI, while others are within the realm of vision science. In the following sections related research possibilities are presented, followed by some opportunities for directly building upon the findings and techniques presented in this thesis.

9.3.1 Related Possibilities

9.3.1.1 Vision Science Opportunity For HCI

Vision science is a rich and diverse field, with many findings that the HCI community can adopt and apply to understanding, evaluating and improving the presentation of information and interfaces.

Others, such as Ware [90], have recognised the possibilities offered by vision science for HCI. Applying vision science to HCI is challenging, because vis-

ion science is a large and technical field with many specialised sub-fields. Establishing ongoing dialogues and more collaborations with the vision science community would enrich the HCI research community.

9.3.1.2 Other Types Of Perceptual Stability

Though perceptual stability is defined in this thesis only for colour perception and spatial perception, the underlying idea of how differences in physiology mediate design experiences can be generalised to other aspects of our physiology.

For example in vision what are the implications for design experiences due to differences in super-acuties, critical flicker-frequency, and temporal integration?

In a broader scenario, how do individual differences in smell, taste, hearing, touch sensitivity, and balance affect our experiences of designed experiences, e.g. measure taste blindness and adapt a cook book for individual taste receptor function?

Measuring other types of perceptual stabilities will enhance what can be achieved with Mass Customisation and Product Personalisation (see Section 1.2.2).

9.3.2 Building Upon Thesis Findings

9.3.2.1 Colour Perception

Improving $PERS_{cp}$ Measure With Surround

$PERS_{cp}$ currently measures the degree to which each pixel independently changes colour. This could be improved by taking into consideration whether a pixel changes colour, such that it becomes closer in colour to the surrounding

colours.

To move away from point / pixel based colour measures and take perceptual interactions between surrounding colours into consideration, other experimental colour models and colour difference measures such as CIECAM02 or S-CIELAB may be of use (see Section 3.6.6).

By taking surrounding pixels into consideration an improved $PERS_{cp}$ would help establish whether colour changing pixels still stand out from their surround, or blend into the surrounding colour.

Individualised Colour Spaces

Differences in colour receptor function exist between everyone, not just for individuals with colour blindness. Techniques for measuring and modelling individual colour spaces could enable the optimisation of colour spaces to suit individual eye function. Though it is questionable whether the differences are significant enough to be of use.

9.3.2.2 Spatial Perception

Automatic Noise Thresholding

Noise Thresholding is introduced in Section 4.2.5 to handle the misleading effects of gradients on calculating $PERS_{va}$.

Are other approaches possible such that NT is not required? One possibility lies in that the amount of change in NNSE entropy a region experiences is ignored (normalised away) between regions, instead the rate of change is used. Instead of normalising away the amount of change, the amount of change could be analysed in conjunction with the rate of change.

Sensitivity Analysis To Preserve Physiological With Semantic Details

As is discussed in Section 8.2 $PERS_{cp}$ is a physiological measure, not a semantic measure. The effect of this is that it can treat a group of visual features as a single visual feature, because the multiple visual features are so close they optically form a single visual feature - when subject to a small amount of optical aberrations.

How $PERS_{va}$ is used can be adapted, such that it can also analyse a visual design to see how sensitive the design is to a small degrees of optical aberrations. A design which is very sensitivity to small amounts of optical aberration is detailed. The correspondance between amount of detail and sensitivity to optical aberrations provides a form of sensitivity analysis, which may be useful. For example, a core focus of research around graphs is how to lay them out to make the less cluttered. Many approaches measure line crossing explicitly and attempt to maximum the distance between nodes.

Sensitivity analysis may enable a new approach to automatically laying out graphs, based on low-level vision rather than counting edge crossings or explicitly maintaining distances between edges and nodes.

When laying out a graph detail preservation is important. By testing whether a graph loses detail quickly when subject to optical aberrations, we potentially are able to learn whether the graph is laid out such that visual features are easy or difficult to see.

Depicted in Figure 9.1 are multiple six node three edge graphs laid out based on $PERS_{va}$ with sensitivity analysis, where no edge crossings are explicitly counted. From left to right and top to bottom the fitness of the graph layout improves, when $PERS_{va}$ with sensitivity analysis is part of the fitness function in a genetic algorithm. Figure 9.2 shows a five node five edge circular graph, and Figure 9.3 shows a seven node seven edge graph, both laid out

using $PERS_{va}$ with sensitivity analysis.

Whether $PERS_{va}$ with sensitivity analysis generalises as a perceptual graph layout technique is an interesting future research direction. If so it has the potential such that $PERS_{va}$ could be enhanced, to make more complex predictions without requiring semantic knowledge about the analysed designs and interfaces.

9.4 Summary

This thesis has introduced and evaluated objective quantifications of “What do people see in a design?”. The techniques for generating the quantifications are capable of modelling individual differences, which is useful for automating design judgements, i.e. automatically compare a range of potential interface designs and make a decision about which is best for a specific user.

In the longer term, as we move into a world where Mass Customisation and Product Personalisation become common place, objective design quantifications are useful for adapting and customising designs to suit individual physiologies, capabilities and preferences.

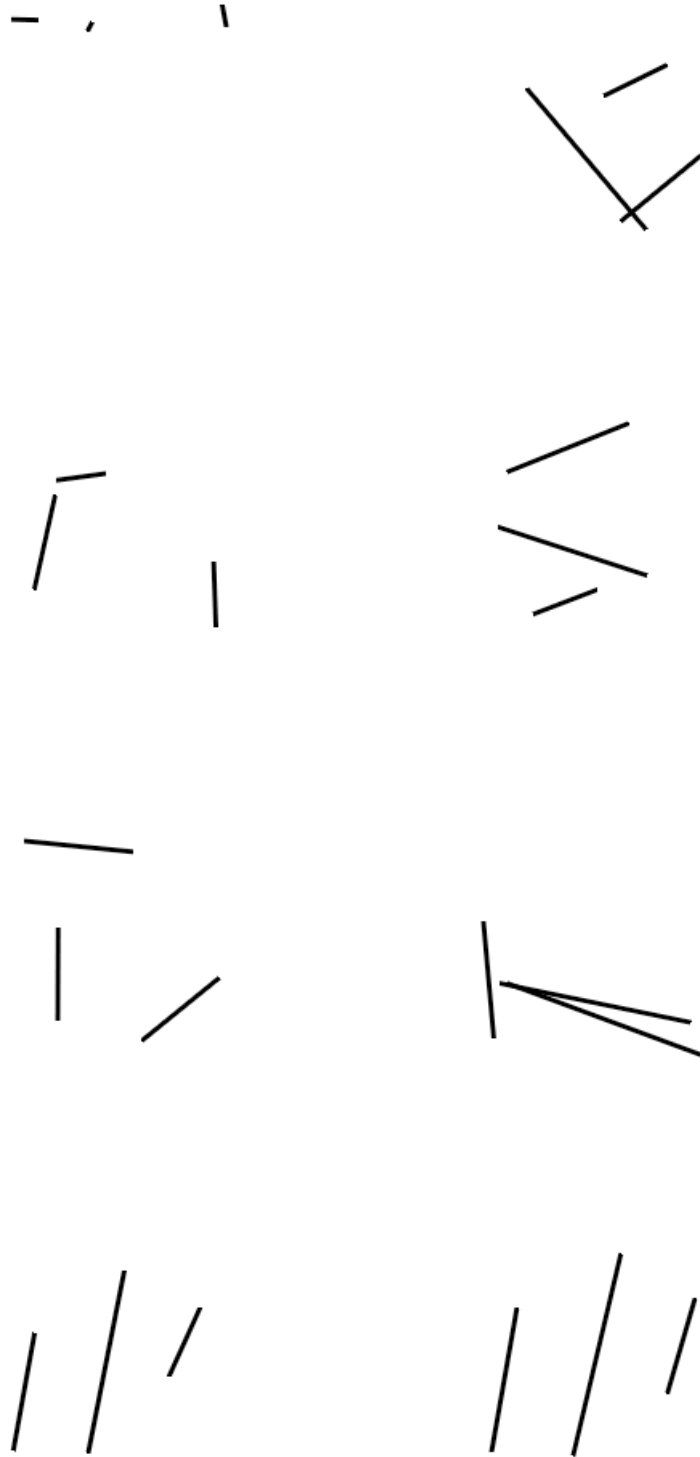


Figure 9.1: Perceptual graph layout - Example of a six node three edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.

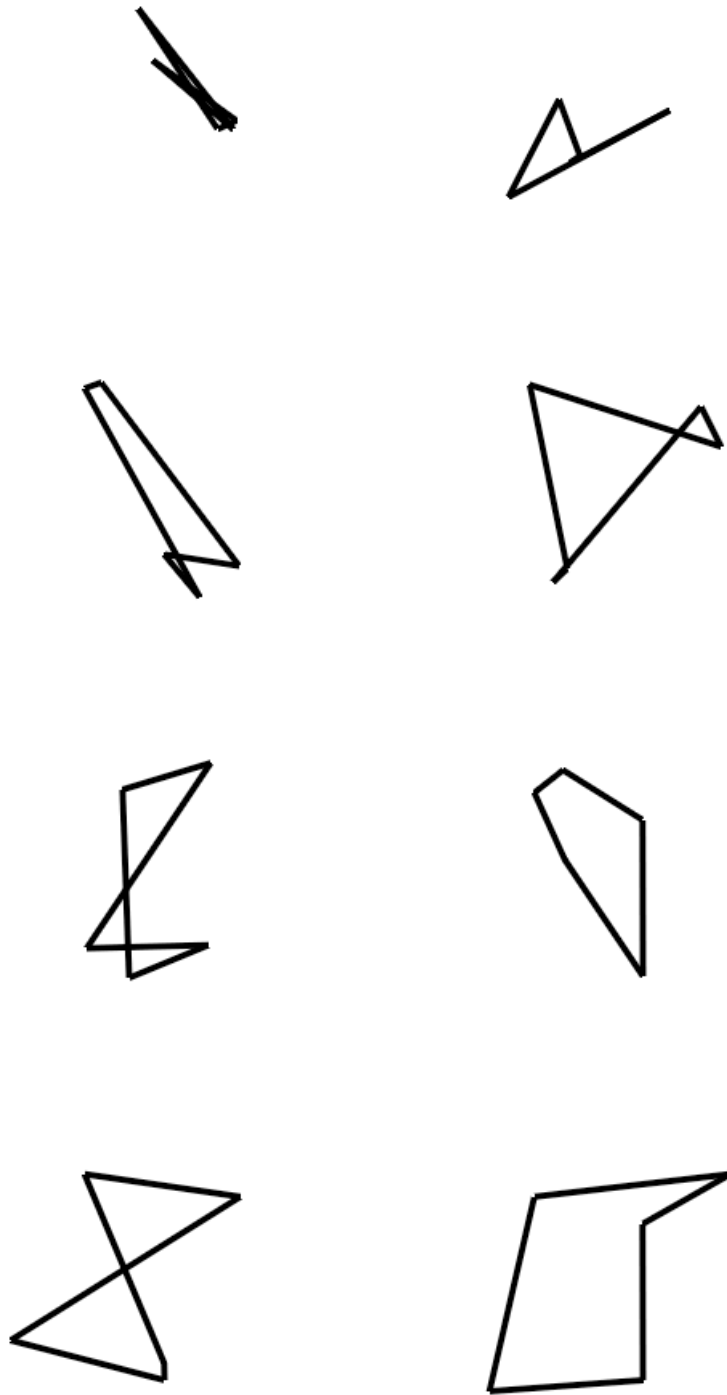


Figure 9.2: Perceptual graph layout - Example of a five node five edge circular graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.

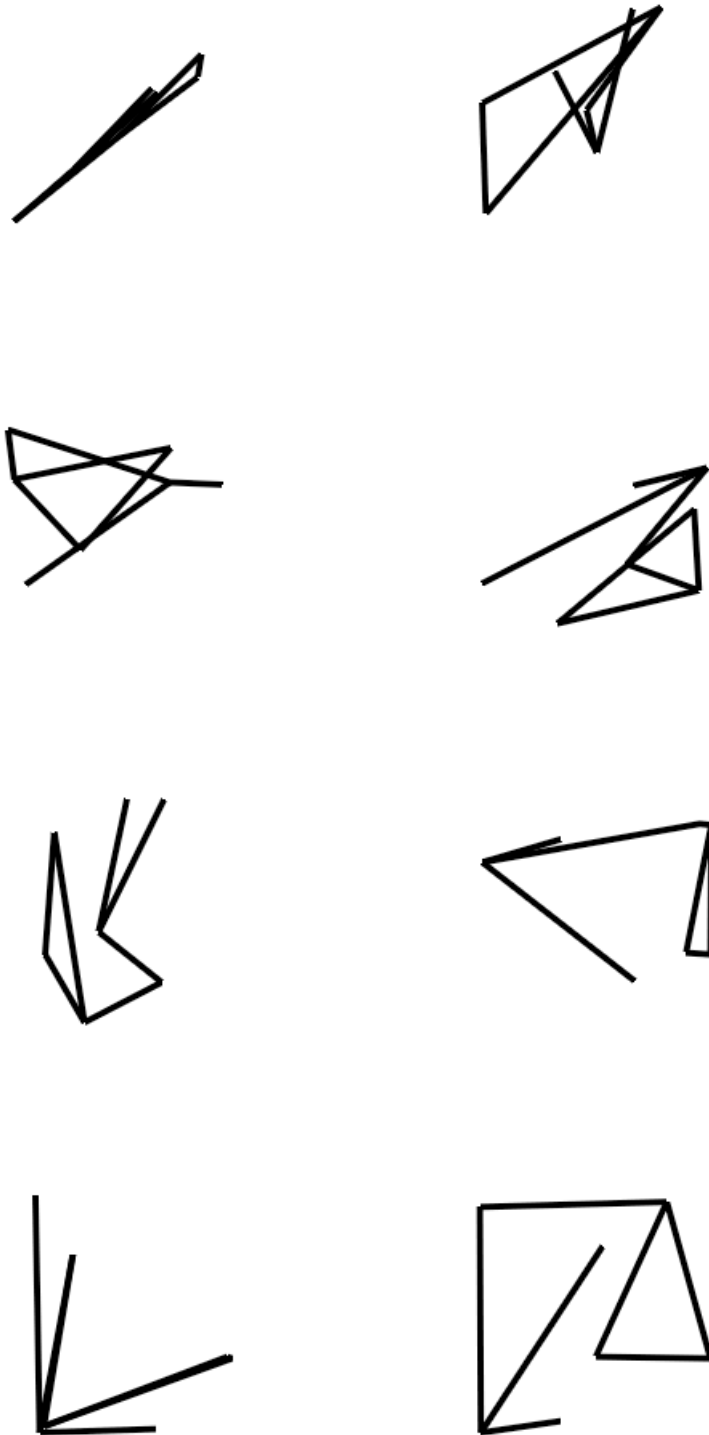


Figure 9.3: Perceptual graph layout - Example of a seven node seven edge graph. Nodes were not drawn during analysis, as $PERS_{va}$ maximization based on edges.

BIBLIOGRAPHY

- [1] Miguel Alonso. *A Method For Enhancing Digital Information Displayed To Computer Users With Visual Refractive Errors Via Spatial And Spectral Processing*. PhD thesis, Electrical Engineering, Florida International University, 2007.
- [2] R.A. Applegate, C. Ballentine, H. Gross, E.J. Sarver, and C.A. Sarver. Visual acuity as a function of Zernike mode and level of root mean square error. *Optometry and Vision Science*, 80:97–105, 2003.
- [3] Joseph J. Atick. Could information theory provide an ecological theory of sensory processing? *Journal of Network: Computation in Neural Systems*, 3:213–251, 1992.
- [4] Ian L. Bailey and Jan E. Lovie-Kitchin. New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*, 53(11):740–745, 1976.
- [5] Richard Barakat. Some entropic aspects of optical diffraction imagery. *Optics Communications*, 156:235–239, 1998.
- [6] Brian A. Barsky. Vision-realistic rendering: Simulation of the scanned foveal image from wavefront data of human subjects. In *Proceedings of the*

- 1st Symposium on Applied Perception in Graphics and Visualization*, pages 73–81, 2004.
- [7] Tony Belpaeme. *Factors influencing the origins of colour categories*. PhD thesis, Artificial Intelligence Lab, Vrije Universiteit Brussel, 2002.
- [8] Mike Bennett and Aaron Quigley. Automatically evaluating the impact of colour blindness on information visualisations. In *CASCON 2006 Dublin Symposium*, Dublin, Ireland, October 2006. IBM Centre for Advanced Studies.
- [9] Mike Bennett and Aaron Quigley. A method for the automatic analysis of colour category pixel shifts during dichromatic vision. In *2nd International Symposium on Visual Computing (ISVC)*, volume 2 of *Lecture Notes in Computer Science*, pages 457–466, Lake Tahoe Nevada, USA, Nov 6-8 2006. Springer Verlag.
- [10] Mike Bennett and Aaron Quigley. Perceptual usability: Predicting changes in interfaces & designs due to visual acuity differences. In *AVI 2008 Conference on Advanced Visual Interfaces*, pages 380–383, Napoli, Italy, May 2008.
- [11] Mike Bennett and Aaron Quigley. Understanding distance & how humans see interfaces & designs. In *VGV Irish Graduate Student Symposium on Vision, Graphics and Visualisation*, Trinity College Dublin, June 2008.
- [12] Brent Berlin and Paul Kay. *Basic Color Terms: Their Universality and Evolution*. London: University of California Press, 1969.
- [13] David Bimler and John Kirkland. Colour-space distortion in women who are heterozygous for colour deficiency. *Vision Research*, 49(5):536–543, 2009.

- [14] David H. Brainard and Brian A. Wandell. Asymmetric color matching: how color appearance depends on the illuminant. *Journal of the Optical Society of America. A, Optics and image science*, 9(9):1433–48, September 1992.
- [15] Hans Brettel, Françoise Viénot, and John D. Mollon. Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America. A, Optics and Image Science*, 14(10):2647–2655, 1997.
- [16] Vicki Bruce, Patrick R. Green, and Mark A. Georgeson. *Visual Perception: Physiology, Psychology and Ecology*. Hove: Psychology, 1996.
- [17] Isabelle Brunette, Juan M. Bueno, Mireille Parent, Habib Hamam, and Pierre Simonet. Monochromatic aberrations as a function of age, from childhood to advanced age. *Investigative Ophthalmology and Visual Science*, 44(12):5438–5446, December 2003.
- [18] C.E. Campbell. Matrix method to find a new set of Zernike coefficients from an original set when the aperture radius is changed. *Journal of Optical Society, Am A*, 20:209–217, 2003.
- [19] F. W. Campbell and J. G. Robson. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 197:551–566, 1968.
- [20] Stuart K. Card, Jock Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1999.
- [21] John M. Carroll, editor. *HCI Model, Theories, and Frameworks: Towards a Multidisciplinary Science*. Morgan Kaufmann, 2003.
- [22] Chaomei Chen. Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16, 2005.

- [23] Xu Cheng, Larry N. Thibos, and Arthur Bradley. Estimating visual quality from wavefront aberration measurement. *Journal of Refractive Surgery*, 19(5):S579–S584, September/October 2003.
- [24] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, September 1984.
- [25] C.A. Curcio, K.R. Jr Sloan, and et al. O. Packer. Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science*, (236):579–582, 1987.
- [26] Catalina M. Danis, Fernanda B. Viegas, Martin Wattenberg, and Jesse Kriss. Your place or mine? Visualisation as a community component. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 275–284, 2008.
- [27] Michael F. Deering. A photon accurate model of the human eye. In *Proceedings of SIGGRAPH Conference on Computer Graphics and Interactive Techniques*, pages 649–658, 2005.
- [28] Andrew Dillon and Charles Watson. User analysis in HCI: The historical lesson from individual differences research. *International Journal of Human-Computer Studies*, 45(6):619–637, 1996.
- [29] Li Ming Dong, Barbara S. Hawkins, and Marta J. Marsh. Consistency between visual acuity scores obtained at different distances. *Archives of Ophthalmology*, 120:1523–1533, November 2002.
- [30] R. Dougherty and A. Wade. Vischeck, <http://www.vischeck.com>, June 2009.

- [31] Thomas Erickson and David W. McDonald, editors. *HCI Remixed: Reflections on Works That Have Influenced the HCI Community*. MIT Press, 2008.
- [32] Mark Fairchild. *Color Appearance Models*. Wiley, 2nd edition, 2005.
- [33] Mark D. Fairchild. Color appearance models: CIECAM02 and beyond. In *Conference Lecture at IS&T/SID 12th Color Imaging Conference*, 2004.
- [34] F.L. Ferris, A. Kassoff, G.H. Bresnick, and I. Bailey. New visual acuity charts for clinical research. *American Journal of Ophthalmology*, 94:91–96, 1982.
- [35] Daniel Dante Garcia. *CWhatUC: Software Tools for Predicting, Visualizing and Simulating Corneal Visual Acuity*. PhD thesis, University of California at Berkeley, 2000.
- [36] A. P. Ginsburg, J. Easterly, and D. W. Evans. Contrast sensitivity predicts target detection field performance of pilots. In *Proceedings of Human Factors Society*, pages 269–273, October 1983.
- [37] A.P. Ginsburg, D.W. Evans, R. Sekuler, and S.A. Harp. Contrast sensitivity predicts pilots’ performance in aircraft simulators. *American Journal of Optometry and Physiological Optics*, 59:105–108, 1982.
- [38] Antonio Guirao and David R. Williams. A method to predict refractive errors from wave aberration data. *Journal of Optometry and Vision Science*, 80(1):36–42, 2003.
- [39] Gunilla Haegerstrom-Portnoy, Marilyn E. Schneck, Lori A. Lott, and John A. Brabyn. The relation between visual acuity and other spatial vision measures. *Journal of Optometry and Vision Science*, 77(12):653–662, 2000.

- [40] Eugene Hecht. *Optics*. Addison Wesley, 4th edition, 2002.
- [41] G. Heron, H. P. Furby, R. J. Walker, C. S. Lane, and O. J. E. Judge. Relationship between visual acuity and observation distance. *Journal of Ophthalmic and Physiological Optics*, 15(1):23–30, 1995.
- [42] Jia-Bin Huang, Sih-Ying Wu, and Chu-Song Chen. Enhancing color representation for the color vision impaired. In *Workshop on Computer Vision Applications for the Visually Impaired (CVAVI 08)*, October 2008.
- [43] R.W.G. Hunt. *Measuring Colour*. Fountain Press, 3rd edition, 2001.
- [44] R.W.G. Hunt. *The Reproduction of Color*. Voyageur Press, 6th edition, 2004.
- [45] Leo M. Hurvich. *Color Vision*. Sinauer Associates Inc., 1981.
- [46] B. Joseph Pine II. *Mass Customization: The New Frontier in Business Competition*. Harvard Business School Press, 1993.
- [47] Luke Jefferson and Richard Harvey. Accommodating color blind computer users. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 40–47, 2006.
- [48] Ronald R. Krueger, Raymond A. Applegate, and Scott M. MacRae, editors. *Wavefront Customized Visual Correction: The Quest for Super Vision II*. SLACK Incorporated, 2004.
- [49] Jeremy Kubica, Arancha Casal, and Tad Hogg. Complex behaviors from local rules in modular self-reconfigurable robots. In *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pages 360–367, May 2001.

- [50] Johan Maurice Gis  le Lammens. *A Computational Model of Color Perception and Color Naming*. PhD thesis, State University of New York at Buffalo (Computer Science), June 1994.
- [51] Gordon E. Legge, Kathy T. Mullen, George C. Woo, and F.W. Campbell. Tolerance to visual defocus. *Journal of Optical Society, Am A*, 4(5):851–863, May 1987.
- [52] Al Lens. *Optics, Retinoscopy, and Refractometry*. SLACK Incorporated, 2nd edition, 2002.
- [53] M. R. Luo. Colour difference formulae: Past, present and future. In *CIE Expert Symposium*, 2006.
- [54] M. R. Luo, G. Cui, and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Application*, 26(5):340–350, October 2001.
- [55] Bruce MacEvoy. Handprint color theory, <http://www.handprint.com/hp/wcl/wcolor.html>, June 2009.
- [56] J. Mansfield, G. Legge, and M. Bane. Psychophysics of reading: XV: Font effects in normal and low vision. *Journal of Investigative Ophthalmology and Visual Science*, 37(8):1492–1501, 1996.
- [57] David Marr. *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman and Company, 1982.
- [58] Gary W. Meyer and Donald P. Greenberg. Color-defective vision and computer graphics displays. *IEEE Computer Graphics and Applications*, 8(5):28–40, September 1988.
- [59] John D. Mollon, Joel Pokorny, and Ken Knoblauch. *Normal and Defective Colour Vision*. Oxford University Press, May 2003.

- [60] Nathan Moroney, Mark D. Fairchild, Robert W.G. Hunt, Changjun Li, M. Ronnier Luo, and Todd Newman. The CIECAM02 color appearance model. In *IS&T/SID Tenth Color Imaging Conference*, pages 23–27, 2002.
- [61] Thomas T. Norton, David A. Corliss, and James E. Bailey. *The Psychophysical Measurement of Visual Function*. Butterworth Heinemann, 2002.
- [62] Cynthia Owsley and Michael E. Sloane. Contrast sensitivity, acuity, and the perception of ‘real-world’ targets. *British Journal of Ophthalmology*, 71:791–796, 1987.
- [63] Stephen E. Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, 1999.
- [64] Rich Pawlowicz. M_map: A mapping package for matlab, 2008.
- [65] Eli Peli. Test of a model of foveal vision by using simulations. *Journal of the Optical Society of America A*, 13(6):1131–1138, June 1996.
- [66] D. G. Pelli, J. G. Robson, and A. J. Wilkins. The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences*, 2(3):187–199, 1988.
- [67] Padmanabhan Pillai, Jason D. Campbell, Gautam Kedia, Shishir Moudgal, , and Kaushik Sheth. A 3D fax machine based on claytronics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4728–4735, October 2006.
- [68] Jennifer Preece, Yvonne Rogers, and Helen Sharp. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley and Sons, 2002.
- [69] Mihail C. Roco and William Sims Bainbridge, editors. *Converging Technologies for Improving Human Performance: Nanotechnology, Biotechnology, Information Technology and Cognitive Science*. National Science Foundation, Arlington, Virginia, United States, June 2002.

- [70] M.L. Rodriguez-Carmona, J.A. Harlow, G. Walker, and J.L. Barbur. The variability of normal trichromatic vision and the establishment of the "normal" range. In *Proceedings of 10th Congress of the International Colour Association*, pages 979–982, 2005.
- [71] Gary S. Rubin, Beatriz Muñoz, Karen Bandeen-Roche, and Sheila K. West. Monocular versus binocular visual acuity as measures of vision impairment and predictors of visual disability. *Journal of Investigative Ophthalmology and Visual Science*, 41(11):3327–3334, October 2000.
- [72] Thomas O. Salmon. Wavefronts, Zernikes and RMS - the basics in plain language. Symposium at COPE 2006, March 2006.
- [73] Thomas O. Salmon and Corina van de Pol. Normal-eye Zernike coefficients and root-mean-square wavefront errors. *Journal of Cataract Refractive Surgery*, 32:2064–2074, December 2006.
- [74] B.A.C. Saunders and J. van Brakel. Are there non-trivial constraints on colour categorization? *Behavioral and Brain Sciences*, 20(2):167–228, 1997.
- [75] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, October 1948.
- [76] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary tests data, and mathematical observations. *Color Research and Application*, 30(1), February 2005.
- [77] M. T. Sheehan, A. V. Goncharov, V. M. O'Dwyer, V. Toal, and C. Dainty. Population study of the variation in monochromatic aberrations of the normal human eye over central visual field. *Optics Express*, 15(12):7367–7380, 2007.

- [78] V.C. Smith and J. Pokorny. Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, 15:161–171, 1975.
- [79] A. Stockman and L.T. Sharpe. Cone spectral sensitivities and color matching. In K.R. Gegenfurtner and L.T. Sharpe, editors, *Color Vision from Genes to Perception*, pages 53–87. Cambridge University Press, Cambridge, UK, 1999.
- [80] Larry N. Thibos. Wavefront data reporting and terminology. *Journal of Refractive Surgery*, 17(suppl):S578–583, 2001.
- [81] Larry N. Thibos, Raymond A. Applegate, James T. Schwiegerling, Robert Webb, and VSIA Standards Taskforce Members. Standards for reporting the optical aberrations of eyes. In *Vision Science and its Applications, OSA Technical Digest*, pages 232–244. Optical Society of America, February 2000.
- [82] Larry N. Thibos, Arthur Bradley, and Xin Hong. A statistical model of the aberration structure of normal, well-corrected eyes. *Journal of Ophthalmic and Physiological Optics*, 22:427–433, 2002.
- [83] Larry N. Thibos, Xin Hong, Arthur Bradley, and Raymond A. Applegate. Accuracy and precision of objective refraction from wavefront aberrations. *Journal of Vision Science*, 4:329–351, 2004.
- [84] Larry N. Thibos, Xin Hong, Arthur Bradley, and Xu Cheng. Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *Journal of Optical Society, Am A*, 19(12):2329–2348, 2002.
- [85] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 2nd edition, May 2001.

- [86] Fernanda B. Viegas, Martin Wattenberg, Frank van Ham, Jesse Kriss, and Matt McKeon. Many Eyes: A site for visualization at internet scale. In *IEEE Transactions on Visualization and Computer Graphics*, volume 13, pages 1121–1128, 2007.
- [87] Françoise Viénot, Hans Brettel, and John D. Mollon. Digital video colourmaps for checking the legibility of displays by dichromats. *Color research and application*, 24(4):243–252, 1999.
- [88] T. Wachtler, Ulrike Dorhmann, and Rainer Hertel. Modeling color percepts of dichromats. *Vision Research*, 44:2843–2855, 2004.
- [89] Ken Wakita and Kenta Shimamura. Smartcolor: Disambiguation framework for the colorblind. In *SIG ACCESS Conference on Assistive Technologies*, pages 158–165. ACM, 2005.
- [90] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2nd edition, 2004.
- [91] Andrew B. Watson and Albert J. Ahumada. Predicting visual acuity from wavefront aberrations. *Journal of Vision Science*, 8(4)(17):1–19, 2008.
- [92] Sheila K. West, Beatriz Muñoz, Gary S. Rubin, Oliver D. Schein, Karen Bandeen-Roche, Scott Zeger, Pearl S. German, and Linda P. Fried. Function and visual impairment in a population based study of older adults. *Journal of Investigative Ophthalmology and Visual Science*, 38(1):72–82, 1997.
- [93] Sheila K. West, Gary S. Rubin, Aimee T. Broman, Beatriz Muñoz, Karen Bandeen-Roche, and Kathleen Turano. How does visual impairment affect performance on tasks of everyday life? *Archives of Ophthalmology*, 120:774–780, June 2002.
- [94] Stefan Winkler. *Digital Video Quality: Vision Models and Metrics*. Wiley, 2005.

- [95] H. R. Wu and K. R. Rao. *Digital Video Image Quality and Perceptual Coding*. CRC, 2005.
- [96] G. Wyszecki and W.S. Stiles. *Color Science Concepts and Methods, Quantitative Data and Formulae*. Wiley Interscience, New York, 2nd edition, 1982.
- [97] Sergej N. Yendrikhovskij. Computing color categories from statistics of natural images. *Journal of Imaging Science and Technology*, 45(5):409–417, 2001.
- [98] Mark Yim, Ying Zhang, John Lamping, and Eric Mao. Distributed control for 3D metamorphosis. *Autonomous Robots*, 10(1):41–56, January 2001.
- [99] ANSI Z80.28-2004. American national standard for ophthalmics - methods for reporting optical aberrations of eyes. ANSI, 2004.
- [100] Jiajie Zhang. The interaction between perceptual and cognitive processes in a distributed problem solving task. In *In Working Notes of the 1993 AAAI Fall Symposium on Games: Planning and Learning*, 1993.
- [101] Jiajie Zhang. The nature of external representations in problem solving. *Cognitive Science*, 21(2):179–217, 1997.
- [102] Xuemei Zhang and Brian A. Wandell. A spatial extension of CIELAB for digital colour image reproduction. In *Proceedings of the SID Symposiums*, pages 731–734, 1996.