Individualized Models of Colour Differentiation through Situation-Specific Modelling

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the degree of Doctor of Philosophy
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By
David R. Flatla

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5C9

Abstract

In digital environments, colour is used for many purposes: for example, to encode information in charts, signify missing field information on websites, and identify active windows and menus. However, many people have inherited, acquired, or situationally-induced Colour Vision Deficiency (CVD), and therefore have difficulties differentiating many colours. Recolouring tools have been developed that modify interface colours to make them more differentiable for people with CVD, but these tools rely on models of colour differentiation that do not represent the majority of people with CVD. As a result, existing recolouring tools do not help most people with CVD.

To solve this problem, I developed Situation-Specific Modelling (SSM), and applied it to colour differentiation to develop the Individualized model of Colour Differentiation (ICD). SSM utilizes an in-situ calibration procedure to measure a particular user's abilities within a particular situation, and a modelling component to extend the calibration measurements into a full representation of the user's abilities. ICD applies in-situ calibration to measuring a user's unique colour differentiation abilities, and contains a modelling component that is capable of representing the colour differentiation abilities of almost any individual with CVD.

This dissertation presents four versions of the ICD and one application of the ICD to recolouring. First, I describe the development and evaluation of a feasibility implementation of the ICD that tests the viability of the SSM approach. Second, I present revised calibration and modelling components of the ICD that reduce the calibration time from 32 minutes to two minutes. Next, I describe the third and fourth ICD versions that improve the applicability of the ICD to recolouring tools by reducing the colour differentiation prediction time and increasing the power of each prediction. Finally, I present a new recolouring tool (ICDRecolour) that uses the ICD model to steer the recolouring process. In a comparative evaluation, ICDRecolour achieved 90% colour matching accuracy for participants – 20% better than existing recolouring tools – for a wide range of CVDs.

By modelling the colour differentiation abilities of a particular user in a particular environment, the ICD enables the extension of recolouring tools to helping most people with CVD, thereby reducing the difficulties that people with CVD experience when using colour in digital environments.

Acknowledgements

I would first like to thank my supervisor, Carl Gutwin, for his amazing guidance, suggestions, and insights into this research project. You have prepared me very well for what lies ahead. Thank you for your patience, corrections, and encouragement!

I would also like to thank my committee for their many questions and suggestions:

Dr. Andrew Sears, Rochester Institute of Technology (External Examiner).

Dr. Ron Borowsy, Department of Psychology, University of Saskatchewan (Cognate).

Dr. Mark Eramian, Department of Computer Science, University of Saskatchewan.

Dr. Anthony Kusalik, Department of Computer Science, University of Saskatchewan.

Dr. Nathaniel Osgood, Department of Computer Science, University of Saskatchewan.

I am also deeply indebted to the members of the HCI lab over the past five years for countless practice talks and pilot studies, not to mention their advice, feedback, and criticism.

I would also like to thank my close friends who have been my sounding board at one time or another, struggling with me through the challenging bits, and celebrating with me through the successes. Thank you Kent, Chelsea, Jesse, Richard, Brennan, Kim, and Nelson.

I also am deeply thankful for the unending support, encouragement, and prayers of my parents Raymond and Audrey and my parents-by-marriage Jim and Wendy. Thank you for your incredible and unending help through one of the most challenging things I have done.

Finally, I want to thank my amazing wife Amanda and our two beautiful boys, Alexander and Andrew. Amanda, I thank you for your support over the last 13 years of marriage (of which I have been in school the entire time)! You amaze me with your tireless love and support. Thank you for putting the boys to bed on so many nights! Thank you for your delicious lunches that make the rest of the lunchroom jealous! Thank you that I have clean underwear! Thank you for your encouragement, endurance, patience, prayers, and faith. Without you, I would not be here.

Alexander and Andrew, you are both still young, but I thank you for understanding in your own ways when Daddy needed to work or go away on yet another trip. Thank you for always welcoming me into whatever you were up to, and focussing on the fun that we have together rather than how busy Daddy is.

Dedicated to

my Heavenly Father for loving me,

my Lord and Saviour for dying for me,

my Comforter and Counsellor for filling me.

Contents

List of Figures

List of Definitions

RGB	The standard way of representing colour digitally as a triplet of three values that represent the amount of red, green, and blue light to be added together to form the colour.
L*u*v*	The CIE 1976 perceptually-uniform colour space with three dimensions: L*=luminance, u* and v* are two chromatic dimensions.
Luminance	The relative perceived luminance of a colour (how bright the colour is, typically compared to white).
Chroma	The perceived colourfulness of a colour; increasing from achromatic (no chroma).
Hue	The dominant wavelength of a colour, e.g., red, yellow, blue, green.
CVD	Colour Vision Deficiency – a condition in which an individual is not able to differentiate colours that are typically differentiable. Can be inherited, acquired, and/or situationally-induced.
Recolouring Tool	An assistive technology that increases the differentiability of colours to help people with CVD.
SSM	Situation-Specific Modelling – a user modelling technique that captures the influence of external and internal factors on the user's performance through an in-situ calibration procedure.
ICD	Individualized model of Colour Differentiation – the application of the SSM approach to modelling a user's colour differentiation abilities.
Discrimination Ellipsoid	The representation of the colour differentiation abilities of a person by surrounding every colour in a colour space with an ellipsoid. Every colour within the ellipsoid is not differentiable from the central colour; every colour outside of the ellipsoid is differentiable from the central colour.
Differentiation Limit	The exact point at which two colours become differentiable as the differentiability of the two colours is progressively increased.

Chapter 1

Introduction

Colour provides most people with invaluable day-to-day information. We use colour to identify things we need (e.g., finding berries), assess health (e.g., diagnosing a rash or sunburn), and to gather information (e.g., identifying traffic signs). In digital environments, colours are often used to encode information. Examples of information encoding include Microsoft Word's differentiation of spelling and grammatical errors by their underline colour, information visualizations and charts using colour to identify different categories of information (e.g., Figure 1.1), and special colours being used to signify missing information in web forms (e.g., when missing entries are marked using red text).

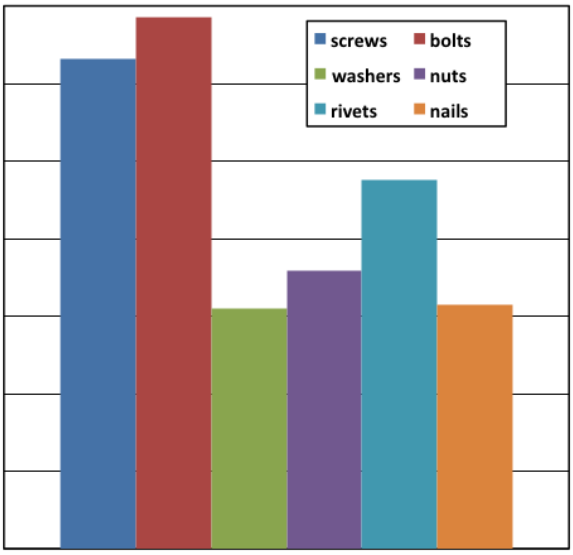


Figure 1.1: Using colour to encode chart information in Microsoft Excel for Mac 2011 (verson 14.1.3). Each chart category (screws, bolts, etc.) is represented using a unique colour.

In order for colour-encoded information in digital environments to be useful to a user, the colours presented must be differentiable: that is, the user must be able to tell the colours

apart. For example, differentiating between spelling and grammatical errors is made possible by utilizing differentiable colours for their respective underline annotations. Likewise, aligning a chart's data values with its legend's data labels is made possible by using a perceptibly distinct colour for each data category.

Although it is typically assumed that all people who use computers have identical colour perception abilities, many users have difficulty perceiving colour differences because of a variety of *colour vision deficiencies* (CVDs – commonly called colour blindness). These can be categorized as inherited, acquired, and situationally-induced CVDs. First, approximately five percent of the population worldwide has *inherited* colour vision deficiency, which typically results in difficulty differentiating between reds and greens (resulting in the common name of 'red-green colour blindness' for this condition) [10]. A simulation showing what someone with one type of inherited CVD sees when looking at the chart in Figure 1.1 is shown in Figure 1.2 (with the original chart included to ease comparison). Second, many individuals have *acquired* colour vision deficiency, most commonly caused by yellowing of the eye lenses with age (leading to difficulties differentiating between blues and yellows). Acquired deficiencies are common: for example, one study found that 64% of British participants over the age of 65 and 32% of African participants over the age of 40 had some degree of acquired CVD [25]. Third, most people experience *situationally-induced* CVD. Situationally-induced CVD arises when an individual's environment causes them to be unable to see colours properly [122], such as when trying to use a mobile device in bright sunlight while wearing sunglasses (see Figure 1.3). Other common causes of situationally-induced CVD include wearing tinted glasses or using improperly configured graphics hardware and software.

Each type of CVD has a variety of subtypes and severities. Inherited CVD can range from mild forms of *anomalous trichromacy* where individuals experience few day-to-day difficulties, to *monochromatism* which limits colour perception to shades of grey [10]. Acquired CVD is often progressive (e.g, as eye lenses become increasingly yellow with age) and can be complicated by conditions such as cataracts or macular degeneration [29] [1]. Situationally-induced CVD can be transient (e.g., bright sunlight) or persistent (e.g., uncalibrated display

[1]Macular degeneration is a medical condition which usually affects older adults and results in a loss of vision in the center of the visual field (the macula) because of damage to the retina. Definition taken from `http://en.wikipedia.org/wiki/Macular_degeneration`

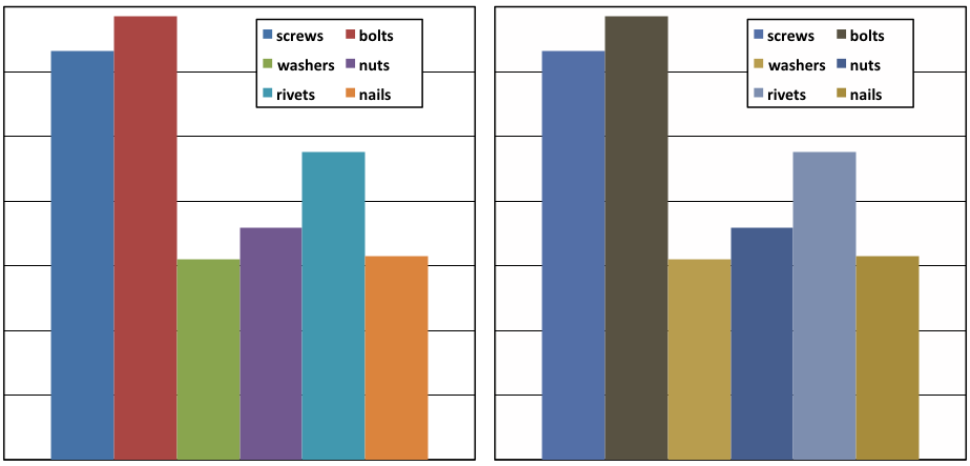


Figure 1.2: Left: the chart presented in Figure 1.1. Right: a simulation of what someone with one type of inherited CVD sees when looking at the chart on the left. The perceptual differences between some category colours (e.g., between 'screws' and 'nuts') is reduced.

hardware), and has a wide range of severities [160]. In addition to the wide variation in the nature of each type of CVD, it should also be noted that they are not mutually exclusive; CVDs can occur simultaneously. For example, a person with inherited CVD may develop acquired CVD as he/she grows older, and may also experience situationally-induced CVD.

In an attempt to address the problems caused by CVD, digital recolouring tools have been developed (e.g., [60, 74, 90, 116, 141]). *Recolouring tools* are software-based accessibility tools that modify the colours presented in an interface or image to improve the differentiability of the colours for the user. For example, a recolouring tool may replace a red-green gradient in an information visualization with a blue-yellow gradient for users who have difficulty differentiating between red and green. Example recolouring results for the chart shown in Figure 1.1 are shown in Figure 1.4.

To perform recolouring, recolouring tools utilize a model of the user's colour differentiation abilities to identify problem colours and to find suitably-differentiable replacement colours. *Problem colours* are any set of colours that are differentiable for people with typical colour vision, but are substantially less differentiable for users with CVD. *Suitably-differentiable* replacement colours are colours that have similar differentiability for the user with CVD as the original colours have for the user with typical colour vision. Suitably-differentiable colours are desired because the goal of recolouring is to restore the differentiability of the

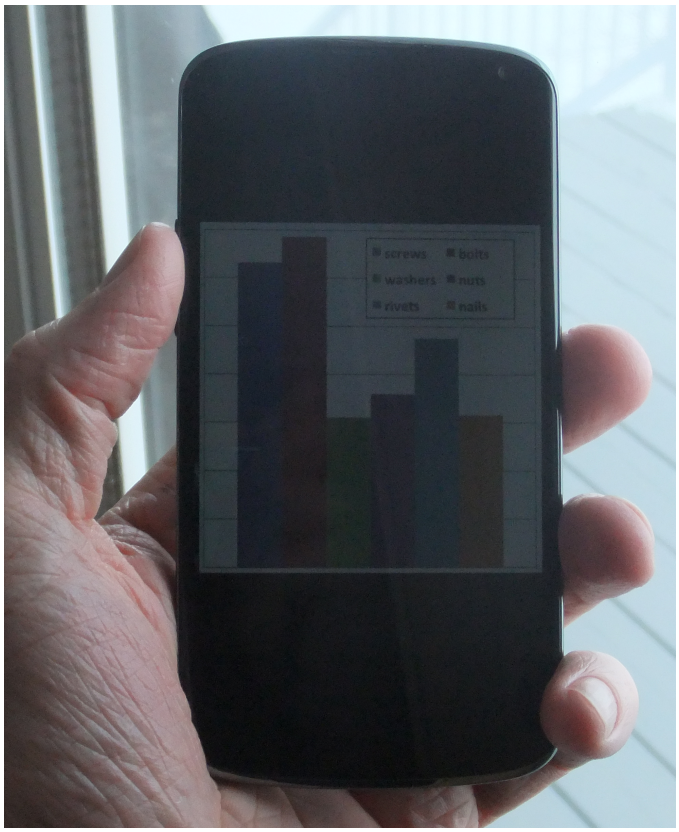


Figure 1.3: The chart shown in Figure 1.1 as displayed on a mobile phone in daylight. The luminance and colour contrasts of the chart is substantially reduced.

original colours for users with CVD. This is because the perceptual differences between colours can have meaning (e.g., when a colour gradient is used to represent ordinal information); failing to use suitably-differentiable replacement colours during recolouring risks destroying the meaning of the differences between the original colours.

As shown in Figure 1.5, recolouring tools are comprised of a 'recolouring algorithm' component and a 'model of colour differentiation' component. To perform recolouring, the recolouring algorithm first decomposes an input image into representative colours and uses the colour differentiation model to identify problem colours in the input image. The recolouring tool then generates candidate replacement colours for these problem colours and tests the differentiability of these candidate replacement colours using the model. This 'generate-and-test' step is repeated until suitably-differentiable replacement colours are found. Finally, the recolouring algorithm replaces the problem colours in the input image with the final replacement colours to produce a recoloured output image.

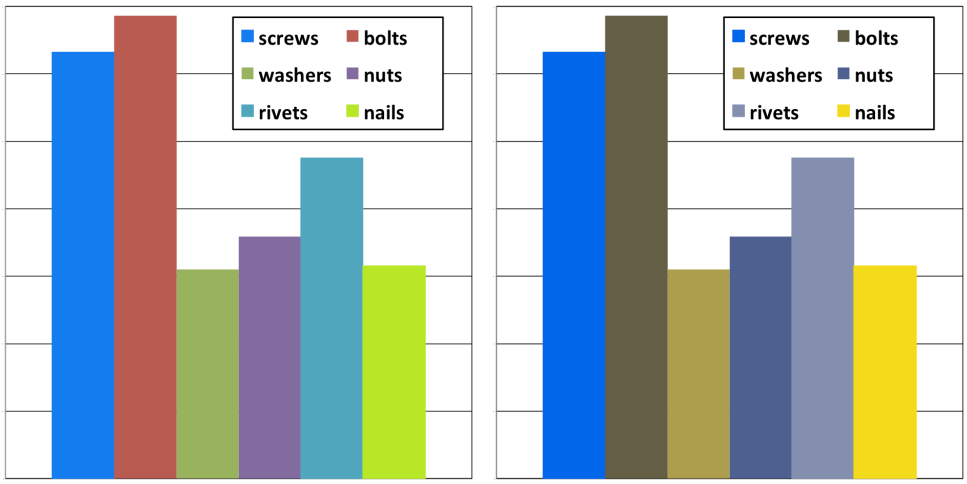


Figure 1.4: Left: recoloured version of the chart presented in Figure 1.1. Right: appearance of this recoloured chart to someone with the same inherited CVD as illustrated in Figure 1.2. Note that each colour is now perceptually distinct for a user with this type of inherited CVD.

However, existing recolouring tools do not help most people with CVD because current models of colour differentiation only represent a very small proportion of people with CVD. Many types of CVD do not have representative colour differentiation models, and for the types of CVD that do have models, the models have overly constrained requirements:

1. *Missing models.* No colour differentiation models have been developed for acquired CVD, situationally-induced CVD, simultaneous CVDs, and some forms of inherited CVD (e.g., monochromatism, extreme anomalous trichromatism [76]). These types of CVD result in unique and complex colour differentiation problems not well-described by existing models.

2. *Impractical models.* Models of inherited *dichromatic* CVD and *anomalous trichromatic* CVD [2] exist but require the specific type and severity of inherited CVD to be modelled. Individuals with inherited CVD are rarely diagnosed beyond having red-green colour blindness because of optometrists' reliance upon Ishihara's Tests for Colourblindness [71], which do not provide the specific information required by these models [10, 23]. As a result, few people with inherited CVD know the type and severity of CVD they

[2] Dichromacy and anomalous trichromacy comprise over 99% of all cases of inherited CVD [10]

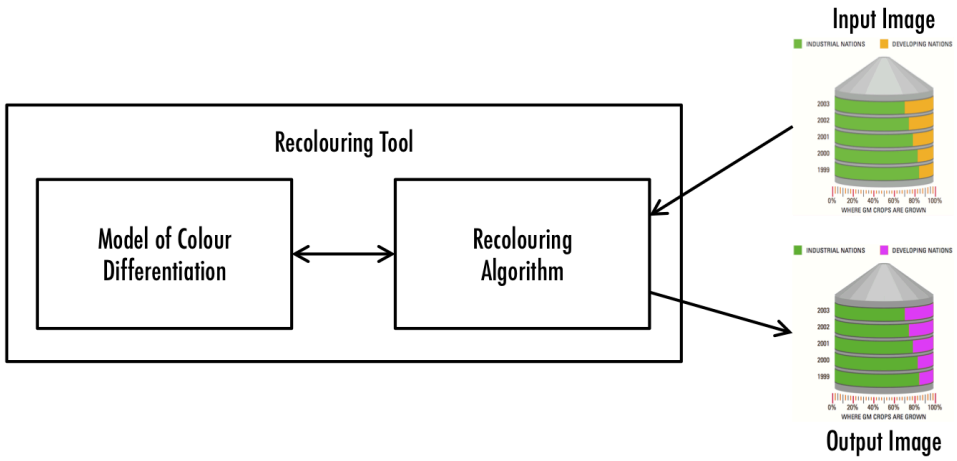


Figure 1.5: Basic recolouring tool architecture. Recolouring tool accepts an input image, consults a model to identify problem colours and find suitably-differentiable replacement colours, and produces a recoloured output image.

have, so cannot provide information necessary for the modelling process. Additionally, these models assume that the user does not have acquired or situationally-induced CVD, limiting the general applicability of these models.

People that are well-represented by the existing models (i.e., those with inherited CVD who know their type and severity and who do not have any simultaneous CVDs), are helped by existing recolouring tools, but this is a small proportion of the entire population of individuals with CVD.

1.1 Problem

The problem to be addressed in this dissertation is: *existing recolouring tools do not help most people with CVD*. Recolouring tools fail to help most people with CVD because the tools rely on models that only represent the colour differentiation abilities of a small proportion of the entire population of people with CVD. When the greater population of people with CVD attempt to use an existing recolouring tool, these individuals must guess the type and severity of inherited CVD that most closely matches their CVD. Either no correct match can be made or an incorrect guess is made, resulting in two general recolouring failures:

1. *Insufficient recolouring.* The model incorrectly identifies colours that are not differentiable for the user as 'differentiable'. This results in recolouring tools performing not enough recolouring – problem colours remain in the recoloured image.

2. *Incorrect recolouring.* The model incorrectly identifies colours that are differentiable for the user as 'not differentiable'. This results in recolouring tools performing the wrong recolouring – non-problem colours are replaced in the recoloured image. The inaccuracy of the model can even result in non-problem colours being replaced with problem colours.

An example of failed recolouring is shown in Figure 1.6.

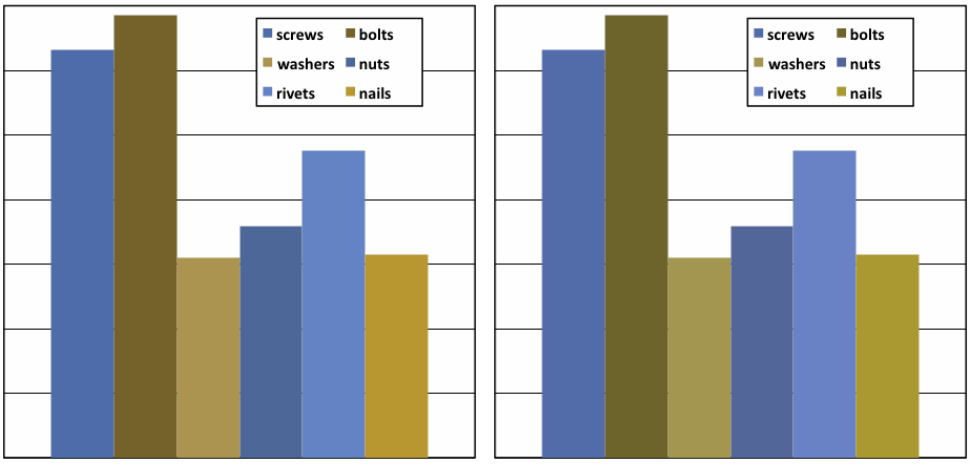


Figure 1.6: Left: Unsuccessful recolouring of the chart presented in Figure 1.1. Right: Appearance of the recoloured chart to someone with the inherited CVD illustrated in Figure 1.2, showing that this recolouring did not result in differentiable replacement colours. Note that every category colour in the recoloured image has been changed (incorrect recolouring), and that the perceptual difference between the category colours for 'screws' and 'nuts' is even smaller than in Figure 1.2 (insufficient recolouring).

1.2 Motivation

The failure of existing recolouring tools to help most people with CVD results in the persistence of difficulties caused by CVD – recolouring tools simply do not help. As a result, people with CVD continue to have difficulty differentiating the colours used to encode information, which results in users with CVD making more errors and taking more time when utilizing

colour-coded information [129] – implications of which have a wide range of severity. Lack of assistance for people with CVD causes minor nuisances such as being unable to tell 'visited' from 'unvisited' links on a webpage and having difficulty finding the missing online form entry that is highlighted using red text, but it also causes major safety concerns such as improperly reading information displays in supervisory control and data acquisition (SCADA) systems used in nuclear power stations and water treatment plants, and misreading navigation icons used by GPS-based navigation aids. The implications of these difficulties are as numerous and diverse as the problems themselves: clicking on an already visited link wastes the user's time, but misreading a colour-coded SCADA display in a nuclear power station could threaten the health and safety of millions of people.

1.3 Solution

To solve the problem of existing recolouring tools not helping most people with CVD, this dissertation presents the idea of *Situation-Specific Modelling (SSM)* and applies this approach to modelling the colour differentiation abilities of a particular individual in a particular situation. By using the resulting *Individualized model of Colour Differentiation (ICD)*, recolouring tools are able to more precisely identify problem colours and suitably-differentiable replacement colours, addressing the problems caused by insufficient and incorrect recolouring, thereby extending recolouring tools to help most people with CVD.

In general, situation-specific modelling is comprised of two steps: a calibration step and a modelling step. The calibration step empirically gathers data from the user that allows some characterization of his/her situation-specific abilities to be formulated. The modelling step utilizes this data to fit a model to the user's particular abilities. The resulting fitted model can then be used to provide personalized adaptations (such as recolouring) to improve accessibility.

To apply the situation-specific modelling approach to colour differentiation, this dissertation presents an in-situ, performance-based calibration technique. Everything that causes the colour differentiation difficulties a user experiences are present within the user (inherited and acquired CVD) and his/her environment (situationally-induced CVD), so calibrations

are performed in the user's situation (in situ). To calibrate, the user performs a series of in-situ colour differentiation tasks reporting whether two on-screen colours are differentiable or not. This performance-based technique frees the user from having to specify his/her type and severity of CVD, and is able to capture any type of CVD the user has (including simultaneous CVDs). The colour pairs shown during the calibration encode the colour differentiation abilities of the user for particular pairs of colours. To translate this data into a full model of colour differentiation, these data points are then extrapolated to all possible pairs of colours in the modelling step. The resulting individualized model of colour differentiation is a model that represents the precise colour differentiation abilities of a particular user in a specific situation.

1.4 Steps in the Solution

Five major steps were carried out to develop the individualized model of colour differentiation (ICD) presented in this dissertation:

1. SSM Idea: I developed the concept of situation-specific modelling including its calibration and modelling components. I then applied the concept of SSM to the ICD and clarified the requirements for utilizing ICD for real-world recolouring tools. This step resulted in a description of the constraints on ICD calibration and modelling techniques.

2. SSM Feasibility: In this step, I determined whether SSM is a feasible approach to achieving ICDs. This involved the development of an initial calibration procedure and model, which was evaluated in terms of calibration time, model sensitivity to environmental and user variations, and model over-specificity. This step resulted in a clear understanding of the feasibility of this approach, including specific identification of areas that required improvement before the application of ICDs to recolouring tools.

3. ICD Improvement #1: Reduce calibration time. The time to perform a calibration is a major factor in the usability of ICD-based recolouring tools. In this step, I explored how an individual's colour differentiation abilities can be rapidly measured, resulting in a

new performance-based calibration technique that is 24 times faster than the calibration technique for the feasibility implementation.

4. ICD Improvement #2: Reduce prediction time and increase prediction power. The faster the ICD model can make a differentiability prediction, and the more powerful each prediction is, the more efficient and powerful ICD-based recolouring tools will be. In this step, I explored how prediction speed and power can be improved by using affine transformations and psychometric-approximating functions in the ICD. This step produced the final implementation of the ICD for this dissertation that executes differentiability predictions 56 times faster than the Step 3 implementation, while providing real-valued predictions instead of the boolean predictions of the previous versions.

5. ICD Application: In this step, I developed and evaluated a recolouring tool based on the individualized model of colour differentiation (ICD). This step developed an ICD-based recolouring tool and compared it to two existing state-of-the-art recolouring tools by examining how well each helps participants match colours in a variety of environments. This final step resulted in a working recolouring tool that utilizes the individualized model of colour differentiation, and a comparative evaluation that showed that the solution presented in this dissertation improves colour matching performance by increasing accuracy by 20% and reducing selection time by two seconds when compared against two existing recolouring tools.

1.5 Evaluation

All but the first step (Step 1 – SSM Idea) presented above were evaluated by conducting lab-based user studies with CVD and non-CVD participants. As the first step presented the idea of situation-specific modelling and the individualized model of colour differentiation, a stand-alone evaluation of this step was not necessary because the evaluation of the remaining steps implicitly evaluated the idea of situation-specific modelling.

To evaluate the feasibility of the SSM approach (Step 2), sixteen participants (eight with CVD, eight without) completed the implemented ICD calibration procedure and then

provided a series of 'ground truth' data points against which the model predictions were compared. This evaluation was focussed on the feasibility of the SSM approach to achieving ICDs, and examined calibration time and quality, model sensitivity to inter-individual and environmental variations, and model over-specificity. This evaluation resulted in three main findings. First, I found that longer calibrations (more calibration data) produced more accurate models, but calibrations longer than 32 minutes did not result in substantial improvements in model accuracy. Second, I found that the feasibility implementation of the ICD was sensitive to differences between CVD and non-CVD participants as well as environmental variations. Third, I developed a tuneable parameter, the limit offset, to control the specificity of the SSM – lower limit offsets result in highly-specific models, but increasing the limit offset reduces the model's specificity thereby reducing the frequency of re-calibration.

To evaluate the reduction in calibration time (Step 3), sixteen participants (eight with CVD, eight without) performed calibrations both for the feasibility implementation and the reduced-calibration time implementation developed as part of this step. I then had participants provide a set of 'ground truth' data points by having them report whether two colours are differentiable or not for a large number of colour pairs. The calibration data was used to determine the calibration time savings accomplished in this step, and the ground truth data was used to compare the accuracy of the model implementations. This evaluation resulted in two main findings. First, I found that the new calibration technique resulted in a 24-fold reduction in calibration time (from over 32 minutes to under 2 minutes). Second, I found that the model resulting from the new calibration data was somewhat more accurate at predicting ground truth responses (76.1% for the feasibility ICD versus 78.7% for the reduced-calibration model developed in Step 3).

To evaluate the prediction time reduction (Step 4), I used the calibration data from the evaluation of Step 3 to generate a reduced calibration time model (Step 3) and the reduced prediction time model (Step 4) developed as part of this step. Differentiability predictions for one million randomly-generated pairs of colours were then conducted and timed to measure changes in prediction time. I also used the ground truth data from Step 3 to compare the prediction accuracy for each model version. This evaluation resulted in two main findings. First, I found that the new model was 56.4 times faster than the reduced

calibration time model at making differentiability predictions. Second, I found no significant differences in prediction accuracy, indicating that the prediction time improvements resulted in no loss of prediction accuracy. To evaluate the increased prediction power (also in Step 4), I visually inspected a sample recoloured gradient visualization and found that the additional prediction power adds value to the recolouring process by enabling the restoration of relative colour differentiability for the user with CVD. The reduced prediction time and increased prediction power improvements were both implemented in a single final version of the ICD.

To evaluate the recolouring tool implemented using the individualized model of colour differentiation (Step 5), I performed a comparative evaluation between this recolouring tool and two existing recolouring tools. Twenty-one participants (nine with CVD, twelve without) performed a series of colour matching tasks on two different sets of colours. The matching tasks were carried out in four environments (control plus three environments that caused situationally-induced CVD), and each colour set was recoloured using the developed recolouring tool and the two existing recolouring tools. This study explored two main questions:

- How is matching accuracy influenced by recolouring tools for a variety of situations, colour sets, and user colour differentiation abilities? Does an ICD-based recolouring tool allow users to match colours more accurately than existing recolouring tools?

- How is response time for a colour matching task affected by choice of recolouring tool?

This evaluation served as the final evaluation of this dissertation research. It answered whether situation-specific modelling-based individualized models of colour differentiation allow recolouring tools to be extended to a wide range of types and severities of CVD.

This evaluation resulted in two main findings. First, I found that the ICD-based recolouring tool resulted in participants making correct matches about 90% of the time, compared with 70% for existing recolouring tools. This 20% improvement was consistent across situations, colour sets, and CVD and non-CVD participants. Second, I found that the ICD-based recolouring tool allowed participants to make their matching selections about two seconds faster than existing recolouring tools. Like the accuracy result, this result was consistent across situations, colour sets, and CVD and non-CVD participants.

These results indicate that a recolouring tool that utilizes the SSM-based ICD as its

underlying model of colour differentiation is able to identify problem colours and find suitably-differentiable replacement colours more effectively than existing recolouring tools. Higher accuracy, fewer errors, and lower selection times indicate that the ICD-based recolouring tool enabled participants to make correct colour matches in less time than existing recolouring tools that do no use the ICD to represent the user's colour differentiation abilities.

1.6 Contributions

The central contribution of this dissertation is the identification of the situation-specific modelling (SSM) approach and the demonstration of how it can be applied to colour differentiation to produce the individualized model of colour differentiation (ICD). The ICD uses in-situ, empirically-based performance tasks to measure a user's ability to differentiate colour, and then utilizes these measurements to construct a prediction model that is specific to the user and his/her situation.

This dissertation also presents a number of secondary contributions:

1. A summary of the wide variety of factors that influence users' abilities to differentiate colours resulting in inherited, acquired, or situationally-induced colour vision deficiency (CVD).

2. An overview of the different applications of colour differentiation in digital systems.

3. An evaluation of the feasibility of using situation-specific modelling to develop individualized models of colour differentiation (ICD).

4. A performance-based ICD calibration technique that rapidly collects SSM calibration data.

5. Improved modelling techniques that reduce the time to perform a prediction and increase the power of a prediction for ICDs.

6. An implementation of an ICD-based recolouring tool.

7. An evaluation of the ICD-based recolouring tool showing ICD extends the applicability of recolouring tools to many types and severities of colour vision deficiency.

1.7 Overview of Dissertation

This dissertation contains the work described in this introductory chapter, presented in the following sequence of nine chapters:

Chapter 1: Introduction

Chapter 2: Related Work: Presents background research related to this dissertation. Includes necessary background on colour perception, factors that influence human colour differentiation, uses of colour differentiation in digital systems, and existing recolouring tools.

Chapter 3: Situation-Specific Modelling and Individualized Models of Colour Differentiation: This chapter presents the concept of situation-specific modelling (SSM). Details about using SSM to develop individualized models of colour differentiation (ICDs) through in-situ, empirically-based calibration are given. Design requirements of ICDs are also considered.

Chapter 4: Feasibility of Individualized Models of Colour Differentiation: Describes the design, implementation, and evaluation of a preliminary 'feasibility version' ICD calibration and model. This chapter describes the calibration and model implementations in detail, and shows that the calibration time, model accuracy, and model sensitivity are appropriate for use in recolouring tools. Areas for improvement are also discussed.

Chapter 5: Reducing ICD Calibration Time: Presents a new technique for modelling a user's colour differentiation abilities that allows a 24-fold reduction in calibration time over the feasibility calibration approach presented in Chapter 4. This chapter also presents a user-study evaluation that demonstrates the calibration time improvement as well as improved prediction accuracy.

Chapter 6: Improving ICD Prediction Time and Power: This chapter describes the final revisions to the ICD model that result in a 56-fold reduction in prediction time over the Chapter 5 ICD version and augment the ICD boolean `areDifferentiable`

function with a more powerful real-valued `howDifferentiable` prediction function with no reduction in prediction time. These extensions are evaluated using user-study data from Chapter 5 and visual inspection, respectively.

Chapter 7: ICD-Based Recolouring: Presents the final evaluation of the solution presented in this dissertation. This chapter describes the implementation of an ICD-based recolouring tool that is then comparatively evaluated against two existing recolouring tools by participants with and without CVD performing a colour matching task in a variety of situations. This chapter also presents results showing that the ICD-based recolouring tool enables significantly higher matching accuracy and significantly reduced match selection time.

Chapter 8: Discussion: Summarizes main findings from previous chapters, presents implications, challenges, and limitations of the SSM approach and ICD implementation presented in this dissertation. This chapter also discusses extensions of SSM beyond colour differentiation models.

Chapter 9: Conclusion and Future Work: Briefly summarizes the work presented in this dissertation and outlines future directions for this research.

Chapter 2

Related Work

The research in this dissertation is based on four main foundational areas: general colour perception, colour vision deficiencies, uses of colour in digital environments, and existing recolouring tools.

2.1 Colour Perception

Colour is an elemental experience for most human beings. In this section, I present fundamental concepts of colour perception and models that represent colour perception.

2.1.1 General Colour Perception

In the typical human eye, three types of photosensitive cells (cones) detect incident light on the retina at the back of the eye and convert this light into neural signals. These neural signals are passed through ganglion cells also located in the retina [2]. Ganglion cells aggregate signals from multiple cones and pass the resulting merged neural signals along the optic nerve to the brain. The optic nerve connects to the lateral geniculate nucleus (LGN) located inside the thalamus in the brain, where the signals are split into different visual pathways. These visual pathways carry visual neural signals to the visual cortex.

Each part of the visual processing pathway described above plays a particular role in colour perception. Different types of cones are sensitive to different wavelengths of light, thereby producing different signals for different-coloured stimuli. The ganglion (and other associated retinal cells) have been shown to contribute to the detection of opposite colours (e.g., red and green) [107]. The LGN treats signals originating from short-wavelength cones differently than signals originating from long- and medium-wavelength sensitive cones. The

visual cortex is responsible for colour identification, which aids with object recognition [5]. Two dominant theories of colour perception have been proposed to explain how humans see colour. Each of these theories is presented below, followed by an overview of digital representations of colour.

Trichromatic Theory of Colour Vision

The visible spectrum consists of electromagnetic radiation with wavelengths between 370nm and 730nm [130]. Each type of cone is sensitive to a different portion of the visible spectrum, with *short-wavelength-sensitive* (peak sensitivity at 419nm), *medium-wavelength sensitive* (peak sensitivity at 531nm), and *long-wavelength sensitive* (peak sensitivity at 558nm) cones [45]. In addition to cones, *rods* are another type of photosensitive cell in the retina. Rods are responsible for vision in low light, and rarely contribute to colour vision [10].

The trichromatic theory of colour vision (also called the Young-Helmholtz Theory) [36] is based on the idea that the three cone types are at the fundamental level of colour perception. The trichromatic theory of colour vision developed out of observations from nineteenth century colour matching experiments in which a match to a pure spectral hue light could be obtained from mixing three other spectral lights in appropriate proportions [45]. Normalized spectral response curves for each of the three types of cone are shown in Figure 2.1.

Opponent-Process Theory of Colour Vision

Although the presence of three types of cones in the eye has been accepted since the 1960s [24, 45], the trichromatic theory of colour vision does not explain some aspects of colour perception such as how some colours seem to be 'opposite' each other (e.g., red and green; blue and yellow). To address this, the *opponent-process theory of colour vision* was developed by Ewald Hering [57], which proposes three fundamental axes of colour perception: one achromatic axis (brightness), and two chromatic axes – one that spans the red-green range of colours, and one that spans the blue-yellow range of colours. Later research has uncovered specialized *ganglion* cells in the retina that receive signals from different cone types and produce inhibitory or excitatory responses depending on which cones are firing [136]. Translations from cone signals to opponent signals have been proposed [130], and one is illustrated in Figure 2.2.

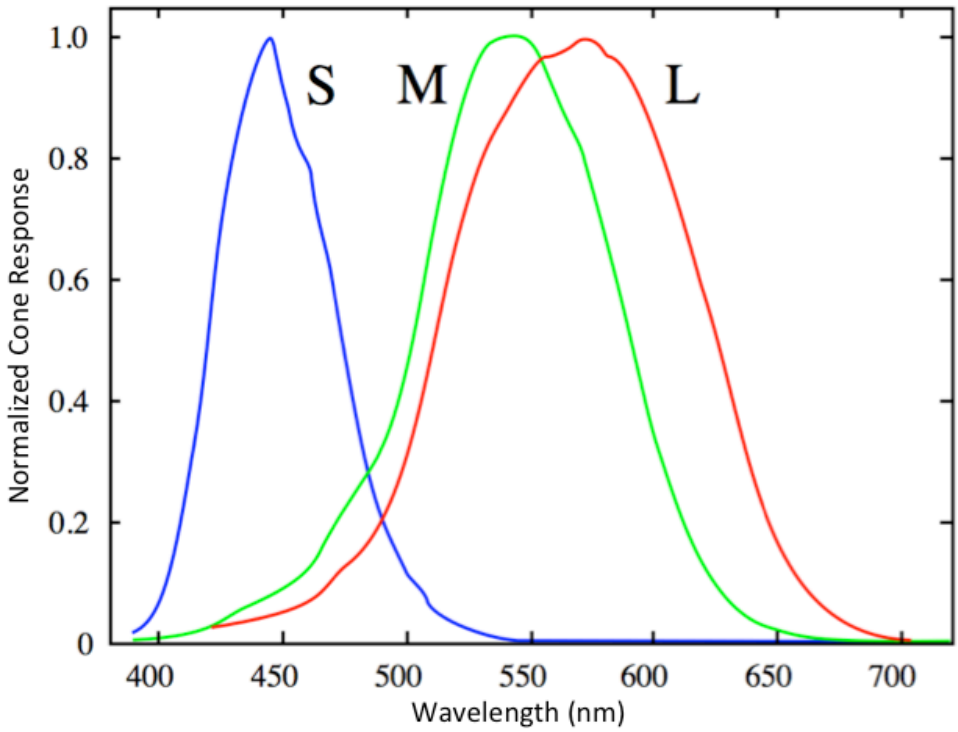


Figure 2.1: Normalized frequency response functions to wavelength for the three types of cones. Blue: short-wavelength sensitive cones. Green: medium-wavelength sensitive cones. Red: long-wavelength sensitive cones. Image from `http://en.wikipedia.org/wiki/Cone_cell`

2.1.2 Encoding Colour in Digital Environments

Digital representations of colour allow visual artifacts such as images to be produced, stored, transmitted, and viewed on digital devices. Cone-level responses and opponent-process opponent axes have been used to develop digital representations of colour. Some of these representations allow the perceptual differences between colours to be modelled for people with typical colour perception and those with CVD. As recolouring tools need to identify problem colours and find suitably-differentiable replacement colours, digital representations of colour are central to this dissertation.

RGB Colour

Stimulating the red, green, and blue (RGB) pixel components of a display produces colour [54]. RGB colour space (and its standardized version, sRGB) is used to represent pixel intensities and colours on typical displays, as well as in images and videos. Input and output images for

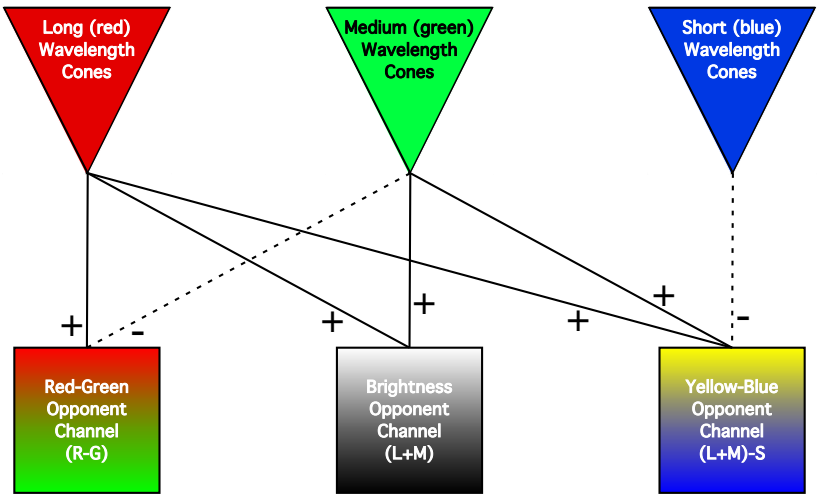


Figure 2.2: Translation of trichromatic cone signals to opponent process signals. The red-green channel is excited by **L** cones and inhibited by **M** cones. The luminance channel is excited by both **L** cones and **M** cones. The blue-yellow channel is excited by the **L** and **M** cones, and inhibited by the **S** cones. Image adapted from [130], pg. 28.

recolouring tools are encoded using RGB, so this colour space is always present at the beginning and end of the recolouring process. In between input and output however, recolouring tools often employ colour spaces that more closely represent how humans perceive colour. More details for the RGB colour space are given in Section 4.2.1.

XYZ, L*u*v*, and L*a*b* Colour Spaces

The International Commission on Illumination (CIE - Commission Internationale de l'Eclairage) has released many models of colour perception. Each model usually specifies colours using three components and so each of these models is known by a triple of characters. Three models that are significant to recolouring tools are the original CIE XYZ [124] specification and the perceptually-uniform spaces CIE L*u*v* [153] and CIE L*a*b* [112] derived from XYZ. The XYZ colour space is based on early colour science research that identified that any spectral colour could be matched by adjusting the intensity of three reference colours. The properties of X, Y, and Z are based on a standardized set of three reference colours, allowing precise communication about colour; colours with identical XYZ colour values are perceived identically. L*a*b* and L*u*v* are spaces that are considered perceptually uniform, and therefore can be used to predict human perception of colour difference using a simple Eu-

clidian distance metric [130]. Both L*a*b* and L*u*v* represent colour in a manner similar to the opponent-process theory with a luminance (L*) axis and two chromatic axis that correlate loosely with the red-green axis (a* and u*) and the blue-yellow axis (b* and v*). The revision of the individualized model of colour differentiation presented in Chapter 5 utilizes the L*u*v* colour space to reduce calibration time.

LMS Colour Space

Meyer and Greenberg [103] give a comprehensive description of a transformation that transforms CIE XYZ colour coordinates to a colour space that represents each colour as the normalized stimulation of the three cones in human colour vision. As discussed above, the three types of cones are classified as long-wavelength, medium-wavelength, and short-wavelength sensitive, hence this representation is called the *LMS* colour space. This colour space serves a particularly important role in simulating the appearance of CVD for individuals who do not have CVD, because the contribution of each cone to colour vision can be reduced or eliminated – which is what people with inherited CVD experience.

2.2 Colour Vision Deficiency

Recolouring tools exist to improve the lives of individuals with colour vision deficiency (CVD). In order to understand the problems experienced by people with CVD as well as potential solutions, overviews of the causes of CVD and models of CVD are now presented.

An individual has *colour vision deficiency* (CVD) when he/she perceives colours in a manner differently than the rest of the population (who have so-called 'typical' colour vision) [22]. Many factors can cause colour vision deficiencies, including *inherited*, *acquired*, and *situationally-induced* issues. Note that these types of CVD are not mutually exclusive and people can have a mixture of more than one type of CVD (called *simultaneous CVD* in this dissertation). There is considerable variation in the particularities of how CVD manifests in an individual [10], but the problems are similar in that they all relate to difficulty distinguishing between colours that the rest of the population has no trouble differentiating. CVD can cause difficulties whenever colour is used to indicate meaning [21, 22, 129].

As discussed further below, the use of colour in digital information displays is ubiquitous, and many presentations require that the user be able to tell different colours apart. For example, information visualization uses colour for categorical encoding, highlighting, popout, and representation of continuous variables [135, 146]. Everyday interfaces also use colour extensively – for example, to show visited links in web pages, or to clearly indicate alerts and warnings. A user with CVD may experience challenges when interacting with any of these interfaces.

2.2.1 Inherited CVD

As discussed above, human colour perception begins with light stimulating the three types of cones in the retina. Any deviation from the 'typical' properties of these three cones can result in variations in colour perception. Sometimes these variations are substantial enough to be classified as inherited CVD, but there are variations within the population of people who have 'typical' colour vision as well.

In non-CVD populations, variations in the genetic codes that determine the spectral sensitivity of cone cells cause people to perceive colour differently [143]. For example, many men have multiple genetic encodings of their long-wavelength and medium-wavelength photoreceptive proteins leading to variations in male colour vision [111]; and there is evidence that some female carriers of CVD genes may possess four types of cones instead of three [72].

There is also considerable variation within the population of those who have CVD. These variations fall into two categories: those with cones that have a shifted peak sensitivity and those missing an entire cone type. There are three types of sensitivity shifts depending on which cone class is affected: *protanomalous* (long wavelength), *deuteranomalous* (medium), and *tritanomalous* (short) [10]. Protanomalous and deuteranomalous CVDs make up almost 75% of all inherited CVD [20]. The severity of each of these three forms of anomalous trichromacy ranges from no difficulties with day-to-day life to frequent colour perception problems [20]. There are also reported cases of individuals who do not fall into one of these three categories, and are diagnosed with extreme anomalous trichromacy [76].

Individuals missing an entire type of cone have dichromatic CVD: *protanopia* (long-wavelength cones), *deuteranopia* (medium), and *tritanopia* (short) [10]. Dichromatism consti-

tutes about 25% of cases of inherited CVD, and manifests in daily life with frequent difficulties identifying, matching, and reproducing colour [20]. Rarer forms of congenital CVD include *cone monochromatism* (two missing cone types), and *rod monochromatism* (no cones at all); individuals with these CVDs only perceive shades of grey [10].

As protan and deutan conditions represent the vast majority of cases of inherited CVD, this condition is commonly termed 'red-green colourblindness'. From the opponent-process schematic presented in Figure 2.2, it can be seen that reduced or missing functionality of either the long-wavelength or medium-wavelength sensitive cones will reduce an individual's ability to distinguish between reds and greens (either the exhibitory or inhibitory component of the opponent-process red-green channel is reduced/missing). This results in individuals with these conditions having difficulties distinguishing between colours that differ solely in their amount of red or green. It also should be noted that these types of CVD do not diminish the ability to distinguish luminance or colours that vary along the blue-yellow channel. An example illustrating the difficulties resulting from protan and deutan CVDs is shown in Figure 2.3.

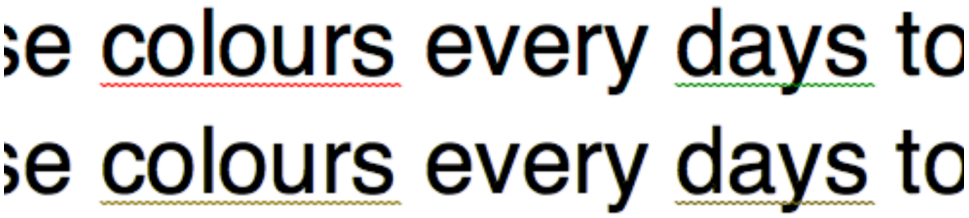


Figure 2.3: Top: Microsoft Word uses red to underline spelling errors and green to underline grammatical errors. Bottom: these colours present difficulties for people missing their medium-wavelength cones (deuteranopia).

2.2.2 Acquired CVD

Acquired CVD results from damage to the visual system from external events such as accidents, disease, or exposure to harmful chemicals. Acquired CVD can result from factors that affect the light entering the eye (e.g., yellowing of the lens with age or cataracts [9, 29, 38]), as well as neurological damage to the retina or visual processing centres of the brain [160]. Retinal damage can result from premature birth, long-term diabetes, hypertension, macular

degeneration, or long-term exposure to organic solvents like styrene [94]. Neurological damage can arise from stroke or aneurism, as well as traumatic brain injury [15]. Acquired CVD can be short-term as well, being introduced by prescription drugs (e.g., Viagra, antidepressants) or as a side effect of depression [56].

Although acquired CVD arises from a diverse set of causes, it typically manifests as difficulties differentiating between blues and yellows. Yellowing of the eye lens and cataracts effectively result in the light entering the eye passing through a yellow filter before hitting the retina. A yellow filter reduces or removes the shorter wavelengths (blues) of the light entering the eye. Similarly, cases of retinal damage often affect the short-wavelength sensitive cones more substantially because these cones are relatively under-represented in the retina (comprising only about 10% of the cones [143]). In both cases, the short-wavelength sensitive cone response is reduced, resulting in decreased perceptual difference between blues and yellows, which is identical to what results from tritanopia inherited CVD, shown in Figure 2.4.

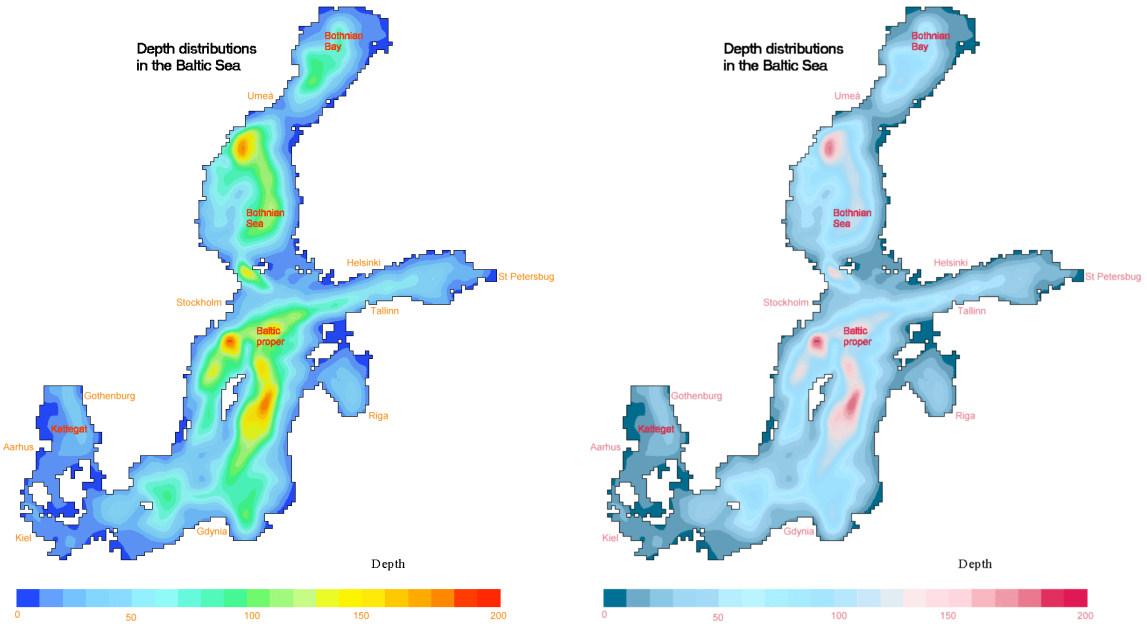


Figure 2.4: Left: using a rainbow colour scheme to illustrate depths in the Baltic Sea. Right: how this visualization appears to someone with severe acquired CVD or tritanopic inherited CVD. Original image from `http://nest.su.se/baltic96/depth.htm`

2.2.3 Situationally-Induced CVD

Aside from the acquired and inherited factors that influence human colour perception after light enters the eye, anything that influences light *before* it enters the eye can also induce a form of colour vision deficiency called *situationally-induced CVD*. This form of CVD is caused by factors in the user's local environment that result in temporary changes to colour perception. If the resulting changes cause problems with colour discrimination or identification, a 'situationally-induced disability' results [122].

Any environmental factor that influences the brightness or spectral distribution of light entering the eye can induce a situational CVD. Vision research classifies environmental brightness into *photopic* (bright), *mesopic* (medium), and *scotopic* (dark) light levels [126]. In scotopic conditions, only the rods function, and no colour perception occurs. At mesopic light levels, both the rods and cones contribute to colour perception, thereby changing how we perceive colours (e.g., the Purkinje Effect when blues appear brighter than normal at dawn and dusk [143])). At photopic light levels, colour perception typically works normally, but very bright environmental light (e.g., midday sunlight) can overwhelm the capabilities of devices such as LCD displays, such that colours are no longer discernible on the display.

In addition, factors that influence the spectral distribution of light entering the eye can change colour perception. This happens when lighting is coloured (e.g., coloured spotlights), or when wearing tinted glasses or contacts [51]. In digital environments, this factor can occur when display hardware or graphics software is not properly configured or is failing. An example of situationally-induced CVD was shown in Figure 1.3 and is reproduced here in Figure 2.5.

2.3 Models of Colour Vision Deficiency

As identified in Chapter 1, recolouring tools need to identify problem colours (colours that are difficult to differentiate for the CVD user) and identify suitably-differentiable replacement colours. In order to do this the digital representations of colour described above are needed, but so are models of how colours appear to people with CVD. To achieve this, simulations of

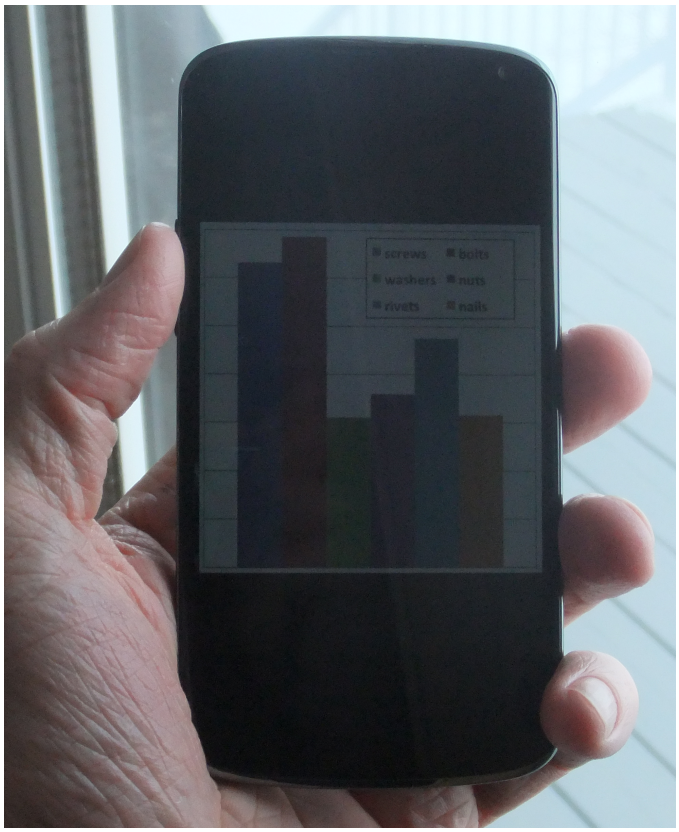


Figure 2.5: The information in a chart displayed on a mobile phone is difficult to read due to situationally-induced CVD caused by the eye adapting to bright outdoor light, thereby reducing the perceived brightness of the display.

the appearance of colours to people with CVD are utilized. These simulations are generally used to help people with typical colour perception understand the colour perception of people with CVD.

Current simulations of CVD are based on an algorithm first presented by Meyer and Greenberg [103] and later updated by Vienot et al. [16, 139]. This algorithm allows the simulation of dichromatic colour perception for individuals without CVD by the following steps:

1. RGB $\rightarrow$ LMS: Using a pre-defined linear transformation, translate the original image pixel colours (encoded in RGB) into a colour representation that encodes colour as stimulation levels for the three types of cones (LMS colour space, described above).

2. LMS $\rightarrow$ LMS*: Manipulate the LMS representation by removing the appropriate wavelength information for the desired type of dichromatic simulation (e.g., long-wavelength

for protanopes).

3. LMS* → RGB*: Using the inverse of the linear transformation of Step 1, translate the modified LMS colours back to the original colour representation.

This simulation technique results in the three dimensional colour space for typical colour perception being reduced to a two dimensional plane. For those with 'red-green CVD', the compression occurs along the red-green axis, thereby reducing the contribution of red-green differences to colour differentiation. For those with 'blue-yellow CVD', the compression occurs along the blue-yellow axis, thereby reducing the contribution of blue-yellow differences to colour differentiation.

To simulate the appearance of an image for someone with dichromatic CVD, the colour of each pixel in the image is replaced by the colour perceived by the person with CVD. As people with CVD perceive fewer colours than people without CVD, this process of mapping original colours to replacement colours is typically compressive – different original colours will map to the same replacement colours, e.g., people with protanopia perceive particular shades of pink, grey, and turquoise all as a single grey. To identify a replacement colour, both the set of colours that are perceived identically, and the colour this set maps to need to be identified.

To find the set of colours that are perceived identically for people with dichromacy, CVD colour confusion lines are utilized. In the CIE L*u*v* perceptually-uniform colour space described above, every colour has a sphere around it that defines a discrimination boundary for that colour for people without CVD – colours inside the sphere are not differentiable from the colour at the center of the sphere and colours outside the sphere are differentiable from the colour at the center of the sphere [114]. When measuring the discrimination spheres for people with dichromacy, it was found that the sphere had been elongated in one dimension to form an ellipsoid [96] that exceeds the L*u*v* gamut, and the direction of elongation was unique to the type of dichromacy. The line defined by the elongation (the major axis of the ellipsoid) is a colour confusion line – the line or set of colours that are perceived identically by someone with dichromacy [130, 153]. Confusion lines for each type of dichromacy are shown in Figure 2.6.

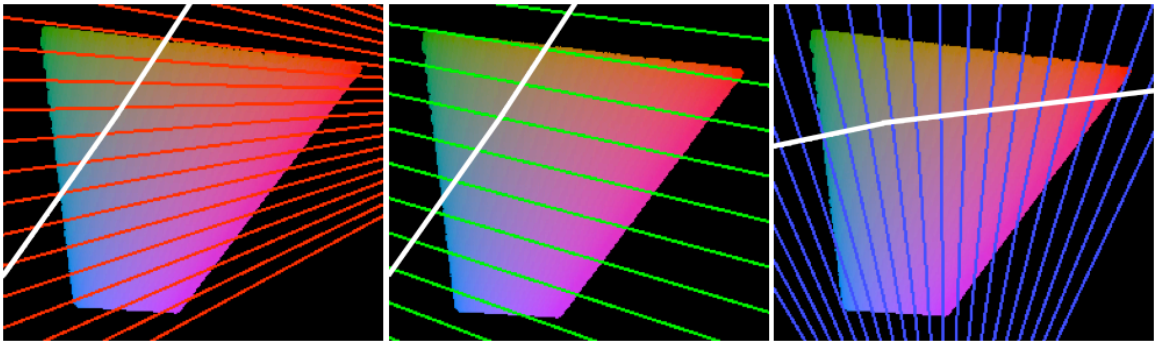


Figure 2.6: Left: red lines are colour confusion lines for protanopia. Middle: green lines are colour confusion lines for deuteranopia. Right: blue lines are colour confusion lines for tritanopia. In each image, white lines represent the set of colours for which there is no difference between how people with inherited CVD perceive each colour and how people with typical colour vision perceive that colour.

Confusion lines allow the identification of original image colours that are indistinguishable for someone with dichromacy, but what colour does someone with dichromacy actually perceive? This is akin to knowing that a man with inherited CVD confuses red and green, but not knowing what he actually sees. To answer this question, a special set of colours that are perceived identically by people with and without dichromacy has been identified using measurements from people with unilateral dichromacy – a condition in which a person has dichromacy in one eye but ordinary colour vision in the other eye. People with unilateral dichromacy have identified that spectral colours of 475nm (blue) and 575nm (yellow) are identically perceived by people without CVD and those with protanopia and deuteranopia. People with tritanopia perceive spectral colours of 485nm (blue-green) and 660nm (red) the same as people without CVD [16, 47, 77, 127]. Each pair of identically-perceived spectral colours define two half-planes in L*u*v* colour space. Each half-plane is defined by the spectral colour and the achromatic axis (grey scale colours from black to white) [16, 103]. These half-planes are shown in Figure 2.7 (protanopic and deuteranopic on the left, tritanopic on the right), and are represented by the white line in each image in Figure 2.6.

To simulate how an image appears to someone with dichromacy, the colour of each pixel in an image is converted into L*u*v* colour space, shifted along its respective colour confusion line to the half-plane for the dichromacy being simulated, and then converted back to RGB. The resulting RGB colour is used to paint the corresponding output pixel in the simulation

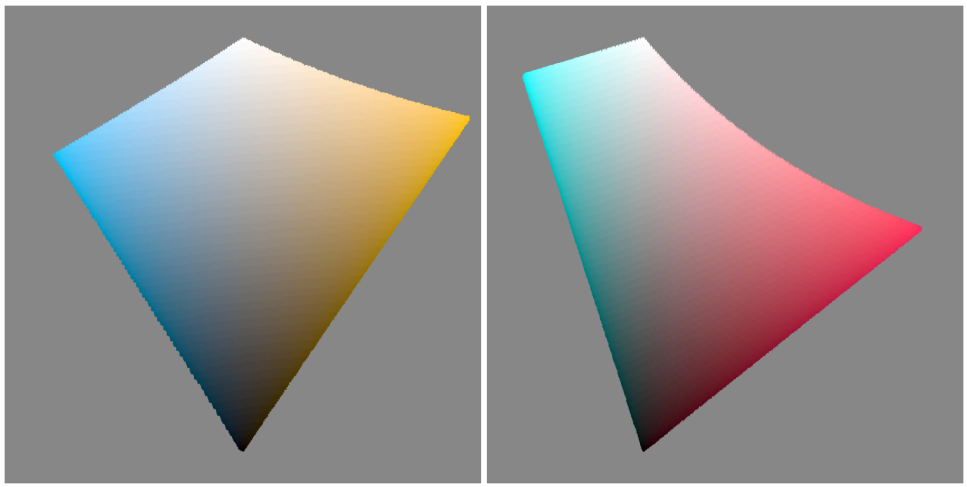


Figure 2.7: Colours perceived identically by people with dichromacy and people with typical colour vision; protanopia and deuteranopia (left) and tritanopia (right).

image. This approach has been extended to produce colour maps for displays which provide a means to simulate the appearance of the entire display to individuals with dichromacy [138, 137].

The conversions from RGB to L*u*v*, L*a*b*, or LMS colour spaces require an intermediate conversion into XYZ colour space. In [59], a method for simulating dichromacy directly within XYZ colour space is presented. This eliminates the conversions to and from these other spaces, reducing simulation computational load. An exact understanding of the set of colours shared between individuals with dichromacy and individuals without CVD is not yet available. To explore this, Capilla et al. [17, 18] propose a new method for estimating this gamut and constructing a new dichromacy simulation based on the estimate. Another shortcoming of the current model is that individuals with dichromacy use colour terms such as 'red' and 'green', yet the current models do not predict that dichromats should use these colour terms. This discrepancy was explored in [140] in which a new model of dichromacy was developed that more accurately predicts when dichromats will use these colour terms. Most recently, Rodriguez-Pardo and Sharma [120] proposed a two-stage model for simulating dichromacy that incorporates a trichromatic sensor stage first, followed by an opponent-process stage.

The simulations described above apply to dichromacy only, but simulations of anomalous trichromacy have been developed as well. These simulations require information about the severity of anomalous trichromacy (measured in shift of peak wavelength for the anomalous

cone). Kondo [84] developed the first simulation of anomalous trichromacy and presented results from taking the Farnsworth-Munsell 100-Hue [31] colour vision test under simulation to show the validity of his approach. This work was extended by Walraven and Alferdinck [142] to produce colour maps that allow automatic dichromatic and anomalous trichromatic simulation of the entire contents of the display. As another means to simplify simulation of anomalous trichromacy, Martin et al. [100] propose a new colour space that allows simple simulation of deuteranomalous colour vision by setting one dimension of this space to a constant value. More recently, work has been done to produce high-resolution, high-speed simulations of many types and severities of anomalous trichromacy [97, 99]. Each of these models can simulate the effects of anomalous trichromacy, but require an exact shift in peak wavelength sensitivity for the anomalous cone, which is not easily obtained.

To address the difficulty of determining the severity of anomalous trichromacy to be simulated, two computer-based anomalous trichromacy tests have been developed. Early work by Wenze, et al. [148] describes a computer-based test for CVD that found that anomalous trichromats and non-CVD participants have the same luminance contrast sensitivity, but different colour contrast sensitivity, especially between reds and greens. Yang et al. [158] proposed a computerized hue test that measures type and degree of anomalous trichromacy, then simulated how their computerized hue test appears to people with CVD.

2.4 Colour Differentiation in Digital Environments

Colour plays a major role in the presentation of visual information, both in everyday graphical interfaces and in specific visualization applications. There are many ways that colour is employed, but three main uses that involve colour differentiation are encoding categories, encoding continuous variables, and highlighting specific items [53, 135, 146].

2.4.1 Categorical encoding

Visual processing in the human visual system allows rapid identification of colours [146]. Labelling objects with colour thus can allow categorical information to be identified quickly and efficiently. In categorical encoding, a unique colour is assigned to each category of data,

and all representations of this category in the visualization will then employ this colour as an identifying characteristic. Colour as category is used in a number of information displays, including charts in spreadsheets, 'link taken' encodings in web browsers, syntax colouring in programming editors, and tagged messages in email clients.

Healy [53] suggests that a maximum of seven category colours can be used if the luminance of the colours is held constant; with variations in luminance allowed, the number of unique categories is likely to increase. Figure 2.8 contains an example with twenty categories, which exceeds these maximum category numbers.

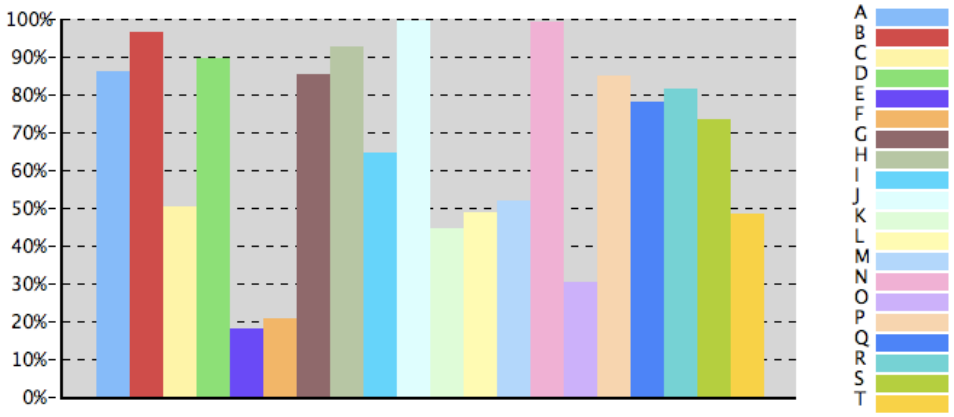


Figure 2.8: An extreme example of categorical encoding using colour.

2.4.2 Popout, Brushing, and Highlighting

A special case of categorical encoding involves temporary changes to the colour of objects that are considered special (Figure 2.9). Colour popout is a visual phenomenon in which colour makes elements stand out in an obvious fashion from the rest of the data. The popout colour must be sufficiently different from other colours in the visualization for the effect to work: generally a saturated, bright, primary colour is used to replace the established element colour. Due to the stark contrast between a popout colour and all other colours in a visualization, the elements marked with the popout colour can be identified with little effort, allowing the marked items to be easily located [146].

Brushing is the interactive application of a popout colour to a visualization with numerous data points. The user of a visualization marks elements of interest so they remain easily

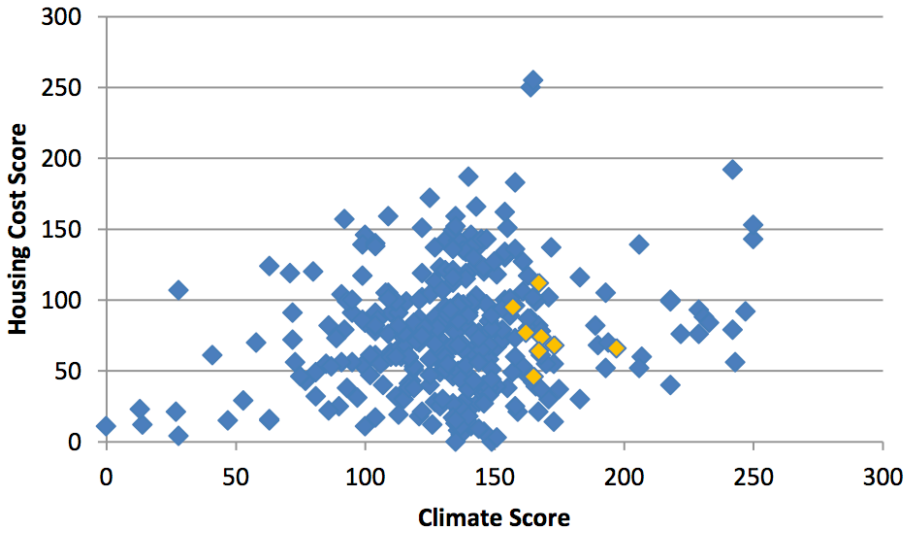


Figure 2.9: Brushing in a scatterplot visualization.

discernible while the data is manipulated (e.g., rotated in a 3D scatterplot).

Highlighting is the use of colour to bring attention to an element or region of a visualization. Unlike popout, highlighting does not replace the element colour in the visualization, but surrounds the element of interest. As a result, desaturated colours are often used to prevent the highlight from occluding the highlighted item.

2.4.3 Encoding continuous variables

Colour is also often used as a means to encode univariate or multivariate data. This process involves the discretization of a continuous data set, where each of the discrete ranges is associated with a given colour, and the data set is painted accordingly. For example, the depth of a body of water can be encoded using shades of blue, where darker blues indicate deep water and lighter blues show shallow water. This approach is used in several techniques: false colour representations (mapping to hue, as in Figure 2.10), continuums (mapping to luminance and saturation), and multidimensional data display (colour as dimension) [147].

Mapping an ordinal set of data to a colour representation requires the invention of a hue scale. This scale must be learned before the visualization can be used, which can reduce usability [130, 146]. Another difficulty with hue scales is simultaneous contrast, which occurs when the perception of a colour is influenced by surrounding colours [130].

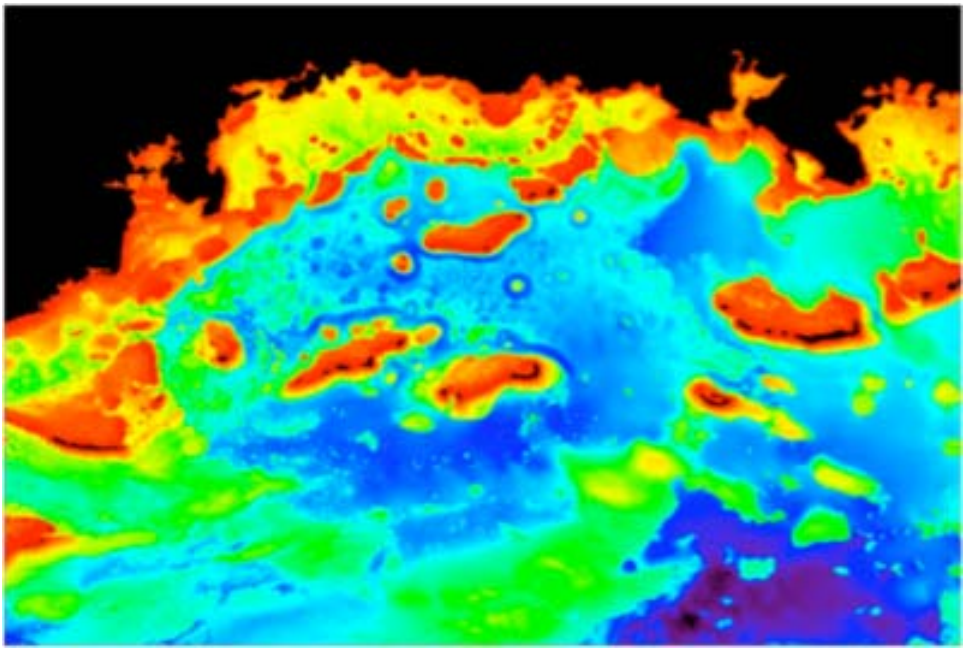


Figure 2.10: False colour image showing water depth in the Bight of Parguera, Puerto Rico. From `www.rsmas.miami.edu/groups/coral-lab/research-projects/la-parguera`.

2.5 Recolouring Tools Overview

Several systems exist to help address the problem of colour differentiation in digital displays. These systems, called recolouring tools, select colours in an image that are likely to be problematic for the user, and switch these to different colours that are more likely to be differentiable. From an early proposal by Meyer and Greenberg [103], several methods have been developed including SmartColor [141] and a number of techniques for dealing specifically with photographic images (e.g., [60, 90, 116, 117]). In some cases the user can participate in the recolouring process – for example, some forms of CVD (such as anomalous trichromatism) are less severe than dichromatism, and an interactive system can allow these users to guide the recolouring process [74, 75].

These tools all replace problematic colours with suitably-differentiable replacements, but some tools increase the intensity of colours to help people with anomalous trichromacy. By inverting the simulation for anomalous trichromacy (described above in Section 2.3), these recolouring tools 'amplify' image colours, making them more colourful. When someone with

anomalous trichromacy looks at the amplified colours, his/her visual system reduces this amplified colourfulness, resulting in a perception similar to what people with typical colour vision experience [104, 105, 106]. However, colour amplification often results in colours being pushed outside of the display's gamut of colours. When this happens, the colours have to be 'pushed back' into the gamut in order to be displayed, thereby nullifying the 'amplification' [103], rendering colour amplification ineffective. Failure of this amplification approach leaves the replacement recolouring tools described above as the only viable option for recolouring tools.

Recolouring tools that replace problematic colours with suitably-differentiable colours all rely on an underlying model of the user's colour differentiation abilities, and most current models are based on the early assumption-based algorithm described above in Section 2.3. This algorithm allows the simulation of dichromatic colour perception for individuals without CVD. Using this simulation, recolouring tools perform recolouring by performing the following four steps:

1. Reduce the colours in the original image to a representative set of colours. This step helps reduce the computational load by reducing the number of colours that need to be considered by the algorithm.

2. Through simulation of the representative colours, identify colours that are less differentiable in the simulation than in the original image.

3. Iteratively search the colour space to find candidate replacement colours. Using the simulation, determine whether the candidate replacement colours are suitably differentiable.

4. When suitably-differentiable replacements have been found, apply them to the original image to produce a recoloured image that is more differentiable for the user with CVD.

The current models of CVD that produce the simulations used in existing recolouring tools use several assumptions, however, that limit the applicability of the model. The various transformations between colour spaces require exact measurements of the colour characteristics for each of the red, green, and blue elements that make up the display, as well

as their response curves (gamma) and the whitepoint of the display (the characteristics of the white colour produced when each channel is set to maximum output). These can be measured and captured in an ICC colour profile through calibration [130], but few people own the specialized equipment required to perform calibration. Additionally, calibration is performed when people notice the colours on their display do not appear correctly, something that someone with a reduced colour perception gamut (e.g., someone with inherited CVD) may not notice. The transformations between colour spaces also assume a 'representative' human colour vision system, but variation between humans is well-documented [111].

Most importantly, the existing simulation techniques only handle certain forms of inherited CVD, and require the precise type and severity to do so. Existing models do not handle other forms of inherited CVD (such as extreme anomalous trichromacy [76] or monochromacy [10]), and cannot represent acquired or situationally-induced CVD.

Existing recolouring tools work well when the underlying model of CVD represents the colour differentiation abilities of the user – and the user knows the type and severity of CVD that they have – but as described above, this is rarely the case. As a result existing recolouring tools fall short of being a broad, reliable, and complete means of helping people with different types of CVD.

2.6 Individual Recolouring Tools

Since the earliest recolouring tools first developed in 2003, many improvements to recolouring tools have been proposed. These improvements can be categorized using the following distinctions:

- *Application*: recolouring tools have been proposed for websites, images, user interfaces, videos, and real-world situations;
- *Naturalness*: recolouring tools attempt to maximally retain original colours during the recolouring process;
- *Speed*: the speed of recolouring has improved from hours to milliseconds, allowing real-time recolouring;

- *Approach*: two main approaches to recolouring have been proposed: *replacing* original colours with other (possibly dissimilar) colours and *amplifying* the original colours. Many variations for these approaches have been proposed.

2.6.1 Early Recolouring Tools - Defining the Landscape

Two recolouring tools were proposed in 2003 [69, 128, 155]. Ichikawa et al. focused on recolouring websites to improve colour differentiability for individuals with dichromacy [69]. To accomplish this, the website's HTML is parsed to construct an 'abstracted image model', which organizes coloured regions into a hierarchy. Using this abstracted image model, the colours for adjacent regions with low colour contrast for individuals with dichromacy are replaced. The authors develop a fitness function that steers a genetic algorithm (self-reproduction with mutation) to find replacement colours that are as similar as possible to the original colours (to maintain naturalness), but still differentiable for individuals with CVD. Through a user study with four participants with CVD, it was shown that this approach provided more differentiable websites. Example recoloured websites are shown in Figure 2.11.

Figure 2.11: Sample website used in evaluation of [69]. Also shows simulated protan and deutan appearance of this website, recoloured versions, and simulation of the recoloured versions. Images copied from [69].

Ichikawa's website recolouring tool uses a genetic algorithm to optimize the replacement colour set according to a number of constraints to restore colour differentiability for users

with dichromacy. Another approach is to apply simple heuristics to identify problem colours and to subsequently modify them [128, 155]. Problem colours in images can be identified by their hue in HSV colour space and modified by adjusting their hue and saturation. This approach is much faster than Ichikawa's optimization approach but can dramatically change colours in an image, reducing its naturalness. This faster approach was evaluated using a quantitative comparison of the root-mean-square colour differences between the original image and the dichromatic simulation of the original image, and between the original image and the dichromatic simulation of the recoloured image.

In addition to proposing an alternative to Ichikawa's dichromatic colour replacement approach, these authors were the first to propose the amplification recolouring approach for people with anomalous trichromacy. This works by modifying the colours in an image to increase the stimulation of the anomalous cone. As anomalous trichromacy results in an anomalous cone that is under-stimulated in comparison with its non-CVD counterpart, overstimulating the anomalous cone results in a user with anomalous trichromacy perceiving an image in a similar manner as individuals with typical colour vision do. This approach was also extended later (details below, Section 2.6.5).

Each of these early approaches identify the areas for recolouring tool improvement outlined above: application, naturalness, speed, and approach.

2.6.2 Applications of Recolouring Tools

Recolouring tools have been applied to many different situations which contain colour – from general images and websites to recolouring objects in the real world. This has been an important development for recolouring tools because the problems individuals with CVD experience with colour are not limited to particular situations, but extend to almost all areas of life.

Static Images

Extending their earlier recolouring tool for websites, Ichikawa et al. created a recolouring tool for any image [70]. To find an optimal recolouring, they utilized random bit-climbing [26] constrained again to maintain naturalness from the original image to the recoloured image.

This recolouring tool was evaluated in a user study in which the naturalness and colour discriminability of the recoloured images were assessed. The recolouring tool was able to improve colour discrimination, but had difficulties maintaining naturalness. An important contribution of this work was the exposition of the basic three-step algorithm for colour replacement recolouring tools, shown in Algorithm 1. Modifications to the steps of this algorithm comprise the majority of the contributions in this research area since Ichikawa et al.'s original work.

1. Reduce the original image to a set of representative *key colours*.
2. Determine a mapping from key colours to replacement colours.
3. Use this recolour map to repaint the original image.

Algorithm 1: Basic three-step recolouring tool approach

A web-based tool called Daltonize [30] was developed to provide a widely-available recolouring tool for protanopic and deuteranopic individuals, by performing recolouring on any uploaded or website image. To perform a recolouring, the user sets the following three variables:

1. *Red-green stretch factor*: how much reds should be made more red and greens should be made more green.
2. *Luminance projection scale*: how much should red-green variations be mapped to luminance variations (e.g., reds get darker and greens get lighter).
3. *Blue-yellow axis projection scale*: how much should red-green variations be mapped to variations on the blue-yellow axis (e.g., reds become blues and greens become yellows).

No additional details are given for how the recolouring is performed, but example images from this recolouring tool are shown in Figure 2.12.

Figure 2.12: Demonstration of the Daltonize [30] recolouring tool. From left to right: original image containing problem colours for red-green dichromats; simulated dichromatic (unspecified) view of original image; Daltonize-recoloured original image; simulated dichromatic (unspecified) view of recoloured image. Images from `http://www.vischeck.com/daltonize/`

Image Search

Image search involves searching for images that are similar to some query image. Work in this field has identified that individuals with CVD may have different expectations for image search results [7, 86], because the similarity of colour between images will be different for CVD users than non-CVD users. Similar findings have been extended to videos as well [88].

To address this problem, image search results can be recoloured to improve their accessibility for CVD users. Kovalev and Petrou [87] proposed a method for constructing a recolour map (Step 2 from Algorithm 1) from 256 'standard' colours to 256 colours differentiable for individuals with CVD that is weighted according to how frequently colours were adjacent in an image database of 12,000 images. One map is generated for each form of dichromatism (protanopia, deuteranopia, tritanopia), and once generated can be used to recolour any image very quickly. To determine each mapping, a randomized-greedy optimization strategy was used, taking between three and five hours to converge. The mapping was optimized to maintain as much as possible the perceptual differences between non-CVD colours in the resulting CVD recolour map colours (e.g., if red and yellow were mapped to blue and grey, the non-CVD perceptual difference between red and yellow was also the perceptual difference between the blue and grey).

This algorithm for adapting the results of image search uses a fixed colour map from original colours to replacement colours. This map does not take individual image characteristics

into consideration, but is calculated over a large set of exemplar images. By considering individual image characteristics, a more differentiable recolouring may be available. By using image-dependent manipulations of the colours in an image, Wang et al. [144, 145] are able to provide CVD-accessible image search that customizes the colours in each image to take advantage of the colour perception abilities of the user. Sample recolouring results are shown in Figure 2.13.

Figure 2.13: Image search recolouring results using the tool developed by Wang et al. [144, 145]. From left to right: original image; protanopic simulated view of original image; recoloured original image. Images copied from [145].

General Documents

SmartColor [141] is a recolouring tool that uses colour-relationship constraints supplied by the author of a document that uses colour, such as a website, visualization, chart or text document. These relationships are used to construct the following constraints over the final recoloured document:

- *Colour contrasts*: the colour perceptual difference between adjacent items should be maintained.
- *Colours as shared properties*: related elements with different locations should share the same colour.
- *Distinguishability*: an element that is coloured to make it stand out should keep this property.
- *Natural colouring*: some elements should not be recoloured (e.g., if textual colour references are made).

The dichromatic set of colours (one for each type of dichromacy) is searched for a set of colours that satisfy these constraints using simulated annealing [80]. The results of this approach were evaluated using visual inspection of a recoloured colour chart and an Ishihara plate [71], which show that the recolouring approach produces a recolouring consistent with the author's desires. Regarding timing, no empirical evidence is supplied, but the authors suggest that the approach is slow, and requires improvement. Sample recolouring results are shown in Figure 2.14.

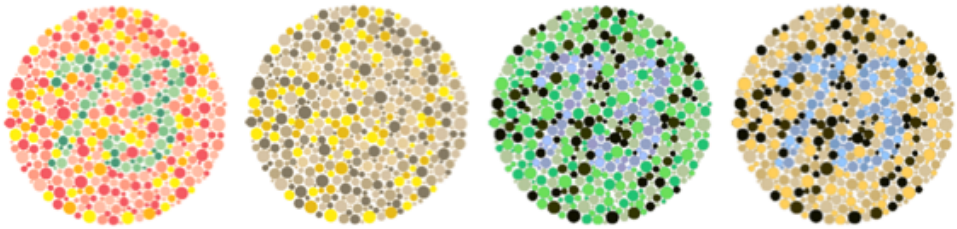


Figure 2.14: Demonstration of SmartColor recolouring [141]. From left to right: original image (Ishihara test plate); protanopic simulated view of original image; SmartColor recoloured version of the original image; protanopic simulated view of the recoloured version of the original image. Images copied from [141].

World Wide Web Content

Iaccarino et al. [67, 68] propose the use of edge services [13] to adapt webpage background colours, text colours, and images in real time. *Edge services* distribute application servers throughout the Internet, thereby redistributing the processing load of a single central server to multiple servers across the network, resulting in a reduced processing load per server. With these reduced loads, each server can perform more work per request, and Iaccarino et al. propose modifying each website requested from a distributed server to improve its colour differentiability for users with CVD. The algorithm they present uses simple RGB comparisons to count red pixels and green pixels to identify whether the image is predominantly red or green. The dominant colour is then shifted in hue (-108° in HSV space), has its saturation decreased by 10% and its lightness increased by 25%. The non-dominant colour (either red or green) has its saturation increased by 10% and its lightness decreased by 10%. Sample recolouring results for this approach are shown in Figure 2.15.

More recently, Mareuţă, et al. [102] proposed a recolouring tool that performs an optim-

Figure 2.15: Recolouring an image of a painting using Iaccarino et al.'s edge-service tool [67]. From left to right: original image; protanopic or deuteranopic (unspecified) simulated view of original image; recoloured version of original image; protanopic or deuteranopic (unspecified) simulated view of the image after recolouring. Images copied from [67].

ization-based recolouring that considers the relative proportion of each colour in the website, as well as the physical proximity of colours on the website (especially foreground text colours and background colours). The authors use these constraints to construct a mass-spring system which contains the set of replacement colours when it stabilizes. The mass-spring system is designed to maintain the original colours as much as possible, as well as attempt to restore colour differentiability for people with CVD. An automated analysis of 87 websites was performed to evaluate this recolouring tool, and showed that the tool helps restore colour differentiability. No user study was performed, and no timing results given.

Red-Green Dichromat Recolouring

Many recolouring tools require the user's type of dichromacy, but as identified in Chapter 1, most people with CVD do not know the specific type of CVD they have. To address this, Nakauchi and Onouchi [109] present another recolouring algorithm that is optimized to produce a single recolouring that is differentiable for individuals with protanopia, for individuals with deuteranopia as well as for individuals with typical colour vision. To achieve this, the authors propose an optimization evaluation function that combines two elements, one to minimize the overall colour changes (to maintain naturalness – see Section 2.6.3, below), and another that minimizes the differences between the perceptual difference between all colours for the dichromatic view and the perceptual difference between all colours for the non-CVD viewer. The dichromatic portion of this evaluation function averages the minimization function for protanopia and deuteranopia, thereby finding a solution for both types of red-green dichromacy.

To quantize the image, Nakauchi and Onouchi [109] use a clustering approach, and deal with clusters of colours, rather than finding key colours. This reduces the computational demands of the optimization procedure as well as makes the final step of Algorithm 1 much easier. This clustering approach was used later by Huang, et. al [62] (described below). Sample results are shown in Figure 2.16.

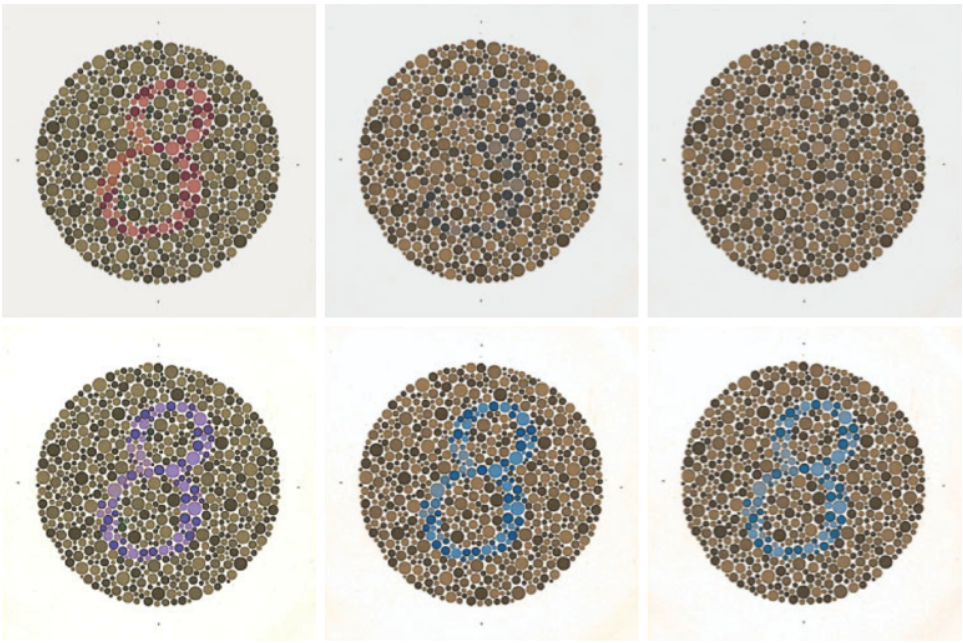


Figure 2.16: Demonstration of Nakauchi et al.'s [109] recolouring approach. Top, left to right: original Ishihara plate; protanopic view of Ishihara plate; deuteranopic view of Ishihara plate. Bottom, left to right: recoloured original Ishihara plate; protanopic view of recoloured plate; deuteranope view of recoloured plate. Note that the identically-recoloured plate is differentiable to both types of red-green dichromacy. Images copied from [109].

User Interfaces

Ichikawa et al. [69, 70] first proposed the use of genetic algorithms for solving the optimization problem often used to achieve step 2 of Algorithm 1. This approach was extended in Troiano et al.'s work [134] to more formally address the aesthetic desires behind palette colour choices in user interfaces. To achieve this, the authors create an objective function based on minimizing global colour changes (to maintain naturalness) and giving the user with CVD the same colour differentiability experience of individuals with typical colour vision (to en-

hance CVD colour contrast). They use a simple genetic algorithm to produce the recolouring results shown in Figure 2.17.

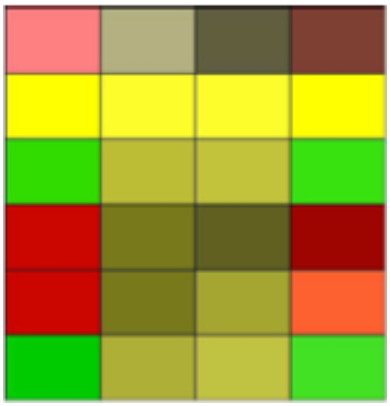


Figure 2.17: Palette recolouring results from Troiano et al. [134]. Palettes organized by columns: leftmost column is the original colour palette; rightmost column is the recoloured palette; second from left is the deuteranopic-simulated view of the original column; second from right is the deuteranopic-simulated view of the recoloured palette. Image copied from [134].

In later work [12, 133], these same authors present the use of interactive genetic algorithms to involve the designer in the colour palette choice for an interface. Intermittently during the genetic algorithm process, the designer is asked to rank five palettes in order of subjective preference. This information is used to steer the genetic algorithm as it attempts to find colour palettes that maximize differentiability for users with CVD while maintaining the aesthetic appeal of a user interface. Results for recolouring using this approach are shown in Figure 2.18.

Figure 2.18: Top row: original six-colour palette from Birtolo et al. [12]. Bottom row: the original palette after recolouring using their recolouring tool based on an interactive genetic algorithm. Image copied from [12].

Video

As part of the earlier research on accessible image search, it was identified that individuals with CVD will have difficulties with video as well as still images [88]. To address this difficulty, Huang et al. [61] developed a recolouring tool designed to improve the accessibility of video for individuals with CVD. Besides the raw processing demands of recolouring thousands of frames of a video, the main concern for video recolouring is temporal consistency of the recolouring. If a standard image recolouring tool is applied to each frame of a video, the recolour mapping may change from frame to frame, leading to objects flickering between different colours. Previous work addressing this [93] achieved the speed necessary to perform real-time recolouring of video (see Section 2.6.4), but was subject to long-term variations in the recolour mapping. This was because the recolour mapping was generated for each shot (sequence of related frames), but was allowed to change between successive shots (albeit smoothly). To address this, Huang et al. generated a recolour mapping for the entire video by distilling key colours from the entire set of frames [61]. This allows a single recolour mapping to be generated for the entire video, thereby ensuring temporarily-consistent recolouring. This method also utilizes the key-colour-priority approach presented in [60] to generate a high-quality, naturalness-preserving recolouring. They evaluated their approach against Rasche's recolouring tool [117] and Liu's recolouring tool [93] frame-by-frame, and found that their approach produced a consistent recolouring for the entire video and was ranked higher in a subjective user study with 11 participants. A summary of the comparative results with this technique is shown in 2.19.

Interactive Visualizations

Earlier work on recolouring interactive information visualizations [98] (see Section 2.6.4) addressed how to provide colour schemes for these visualizations that were accessible for people with CVD. One issue remaining with this earlier approach was the problems arising from the blending functions used to combine two overlaid patches of translucent data. Typical blending functions assume non-CVD colour perception [115], so these functions result in blended colours that may cause difficulties for users with dichromacy. To address this, a

Figure 2.19: Huang et al.'s [61] approach for temporally-consistent recolouring of videos. The numbers at the top are frame numbers for successive frames of a single video. Rows from top to bottom: original video frames; protanopic-simulated view of original video frames; video frames recoloured using Rasche et al.'s method [117]; video frames recoloured using Liu et al.'s method [93]; video frames recoloured using the method proposed by Huang, et al. [61]. All images copied from [61].

new colour blending function was proposed that blends colours so that the resulting colour is limited to the colour palette perceptible to a dichromatic user [19]. Coupling this with restrictions regarding the colours used for datapoints, they are able to produce more useable transparent regions as shown in Figure 2.20.

Real World Recolouring

A recent innovation in the application of recolouring tools has been using them to recolour objects in the real world. Two projector-based systems have been proposed to date [1, 154] and one mobile-phone system [113].

The projector-based systems utilize a camera-projector system to detect problems in the environment and to project relevant colour information into the environment. Amano et al. propose the use of a camera-projector system to recolour the user's surroundings using Jefferson's real-time interactive recolouring tool [75] (see Section 2.6.3). Results are shown

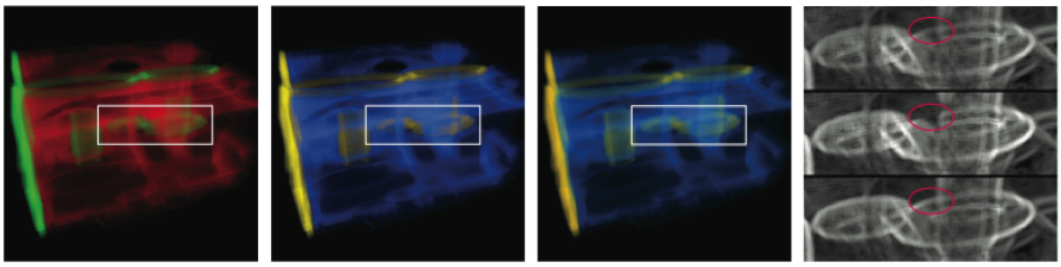


Figure 2.20: Direct volume rendering for users with CVD. From left to right: original visualization; visualization recoloured using Machado's method [98]; visualization recoloured using Chen's method [19]; magnification of rectangular region (top: original, middle: Machado, bottom: Chen) showing improved recolouring of Chen's method. Images copied from [19].

in Figure 2.21.

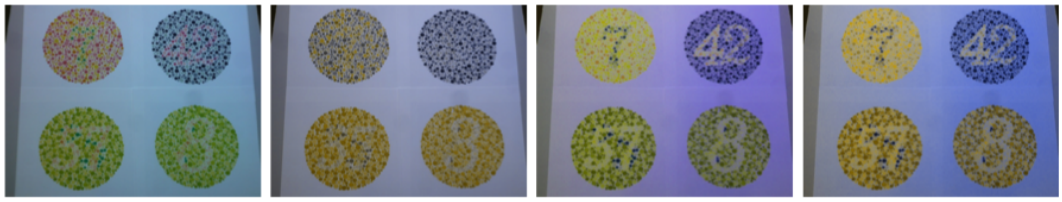


Figure 2.21: First projector-based recolouring tool [1]. From left to right: original image; protanopic simulated view of original image; recoloured original image; protanopic view of recoloured image. Images copied from [1].

Extending this earlier projector-based recolouring tool, Yamashita et al. [154] propose and implement the following three forms of projector feedback:

1. Drawing boundary lines between objects whose colours are not differentiable.

2. Drawing colour names on objects whose colour may be difficult for people with CVD to discern.

3. Projecting replacement colours onto objects in the environment to provide a new colour perception experience for the user.

Figure 2.22 shows each of these three projections at work. It should be noted that the colour choice algorithm is very simple, essentially using a red projection when the object under the projection is blue and a blue projection when the object under the projection is not blue. No evaluation of the system is provided, but its operation speed of four frames per second is given.

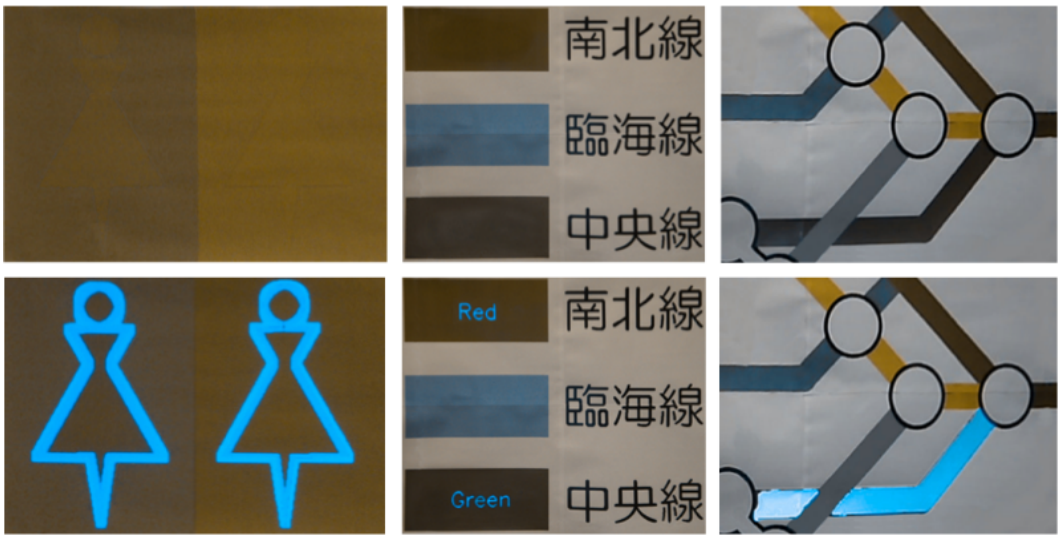


Figure 2.22: Projector-based recolouring with Yamashita et al.'s recolouring tool [154]. Top row contains three images that will cause problems for deuteranopic users: two red-green contrast signs; a chart or map legend; and a subway map. Bottom row shows projected overlays on each image type: boundary lines; colour names; replacement colour. Images copied from [154].

Lastly, recolouring tools have been implemented on mobile phones to provide a less-intrusive form of environmental recolouring [113]. To implement this, a region-growing method is used to identify regions with similar colours. These region colours are analyzed to see if any colours lie along the same confusion line. Problem colours are recoloured using the hue-rotation approach proposed in [63]. Sample iPhone recolouring images are shown in Figure 2.23.

2.6.3 Preserving Naturalness During Recolouring

Some approaches to recolouring can produce recoloured images that use dramatically different colours than the original image. Even though colour discriminability may be increased for someone with dichromacy, the usefulness of the image may be reduced because little of the original or natural colour experience remains. To address this, recolouring tools targeted at preserving the naturalness of the original image have been developed.

Typically, naturalness is employed to minimize the perceived difference between the original image and the recoloured image for the individual with inherited CVD. According to Figure 2.7, every colour that lies along a single confusion line is perceived as a single colour

Figure 2.23: Mobile-phone recolouring tool [113]. Left to right: original image of an Ishihara plate; protanopic simulated view of original image; protanopic simulated view of recoloured image. Images copied from [113].

by someone with dichromacy. As a result, colours can be arbitrarily moved along confusion lines during recolouring without influencing the perceived difference between the original and recoloured images for users with dichromacy. Some recolouring tools take advantage of this property to reduce the search space for the set of replacement colours by only considering colours that are perceived identically by people with and without CVD (the white lines in Figure 2.7).

An implication of this approach is that the perceived difference between the original image and the recoloured image can be quite large for an observer with typical colour vision. To address this, some recolouring tools do not modify every colour in an image, but only colours identified as problem colours (colours that are indistinguishable for people with CVD, but are distinguishable for users with typical colour vision). These recolouring tools aim to make only the problem colours more differentiable for people with CVD, leaving the remaining colours unchanged. This still results in naturalness preservation for users with CVD (because a majority of the colours are unchanged), but also helps preserve the original colours for people with typical colour vision. This second technique is employed in the recolouring tool presented in this dissertation (Chapter 7).

When recolouring tools use optimization steps similar to Ichikawa's proposed use of genetic algorithms to optimize a recolouring [69], the optimization can be steered to maximize

retention of original colours. This is usually achieved by adding a constraint to the objective function that penalizes moving colours away from their original colour coordinates.

Rasche et al.'s first recolouring tool did not preserve naturalness during recolouring (see Figure 2.31). To address this, the constraints for the optimization step of the recolouring algorithm were refined to include a constraint that restricted the output luminance to be equal with the input luminance [117]. As a result, the luminance of the image was maintained, preserving naturalness to a certain degree. This second approach uses multi-dimensional scaling to collapse the full-dimensional colour space of the original image to the reduced-dimensional colour space of dichromacy. Sample recolouring results are shown in Figure 2.24.

Figure 2.24: Demonstration of Rasche et al.'s second recolourer [117]. Left is original image containing red-green contrasts. Center is the original image recoloured for deuteranopic viewers. Right is the deuteranopic simulated view of the recoloured version of the original image. Note that luminance is maintained. Images copied from [117].

Jefferson and Harvey [73, 74] proposed the first recolouring tool based upon the World Wide Web Consortium's (W3C) Web Content and Accessibility Guidelines (WCAG) [48]. These guidelines are used to construct constraints designed to maintain sufficient luminance and colour contrast so that a document's colours are useable for any individual (CVD or not). These constraints are modelled simply as luminance difference (Y component difference in YIQ colour space) and colourimetric difference (Manhattan distance in RGB colour space). Jefferson and Harvey also add two more constraints to prevent replacement colours from being out of gamut of the display device and to penalize the replacement of colours that are deemed too vital to replace (white, black and greys). These constraints are expressed as a cost function which is minimized using the Matlab function *fminunc*. This function takes a cost function and an initial value, and determines the local minima for that function, starting

from the initial value. Multiple random starting points are executed, and the lowest-cost solution is selected. By maintaining these vital colours, naturalness is preserved somewhat, but not as effectively as Rasche's approach described above.

To select the key colours for the input image (step 1, Algorithm 1), Jefferson and Harvey use an RGB equal-volume binning of the colours and use the mean of each bin to represent its member colours, similar to [116]. One difference however, is that Jefferson and Harvey produce a similar binning for the dichromatic simulation of the image and only recolour those bins which are not in the dichromatic binning. To perform the interpolation step of the basic recolouring algorithm (Step 3, Algorithm 1), Jefferson and Harvey use Shepard's method [123], but leave it to the reader to explore the related literature. Sample recolouring results are shown in Figure 2.25.

Figure 2.25: Demonstration of Jefferson and Harvey's recolouring tool [74]. From left to right: original image; tritanopic simulated view of original image; recoloured version of the original image; tritanopic simulated view of the recoloured version of the original image. Images copied from [74].

Maintaining luminance [117] and essential colours [74] are important parts of preserving naturalness, but can be extended further. Huang et al. [63] were the first to propose a recolouring approach that attempts to make an image more differentiable for individuals with CVD, but also attempts to maximize naturalness as much as possible, so the image retains realism for the non-CVD viewer. To maintain naturalness, the authors propose three requirements for the recolouring:

1. Replacement colours have identical luminance as the original image representative colours.

2. Colours with the same hue in the original image all have the same hue in the recoloured image.

3. Replacement colours have the same chroma (colourfulness) as the original image colours.

To achieve recolouring while maintaining these three rules, Huang et al. propose the use of a rotation around the achromatic axis within the *a-b* plane of CIE L*a*b* colour space [130]. Any *a-b* plane is isoluminant (because L^* encodes luminance and is invariant for an *a-b* plane, so this meets requirement 1, above. Chroma is defined by distance from the achromatic axis in the a-b plane, so rotating around the achromatic axis will maintain chroma, meeting requirement 3. Hue is defined by angle from the *a axis* of the *a-b* plane, so rotating around the achromatic axis will move all colours of one hue to another hue (while maintaining their chroma), meeting requirement 2 from above. This rotation approach is illustrated in Figure 2.26.

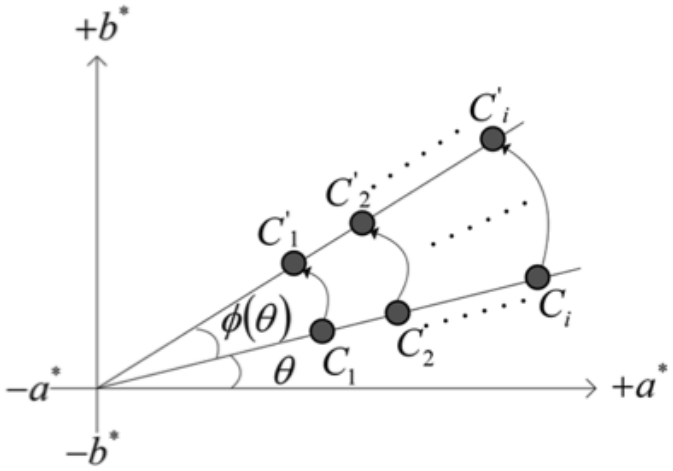


Figure 2.26: Example rotation of colours C_1, C_2, ... C_i about the achromatic axis of L*a*b* colour space by angle $\phi(\theta)$. Achromatic axis is coming out of the screen or paper. Image copied from [63].

To achieve a recolouring that maximizes differentiability and naturalness, the authors develop two criteria, one for maximizing colour differentiability and the other for naturalness preserving. These criteria are used to steer an optimization (Fletcher-Reeves conjugate-gradient method) to find the optimal recolouring. A variable is also introduced into the objective function to allow the user to select between emphasizing colour differentiability or maintaining naturalness. A user study measuring the naturalness and differentiability of recoloured images ranked this method above the results produced by the Daltonize [30] and Rasche's [116, 117] methods. A sample recolouring for protanopic individuals is shown in Figure 2.27.

Figure 2.27: Recolouring to improve dichromatic differentiability while maintaining naturalness from [63]. From left to right: original image; protanopic simulated view of original image; recoloured original image; protanopic view of recoloured image. Images copied from [63].

These same authors propose yet another recolouring tool in [62] that makes improvements to the first and final steps of Algorithm 1. To improve the selection of key colours from an image (step 1), the authors propose representing image colours with a Gaussian Mixture Model [101], in which the distribution of colours is assumed to be well-approximated by a collection of Gaussian distributions. The mean of each distribution is considered a key colour. This representation also helps with step 3 of Algorithm 1 because it automatically provides the distribution of all colours that are represented by a key colour, so they can be easily aligned to the key colour after it is recoloured. To achieve step 2 of Algorithm 1, the authors once again use a rotation about the achromatic axis as they did in [63] and [64], but this time use the LCH (Luminance-Chroma-Hue) representation of L*a*b* colour space. By maintaining the distribution of colours throughout the recolouring process, the variance of the original image can be maintained, thereby improving naturalness preservation. Example recolouring results are shown in Figure 2.28.

Figure 2.28: Recolouring by hue transformation using Gaussian Mixture Model [62]. Left: original image; center: deuteranopic simulated view of original image; right: recoloured original image. Images copied from [62].

Similar to Huang et al.'s work [63], Kuhn et al. [89, 90] developed a recolouring tool

that sought to preserve the naturalness of the image as much as possible while reducing the time required to perform recolouring. Kuhn et al. utilize a GPU-optimized mass-spring system to rearrange representative colours along the line representing the intersection of each representative colour's luminance plane, and the set of colours uniquely discernible for someone with dichromacy (approximated using a plane). The solver is constrained to try to maximize colour discriminability for users with dichromacy while reducing unnecessary modification of colours to maintain naturalness. GPU-parallelization of their solution allows about a 2000x speedup over Rasche's solution presented in [116]. They evaluated their solution using fourteen participants with CVD, and found that their recolouring technique was preferred over Rasche's in terms of naturalness preservation, contrast enhancement, and personal preference. Sample recoloured images are shown in Figure 2.29.

Figure 2.29: Image recolouring using Kuhn et al.'s [90] recolouring tool. From left to right: original image; protanope simulation of original image; original image recoloured with Rasche; original image recoloured with Kuhn. Images copied from [90].

In [60], Huang et al. present the use of 'key colour priority' as a way to make recolouring efficient (in terms of computation time) and naturalness preserving. To accomplish this, key colours are first selected by a two-stage process: 256-colour RGB quantization [55] followed by clustering with affinity propagation [37] to produce clusters of related colours. The center of each cluster is then used as a key colour. The key colours are then ordered according to the perceptible difference between non-CVD perception of the colours and dichromatic perception of the colours. This is the key colour priority referred to above; key colours are given a priority that correlates with how perceptibly different they are for people with dichromacy. Colours that are identically perceived in both dichromacy and typical colour vision (e.g., greyscale colours) are processed first by copying them as they are to the recoloured image. Colours that are almost identically perceived (e.g., blues and yellows for deuteranopia), are processed next and may experience minor modification before moving into the recoloured image. Finally, the colours with the largest differential perception are processed last. This priority-based

approach helps maximize the preservation of naturalness between the original and recoloured images. To paint all colours to the recoloured image, Gaussian Mixed Models [101] are used just as they are in [62], presented above. The authors compared their method with Rasche et al.'s [117] and found by visual inspection that the results from their recolouring approach were better in terms of naturalness preserving. Sample recolouring results are shown in Figure 2.30.

Figure 2.30: Huang et al.'s [60] recolouring approach. Top and bottom rows, left to right: original image; simulated deuteranopic view of original image; eecoloured with Rache's recolourer [117]; protanopic simulated view of original image recoloured with Huang's recolourer [60]. Images copied from [60].

Preserving naturalness through recolouring is inherently difficult, because increasing colour differentiability for people with CVD necessitates changing the original colours to colours that are more differentiable. This results in two typically competing constraints influencing the recolouring process. Although the results presented here show that naturalness can be well preserved (at least in the images presented), it is almost always based on an optimization procedure, which can induce unnecessary computational overhead. Next, methods introduced for reducing the time to recolour are introduced.

2.6.4 Recolouring Speed

Many recolouring approaches presented so far require considerable computational resources to perform. As recolouring is typically performed in a user-demanded fashion (recolouring occurs when the user requests it), lengthy recolouring can lead to the user abandoning the

technology as it imposes too much of a time penalty to use [27]. To address this, many techniques have been proposed that reduce the time to perform recolouring. This has resulted in substantial time savings, such that recolouring can now be performed in real time (e.g., on interactive visualizations [98]).

Building on the early work of Ichikawa [69, 70], Rasche et al. proposed another optimization-based recolouring tool based on dimensionality-reduction work moving colour images to greyscale [116]. As optimization-based recolouring algorithms are generally computationally expensive, the authors use a simple equal-volume binning of the image colours and use the mean of each bin to represent its member colours. This reduces the number of colours in each image to about 1000 (Step 1, Algorithm 1). To find the recolour map (Step 2, Algorithm 1), the Fletcher-Reeves conjugate-gradient method [125] is used to optimize a recolour mapping to maintain the internal differences between the original colours and the replacement colours. The initial mapping optimized by Fletcher-Reeves maps the first two components of a principal component analysis of the original image to the a and b dimensions of CIELab colour space. No details for applying this recolour mapping to repaint the image are provided (step 3, Algorithm 1), but this approach takes less than ten seconds to execute on images of unknown size. Sample recolouring results are shown in Figure 2.31.

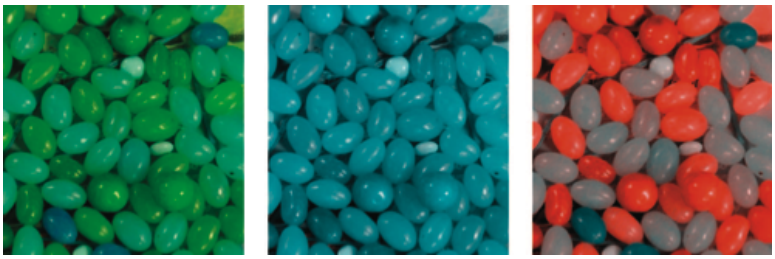


Figure 2.31: Demonstration of Rasche et al.'s first recolourer [116]. Left is original image containing blue-yellow contrasts. Center is the tritanopic simulated view of the original image. Right is the tritanopic simulated view of the recoloured version of the original image. Images copied from [116].

Breaking away from the algorithm presented in Algorithm 1, Jefferson and Harvey produced a real-time dichromatic recolouring tool that relies on the user to find a useful recolouring, rather than an optimization algorithm [73, 75]. An image is captured and converted to LMS coordinates (see Section 2.1.2). The user supplies his/her form of CVD (protanopia, deuteranopia, or tritanopia) and the colour information for the missing cone is transferred to

the remaining two cones. How this colour information is transferred is governed by particular values that are derived from the value of a single slider that the user adjusts. As the user adjusts this slider, the values are manipulated and the resulting recolouring is modified. In this manner, the user can explore the available recolouring options and select the best one. Jefferson and Harvey evaluated this approach by running 25 participants through 45 Ishihara plates [71]; 15 original, 15 with simulated protanopic vision, and 15 after recolouring. It was found that this recolouring tool significantly reduced error scores from the simulated condition to the recoloured condition. Sample recolouring results are shown in Figure 2.32.

Figure 2.32: Example recolouring from Jefferson and Harvey's real-time recolouring tool [75]. Left: original image; center: protanopic simulated view of original image; right: protanopic simulated view of recoloured image. Images copied from [75].

In [28], Deng et al. explore the generation of a fixed recolour mapping that compressively maps the full gamut of colours to the restricted set of colours perceived by individuals with dichromacy. The full gamut of colours is subsampled to produce a reduced-size representative colour set. This smaller set is mapped to the set of colours that dichromats can distinguish using an optimization based on similarity matrices [52]. This fixed map is generated once and then utilized to perform real-time recolouring by first reducing the image to only use the compressed full-gamut colours (akin to Step 1 in Algorithm 1), then using the mapping as a look-up-table to recolour this compressed colour image to produce the final recoloured image. By utilizing a single look-up-table for every recolouring, the temporal consistency concerns of video recolouring ([93]) are addressed, but because every colour in the source image is mapped to a new colour, images can be quite dramatically altered through recolouring, as shown in Figure 2.33.

Building on their previous recolouring tool, Huang et al. extended the recolouring algorithm to run in near real time (0.5 seconds to recolour a 200x200pixel image) by performing

Figure 2.33: Illustration of recolouring process proposed by Deng et al. [28]. From left to right: original image; compressed colour original image; protanopic simulated view of compressed image; recoloured compressed image (also serves as a simulation because it only uses colours that are identically-perceived by non-CVD individuals and individuals with protanopia). Images copied from [28].

an image-dependent transformation of hue in HSV colour space [64]. This transformation is based on generalized histogram equalization, and seeks to distribute the hues of an image uniformly around the circle of hues with similar saturation and lightness to maximize hue contrast. The authors compare the recolouring results of this technique with Rasche's [116] and Jefferson's [74] recolouring techniques to show that the recolouring produced by this technique is more natural and consistent, but no quantitative results are presented. Example results from this recolouring tool are shown in Figure 2.34.

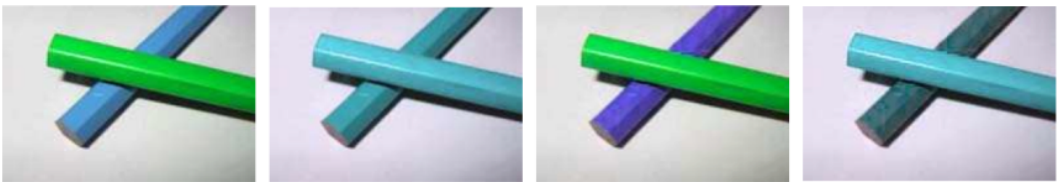


Figure 2.34: Recolouring by hue transformation using generalized histogram equalization from [64]. From left to right: original image with colours that cause problems for individuals with tritanopia; tritanopic simulated view of original image; recoloured original image; tritanopic view of recoloured image. Images copied from [64].

Besides optimization steps, another source of computational overhead is moving between multiple colour space representations of the image. To eliminate this overhead, recolouring can be performed using simple RGB manipulations [4]. This approach was also used in Iaccarino's edge-services work for websites [67, 68]. To accomplish this recolouring, all RGB colours are mapped to 2D planes in RGB colour space that approximate the planes of colour used in simulations by Brettel et al. [16]. Image colours are then mapped to these planes. In comparison to the author's previous recolouring tool [3], this new approach is shown to

execute more quickly and also to produce more natural output, as shown in Figure 2.35.

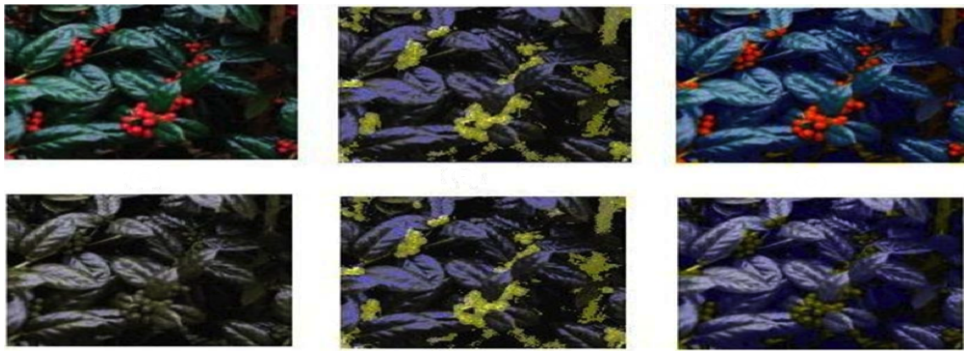


Figure 2.35: Original and improved recolouring results presented by Bao et al. [4]. Top row, left to right: original image; original image recoloured with original recolourer; original image recoloured with improved recolouring tool. Bottom row contains simulated protanopic views of the corresponding image in the top row. Images copied from [4].

Another approach not based on optimization algorithms for Step 2 of Algorithm 1 is presented by Wong and Bishop [152]. Their approach uses the HSV representation of RGB colour space. To recolour, a non-linear rotation of individual colours around the hue circle (maintaining HSV saturation and value) is performed to increase the angular hue distance between some colours while decreasing the angular hue distance between others. For protanopic and deuteranopic users, the red-yellow-green end of the spectrum is stretched, and the blue end of the spectrum is compressed. The amount colours are compressed or stretched rotationally depends on the colour qualities of the input image. No timing results are given for this approach, but as it contains no optimization step and uses simple rotations in a colour space quickly converted to and from RGB, it should be relatively fast. Sample images are shown in Figure 2.36.

Figure 2.36: Wong and Bishop's recolouring approach presented in [152]. Left: original image; center: simulated deuteranopic view of original image; right: simulated deuteranopic view of recoloured image. Images copied from [152].

Early recolouring work looked to produce a single recolour map from all possible colours to the set of colours that dichromats can distinguish. By having a single map, any image can be recoloured very quickly by simply using the map as a look-up-table to look up the replacement colour for a given pixel [28, 87]. The first to apply a self-organizing map (SOM) to generating such a recolour map are Ma et al. [95]. To perform a recolouring, the authors first randomly subsample the original image. This acts as a simple Step 1 of Algorithm 1. The sampled RGB colours are input into an SOM that reduces the three-dimensional RGB representation of the colours to a two-dimensional planar representation of the input colour set. The resulting plane of colours is then mapped to a replacement set of colours drawn from a colour plane that contains colours along a black-white axis and a blue-yellow axis. To perform Step 3 of Algorithm 1, each original image pixel is examined to determine which of the sampled colours is closest ('nearness' is not defined – but probably is Manhattan or Euclidean distance in RGB space), and the colour that this sample colour maps to is used as the new pixel colour. This leads to recoloured images that are limited to black, white, yellow, and blue colours that are subject to banding because the number of replacement colours is dependent on the number of subsampled colours, which is kept small to reduce computational cost. Example images are shown in Figure 2.37.

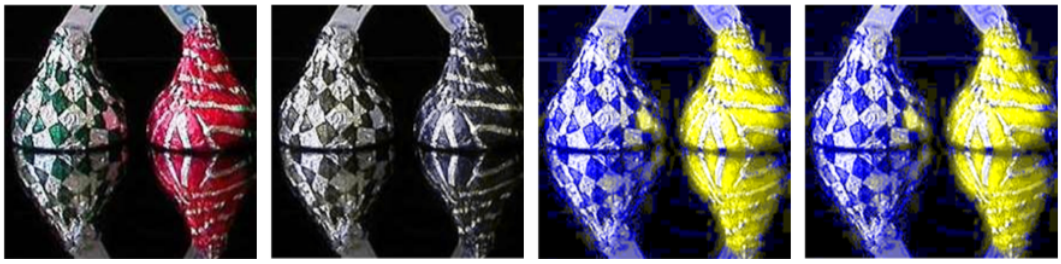


Figure 2.37: Ma et al.'s recolouring approach [95]. Left to right: original image; simulated protanopic view of original image; recoloured; protanopic simulated view of recoloured image. Images copied from [95].

Although some of the solutions presented so far have achieved fast computation of recoloured images, Liu et al. [93] propose the first method intended for use on video streams in real time. Their approach to recolouring uses rotations within L*a*b* colour space, much like Huang et al. [63]. The first rotation maps colour variations along the a^* axis onto variations along the b^* axis because protanopic and deuteranopic colour vision can be approximated as

a reduced ability to perceive variations along the a^* axis. The second rotation collapses all colours to their a^*-b^*-plane coordinates and rotates these colours so that the first component of a Principal Component Analysis with the axis of maximum differentiability (which doesn't align perfectly with the L*a*b* b^* axis for each type of dichromatism). As these series of rotations are not computationally intensive, a recoloured image can be produced very quickly. By reducing the total number of colours to 32,768 through quantization, the authors are able to apply the recolouring algorithm proposed here to real-time video. To maintain recolouring consistency (preventing sudden colour changes between frames), they propose using the same recolour mapping for all frames in a single shot, and smoothly changing the recolour mapping between shots. The authors evaluate their static image and video recolouring technique with a user study. Subjective results show their tool is superior to Huang et al.'s previous two methods ([63] and [64]). Figure 2.38 contains two sample images recoloured with this method along with sample recolouring results from ([63] and [64]).

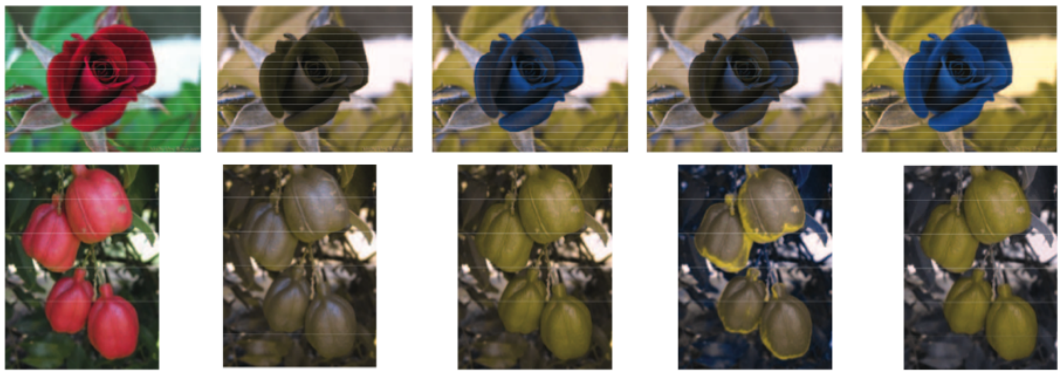


Figure 2.38: Demonstration of video-stream recolouring tool by Liu et al. [93]. Top row and bottom row left to right: original image; simulated deuteranopic view of original image; original image recoloured with Huang et al.'s [63] first approach; original image recoloured with Huang et al.'s [64] second approach; original image recoloured with Liu et al.'s [93] approach. Images copied from [93].

Improving real-time recolouring for information visualizations, Machado et al. [97, 98] also show they can attain temporally-consistent real-time recolouring of dynamic content (in an interactive visualization system) similar to Liu et al. [93]. Their recolouring tool maintains greyscale colours in the image and consistently enhances the colour contrast for dichromats. The authors optimized their technique to take advantage of recent GPU parallelization technology and the resulting optimized recolouring tool can recolour 800x800 pixel images in

under 0.03 seconds, allowing real-time recolouring to be performed. Similar to Kuhn, et al.'s recolouring tool [90], the system presented here is able to perform an 'enhanced contrast' recolouring too. The authors use visual comparison to evaluate their solution – results of which are shown in Figure 2.39.

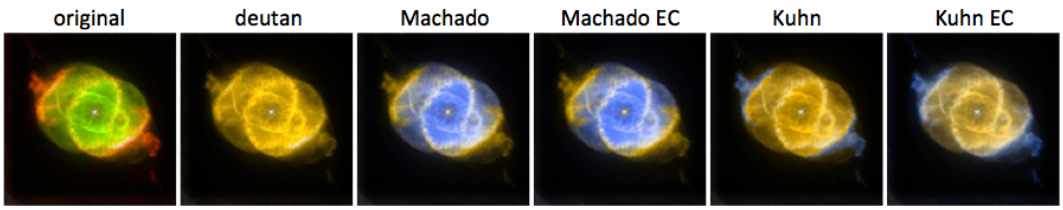


Figure 2.39: Demonstration of Machado's real-time temporally-consistent recolouring tool [98]. From left to right: original image; deuteranopic simulated view of original image; Machado's proposed recolouring; Machado's proposed contrast enhancing recolouring; Kuhn's static image recolouring [90]; Kuhn's enhanced contrast static image recolouring [90]. Images copied from [98].

Returning to images on the web, Ruminiski et al. [121] pick up where Ichikawa and Iaccarino left off with their work on recolouring web content ([67, 68, 69]). To extend this previous work, Ruminski et al. focus on attaining real-time performance of their recolouring tool to allow rapid recolouring of web content as it is accessed. To accomplish this, they propose two simple recolouring algorithms which each lead to a recolouring tool:

1. This algorithm works in RGB colour space only, and acts to amplify the 'redness' of red areas and suppress the 'greenness' of green areas, much like Iaccarino, et al. [67, 68].

2. This algorithm operates in HSV colour space by decreasing the saturation of red areas of the image and decreasing the brightness of green areas.

It is arguable that these simplistic approaches would not compete with the state-of-the-art (e.g., as demonstrated by Huang [60]), but the resulting recoloured images increase dichromatic contrast and maintain naturalness quite effectively (Figure 2.40). Due to the simplicity of these proposed recolouring tools, their execution time is quite low; under 500 ms for 1500x1500-pixel images.

Using simplified recolouring algorithms or precomputed recolour maps, recolouring can be made to be fast enough to operate in real time, allowing interactive recolouring tools. Current

Figure 2.40: Demonstration of Ruminski et al.'s two proposed recolouring algorithms [121]. Top row, left to right: original image, deuteranopic simulated views of Rasche's recolouring [117] and Kuhn's recolouring [90]. Bottom row, left to right: deuteranopic simulated views of the original image after Daltonize recolouring [30], Ruminski method one, and Ruminski method two. Images copied from [121].

approaches only address dichromatic users, but anomalous trichromacy is three times more common than dichromacy [20], yet is not as severe as dichromacy. As a result, the real-time recolouring tools presented above perform 'too much' recolouring. In the next section, an alternative to recolour mapping is presented as a solution for individuals with anomalous trichromacy.

2.6.5 Recolouring Approach

In general, two main approaches to recolouring have been proposed: *colour replacement* and *colour amplification*. Colour replacement involves completely changing the colours in an image to improve its colour differentiability; this solution is typically used when the user has dichromacy. Colour amplification involves increasing the saturation and/or luminance of colours in an image (but retaining the hue of the colours) to make the original colours more discernible; this solution is typically used when the user has anomalous trichromacy. The

other sections dealing with application (2.6.2), naturalness (2.6.3) and speed (2.6.4) deal with recolouring tools that perform colour replacement, so recolouring tools that perform colour amplification will be presented here.

By modifying the spectral frequency distribution entering the eye, different cones can be stimulated in varying amounts. As discussed in Section 2.2.1, individuals with anomalous trichromacy have all three cone types, but one of the cones (the anomalous one) has its peak sensitivity shifted toward the peak sensitivity of another cone. As a result, these two cones are subject to less differential stimulation than in typical colour vision, thereby reducing the ability to discriminate between colours that are identified via differential stimulation of these two cones.

By modifying the spectral frequency distribution before it enters the eye, the anomalous cone can be stimulated at a higher level than usual, allowing the individual with anomalous trichromacy to experience the typical colour vision. The first advances in this area used filters. Early on, approximate solutions were achieved using long-wavelength pass filters [58] to super-stimulate the long-wavelength sensitive cones in individuals with protanomalous CVD. Later, custom-constructed filters tuned to the phosphor emissions of the CRT display and the type and severity of anomalous trichromatism [85] were developed to achieve a more precise solution.

Although filters modify the spectral distribution in an appropriate fashion, filters operate by suppressing the intensity of the light for undesired parts of the spectrum. This results in an unnatural darkening of of the user's vision. In addition, the filters were often implemented as glasses, conflicting with existing prescription glasses. To address this, the light emitted by the display itself can be modified to compensate for anomalous trichromatism. To perform this modification, a measurement of the severity of anomalous trichromatism is needed. This is specified as the amount of peak sensitivity wavelength shift in the anomalous cone, measured in nanometers. With this information, the RGB values of an image can be modified to increase the isolated stimulation of the anomalous cone [110, 119, 128, 155, 156].

Building on their earlier proposals for colour amplification [128, 155], Yang et al. first evaluated their proposal using a subjective user response study in which participants ranked adapted images as 'better', 'equal', or 'worse' than the original image [119]. Utilizing a

between-participants design, it was found that the average number of participants who found that the adapted images were 'better' was greater than those who found it to be 'equal' or 'worse'. These experiments were repeated again and presented in [156], this time with a Z-test comparison showing that the number of participants who chose 'better' was significantly higher than the number of participants who chose 'equal' or 'worse'. These findings were reported once again in [110] which extends the colour amplification approaches discussed so far to incorporate user preference for colour temperature of the final image.

A necessary input to these adaptations is the type and severity of anomalous trichromacy to accommodate in the recolouring tool. The type of anomalous trichromacy is diagnosable using standard CVD screening tests (e.g., Ishihara plate test [71] or HRR pseudoisochromatic plates [50]), but to determine the severity of anomalous trichromacy, the Farnsworth-Munsell 100-Hue Test [31] is used in [157] and shown to work well as input to their colour amplification recolouring tool. These authors later present their own computer-based colour vision test (called a Computerized Hue Test - CHT) [158] and use this as the basis for a colour amplification recolouring tool [159].

Mochizuki et al. [104] identify the same limitation identified by Yang et al. [157] in that information regarding the severity of anomalous trichromatism is not easily obtained. To address this, Mochizuki utilizes measurements of discrimination ellipses [114] to approximate the spectral sensitivity shift of the anomalous cone. This is used to perform a partial translation of colours along confusion axes. Applying the translation so that colours move towards the two half-planes results in a simulation of anomalous trichromatism [16, 99]. Applying the inverse of the transformation moves colours away from these half-planes, thereby amplifying them, resulting in the over-stimulation of the anomalous cone. This over-stimulation allows an individual with anomalous trichromacy to experience typical colour perception, because his/her cones are differentially stimulated in a manner similar to how they are stimulated for someone with typical colour perception. Examples of simulation and recolouring are shown in Figure 2.41.

Mochizuki et al.'s early approach [104] is expanded in [106] by more clearly describing how measurements of a user's discrimination ellipse can be extended to perform full simulation within three dimensions (where each dimension is defined by a confusion line). This extension

Figure 2.41: Simulation and amplified recolouring produced by [104]. Left to right: original image; anomalous trichromacy (protanomalous) simulation; protanomalous amplification recolouring. Images copied from [104].

allows colour perception modelling of individuals who experience multiple types of CVD (e.g., an older protanomalous man experiencing tritan effects from yellowing of the lens with age [38]). An example of a colour-amplification recoloured image for one of Mochizuki et al.'s study participants is shown in Figure 2.42.

Figure 2.42: Simulation and amplified recolouring produced by [106]. Left to right: original image; anomalous trichromat participant colour vision simulation; colour amplification recolouring for this participant. Images copied from [106].

Mochizuki et al. further extended the approach described above by simplifying the transformations between the original colour space and the simulated and amplified versions [105]. Mochizuki et al. claim that these modifications improve the execution speed of the transformations, although no timing results are given. They also claim that the simplification allows a more-straightforward implementation of the transformation, but how this is achieved is unclear from the manuscript. Sample images are presented in Figure 2.43.

By amplifying the perception of particular hues, the colour experience of non-CVD individuals can be restored to individuals with anomalous trichromacy. In spite of the apparent value of this approach, two limitations restrict its effectiveness. First, the techniques to precisely measure the type and severity of anomalous trichromacy need to be improved to reduce

Figure 2.43: Simulation and amplified recolouring produced by [105]. Left to right: original image; simulated anomalous trichromat colour vision; colour amplification recolouring. Images copied from [105].

the time taken to perform this measurement. Second, the technique of amplifying colours is fundamentally limited by the gamut of the display device. Typically, realistic images contain colours that are near the outer limit of a display device. Amplifying these colours will often cause them to be shifted outside of the gamut. To address this, display hardware moves the colours back into the gamut, thereby nullifying the benefits of the amplification process. Some examples of this can be seen in the examples of Mochizuki's work above, e.g., Figure 2.41. Introducing larger gamut display devices does not address this problem either, as the gamut will eventually be bounded by the gamut of colours perceivable by the human visual system.

2.7 Conclusion

This chapter presented the foundational knowledge required to understand Situation-Specific Modelling (SSM), and Individualized Models of Colour Differentiation (ICD) as presented in this dissertation. In the next chapter, the SSM and ICD approaches will be explored and more formally defined.

Chapter 3

Situation-Specific Modelling and Individualized Models of Colour Differentiation

As discussed in Section 2.2, many internal and external factors influence an individual's ability to differentiate between colours. When these factors deviate from typical colour vision, inherited, acquired, situationally-induced, and simultaneous colour vision deficiencies (CVDs) result. CVD is common and results in reduced colour differentiation abilities.

To assist people with CVD, *recolouring tools* have been proposed that modify interface and image colours to make them more differentiable. Recolouring tools rely on models of colour differentiation to identify problem colours and to find suitably-differentiable replacement colours; however, existing models of colour differentiation do not represent the full range of differentiation abilities. This results in existing recolouring tools not helping most people with CVD. In this dissertation, I present Individualized Models of Colour Differentiation (ICDs) through Situation-Specific Modelling (SSM) as a solution. SSM-based ICDs capture the influence of the full range of internal and external factors that influence colour differentiation, thereby providing recolouring tools with a model that can represent the full range of colour differentiation abilities.

In this chapter, I present a more complete picture of Situation-Specific Modelling (SSM) and the use of SSM to develop Individualized Models of Colour Differentiation (ICDs). First, I present an explanation of why traditional modelling approaches are ill-suited to ICDs. Next, I give an overview of SSM and its application to ICDs. I then compare SSM to previously-published related approaches, and then present some details pertaining to the application of SSM to ICDs. I conclude this chapter with some design considerations for SSM-based ICDs, with pointers to the relevant chapters in the remainder of this dissertation.

3.1 Traditional Modelling Approach

Section 2.3 provided an overview of existing models of colour vision deficiency. In general, existing models of CVD assume that the only factor influencing an individual's ability to differentiate colours is the absence of one type of cone (for dichromacy) or a shift in peak sensitivity of one type of cone (anomalous trichromacy). However, many other factors influence human colour differentiation abilities, as described in Section 2.2.1, such as genetic variations in 'normal' human colour perception, acquired colour vision deficiency, and CVD induced by different environmental factors (situationally-induced CVD). Current models of CVD do not account for these additional factors.

To extend existing models of CVD to accommodate the wide variety of factors that influence colour differentiation, a model of each factor's influence on colour perception is necessary. These models would need to be able to take some input (e.g., lighting level for a light-level model) and provide some indication of how the entered value influences colour differentiation. Examples of models that would be necessary include the effects of light levels, light quality (i.e., white balance of ambient lighting), eye diseases, aging, neurological damage, coloured eyewear, and air pollution on the colour differentiation abilities of an individual. Some of the factors influence a user's colour differentiation abilities are illustrated in Figure 3.1.

In addition to modelling each individual factor, a method for combining each of these models is required to amalgamate each individual model into an aggregate model that describes the colour differentiation abilities of a particular individual in a particular situation. This amalgamation process would be complicated by interactions between the different factors, e.g., ambient lighting with different white balances may cancel-out or magnify colour differentiation difficulties caused by yellowing of the lens with age.

Even if each factor that influences colour differentiation was represented by a model, and a method for combining these models into an aggregate 'super-model' existed, two additional constraints prevent this from being a feasible approach. First, each model requires its own input, so mechanisms would need to be developed for collecting the input for each model as it changes from situation to situation and over time for each user. It is important to note

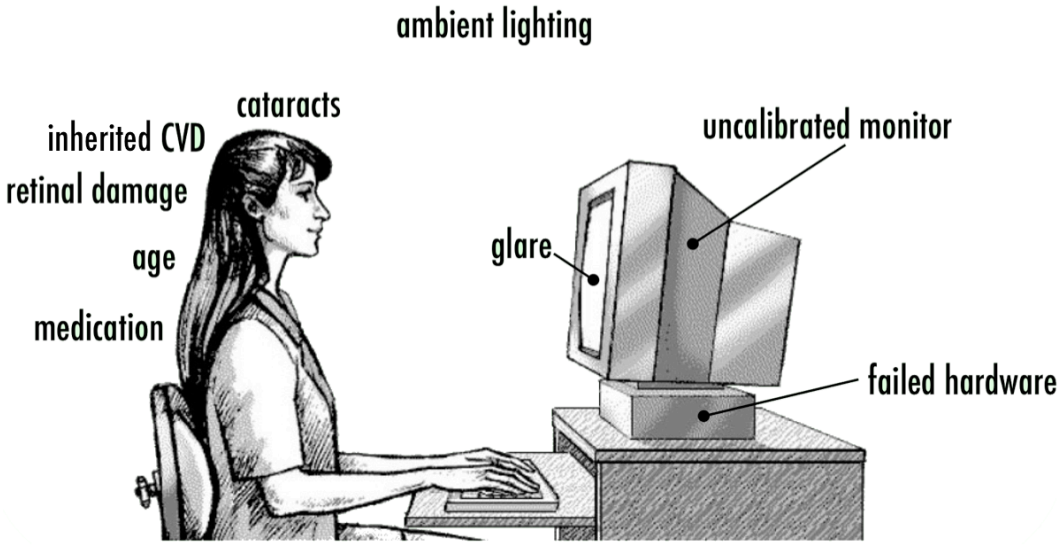


Figure 3.1: Some of the factors that influence a user's colour differentiation abilities. Factors can be internal (inherited or acquired) or external (situationally-induced) to the user.

that each model may require complex input (e.g., degree, type and location of neurological damage), confounding this input collection mechanism further. Second, there may remain unknown factors that influence colour differentiation that have yet to be discovered and modelled yet. For example, a new class of anomalous trichromacy called *extreme anomalous trichromacy* was only discovered in 2003 [76]. Another example is the lack of clear understanding of whether a colour is processed in a single location or is distributed throughout the visual cortex [45]

As a result of these major hurdles, this dissertation presents SSM-based ICDs as a practical alternative to extending traditional modelling approaches for developing models that represent the full range of factors that influence colour differentiation.

3.2 Situation-Specific Modelling (SSM)

A key observation regarding the wide variety of factors that influence a user's colour differentiation abilities is that everything that influences the colour perception of the user shown in Figure 3.1 are influencing her colour differentiation abilities as she uses her computer. The colours generated by the system will be influenced by any hardware or software malfunc-

tions, and improper calibration and ambient lighting will influence the colours as they are displayed on the screen. Any acquired or inherited CVD that the user has will influence the colour perception of the user. Finally, the user's ability to differentiate between colours will influence how he/she interacts with what is on the screen; whether he/she can read text, or use a visualization, for example.

At its foundation, Situation-Specific Modelling (SSM) is based on capturing the user's performance of some task *in situ* – within the user's current situation. In this dissertation, the process for measuring a user's performance is called an in-situ calibration, because it is through the measurement process taking place in situ that a general model of user performance can be calibrated to that particular user in his/her particular situation. In this dissertation, SSM is used to measure and model a user's colour differentiation abilities, but this is not its only application. For example, SSM could also be used to model a user's selective hearing loss (because of inherited, acquired, or situationally-induced causes), or to model inter-user differences in visual acuity.

As mentioned above, measurements from the in-situ calibration are used to calibrate a general model to a particular user in his/her particular situation. In general, the calibration data represents the user's performance on some sample of tasks and the modelling process takes these samples and extends them to develop a full picture of the user's abilities. More details about this process for colour differentiation are given below.

Situation-specific modelling therefore requires two main components: an in-situ calibration component that measures a user's abilities or performance, and a modelling component that uses measurements from the calibration to build a model that represents the user's abilities.

3.2.1 SSM Calibration Component

The in-situ calibration component of SSM presents particular stimuli to the user and records the user's response to those stimuli. Stimuli and responses are now explored in more depth.

Stimuli Presented to the User

To perform a calibration, SSM presents stimuli to the user and records his/her response to the stimuli. As current computer technology produces visual output through the display and auditory output through speakers, these are the principle stimuli presentation modalities for SSM. Examples of each are images, colour swatches, and moving icons for visual stimuli; and sound samples at particular frequencies or volumes.

When considering stimuli to be presented to the user, SSM can utilize *active* or *passive* stimuli. Active stimuli require an explicit calibration procedure in which the user or system initiates a calibration and the user knows they are performing the calibration. Passive stimuli allow calibration data to be collected without the user's knowledge. An example of active stimuli involves the calibration procedure used to align the input space on a table with the output display space for the table [91, 149], where the user knows that the calibration is taking place and so is actively involved in the calibration. An example of passive stimuli is the use of a game to gather calibration data – the user will be actively engaged with the game rather than the calibration task, so the calibration data can be gathered passively as a side-effect of the gameplay [35]

Responses Provided by the User

Users will typically provide input via the keyboard or mouse (or through touch on touch-sensitive displays). Despite the stimuli being limited to only visual and auditory modalities, the responses provided by the user can reflect a wide range of measurable variables. For example, showing an image on the display can be used to determine the user's visual acuity [6, 8], colour differentiation abilities, response time, face-recognition abilities, or whether the user is field-dependent or field-independent [46] Likewise, auditory stimuli can be used to determine selective hearing loss, response time, or spatial awareness.

The nature of the responses provided by the user can be divided into two main types: judgement tasks and performance tasks. *Judgement tasks* are tasks in which the user needs to mentally process the stimuli and provide a judgement about the stimuli. For example, to determine colour differentiation abilities, two colour swatches can be displayed on a display

and the user provides a judgement about whether he/she perceives the two colours as the same or different. In contrast, *performance tasks* allow the user to provide input to the calibration simply by performing a task; the user's success or failure in performing the task provides the information necessary for calibration. An example of a performance task for measuring colour differentiation is to embed a shape of one colour within a distractor background of a different colour; if the user is able to identify the shape, then he/she can distinguish the difference between the shape and background colours.

3.2.2 SSM Modelling Component

The modelling component of SSM takes measurements obtained during the calibration and constructs a full representation of a particular user's abilities in a particular situation. The nature of this component is very dependent upon what the model is representing, so only a brief description will be given here. A full description of the modelling process for CVD is given throughout the remainder of this dissertation.

Calibrations typically gather user responses for particular samples in the modelling space of interest. For example, when using SSM to model selective hearing loss, the minimum volume required for a user to hear a particular frequency may be assessed during calibration. As sound frequency is continuous, it is impossible to test the user's hearing ability for every single frequency. As a result, the user's hearing ability will be measured at particular frequencies and these measurements will be generalized to all frequencies. The process of generalizing calibration sample measurements to all possible values is the responsibility of the SSM modelling component.

In general, mathematical models can be used to balance the number of samples against the accuracy of the model. For example, in the selective hearing loss example just discussed, calibration samples can be taken at fixed frequency intervals and these measurements can be generalized using a cubic spline to generate a smooth function that describes the user's ability to hear every frequency. Clearly the number of calibration samples will affect the quality of this approach – the cubic spline will get progressively closer to the user's 'true' frequency hearing ability as calibration samples from more frequencies are gathered.

Generalizing calibration samples to a full representation of the user's abilities has impli-

cations for the design of the calibration and the modelling process. Calibrations must be designed to gather as much information as possible (to improve the modelling process) in the least amount of time (to reduce user fatigue). In turn, the modelling process must be carefully designed to both reduce the number of calibration samples taken (to reduce calibration time) and to maximally extract value from every calibration sample.

Iterative Modelling

Nothing in the design of the modelling process of the SSM requires that one model is generated for every calibration, but a single model can be continuously refined as new calibration data is gathered. This is particularly valuable when there are base conditions effecting a user's ability (e.g., an inherited condition that does not vary over time), that mix with other conditions that do vary over time (e.g., situationally-induced impairments).

An initial calibration could be performed in a situation that does not cause situationally-induced impairments (e.g., for an SSM-based model of a user's hearing, the initial calibration can be performed in a quiet room with little to no ambient sound). Then as the user's situation changes, new calibrations can be performed, and the base model updated to reflect the union of the user's internal invariant conditions and the external situationally-induced conditions. As a result, the model would be iteratively adapted to different situations, but could be reverted to the original initial calibration state when no situationally-induced impairments are present.

Iterative modelling also offers the benefit of reducing the calibration data necessary to adapt the model to a situation. For example, if it is determined from the initial calibration that a user has protanopia (missing long-wavelength sensitive cones), then this does not have to be re-tested in future calibrations as protanopia typically does not vary from condition to condition. As a result, calibrations that occur after the initial calibration could be optimized to gather data that reflect situationally-induced CVD (e.g., the latter calibrations could be restricted to simply measure luminance contrast perception abilities, which would require less calibration time than performing a full colour differentiation assessment). It is also possible that some of this information could be gathered using sensors, potentially eliminating calibration altogether.

3.3 Approaches Similar to SSM

In this section, approaches that are similar to the situation-specific modelling approach presented in this dissertation are discussed and compared to SSM.

3.3.1 Ability-Based Design

SSM is related to the general design approach called Ability-Based Design developed by Wobbrock et al. [151]. Ability-based design (ABD) is an approach that focusses on a user's abilities rather than his/her disabilities, and advocates a system capturing, measuring, and modelling the abilities of a user in order to allow the user to take advantage of their abilities when interacting with the system. The calibration component of SSM is akin to the capturing and measuring components of ABD; the modelling component of SSM reflects the modelling component of ABD.

ABD advocates integrating capturing, measuring, and modelling a user's abilities directly into the design of a system to allow the system to be adapted, or adapt itself, to the user. This is akin to incorporating the SSM-based individualized model of colour differentiation into a visualization system, thereby allowing the system to utilize the individualized model to select colours for use in a visualization. Pairing SSM with a recolouring tool (see Chapter 7) falls outside the scope of ABD however, because this constitutes an adaptation tool – a tool that adapts the user to the system and vice versa. ABD brings adaptation into the original system's design, thereby eliminating the need for explicit adaptation tools.

In developing ability-based design, Wobbrock et al. consulted a number of existing systems similar to SSM that model the abilities of a particular user. These systems are now briefly explored, discussing their relationship to SSM.

3.3.2 Existing Ability Modelling Systems

Keyboard text entry and command selection ability can vary considerably between users, especially those with motor impairment. This can lead to users unintentionally double-tapping keys, or dwelling longer than the system expects for a single key press. To address

this, Trewin and Pain [131, 132] developed a dynamic keyboard model that monitors a user's keyboard activity and constructs a model specific to the user. This model is then used to make suggestions about how to adjust system settings for key repeat rates and double-tapping sensitivity. This approach utilizes a passive calibration step (measurements are made without an explicit calibration step), that features iterative modelling.

Another system designed to model a user's typing is the Input Device Agent [81, 82, 83], except this system also models mouse pointing performance. This system incorporates an active calibration approach called *Compass* that is used to measure a user's performance. These measurements are used to develop a user-specific model, which is then used to suggest keyboard repeat rate, repeat delay, and settings for Sticky Keys. For mouse pointing, the system is used to recommend an optimal control-display (c:d) ratio [1] for mouse movements. In relation to SSM, Compass is an active calibration procedure used to develop a user-specific model. Neither Trewin and Pain's model or the Input Device Agent explicitly model situationally-induced impairments, which SSM is designed to accommodate, but each system could potentially be extended to address this.

Dynamically modelling a user's pointing ability and adjusting the control-display ratio to accommodate difficulties a user experiences is developed in Angle Mouse [150]. This system dynamically tracks angular deviation of mouse movements, increasing the c:d ratio when the deviation is low and decreasing the c:d ratio when the deviation is high. This has the effect of allowing rapid large mouse movements (e.g., ballistic movement towards a target), but slowing cursor movement when a user performs precise movements (e.g., attempting to select a small target). This system is somewhat similar to SSM in that user performance is monitored (like passive calibration), but user-specific models are not persistent: a relatively small time series of mouse movements is used to continually and dynamically develop a new model.

The system that is closest to the SSM approach employed in this dissertation is SUPPLE [39, 40, 41, 42, 43, 44], developed by Gajos et al., but this system does not model colour vision. SUPPLE utilizes an active calibration procedure that gathers a user's motor

[1] The ratio between physical mouse movement and on-screen cursor movement. Higher ratios result in small mouse movements producing large cursor movements on-screen. Lower ratios result in large mouse movements translating into small on-screen cursor movements.

abilities through a series of pointing tasks. This data is then used to generate a user-specific model of pointing performance. Once generated, this model is used to inform an interface layout optimization algorithm to construct a GUI interface that takes fullest advantage of the user's pointing abilities. Like the ICD described in the remainder of this dissertation, an active calibration process is used to develop a user-specific model which is then used to modify the interface to improve user performance. The main difference (besides SUPPLE's focus on motor impairments) is that the SSM proposed here requires a short calibration procedure as the calibration will need to be run whenever the situation changes to accommodate situationally-induced impairments. SUPPLE was designed to utilize a single calibration, and therefore is not optimized to dynamically modelling situational variations.

Similar to Angle Mouse, Hurst et al. developed a dynamic modelling approach to mouse movement, but this approach constructed a persistent (although changing) model of the user [65, 66]. A passive calibration process was used (pointing performance measurements were gathered during regular mouse use) to iteratively model the user's abilities. This model was used to classify users as novice or skilled; as having or not having pointing problems; as young or old; as needing or not needing specific adaptations; and as having or not having Parkinson's Disease, all with accuracies above 90%. Using empirically-gathered data to iteratively model a user's abilities and then make predictions about the user and his/her performance are all very similar to the SSM approach.

None of the systems presented so far have explicitly addressed situationally-induced impairments. To help resolve the situational impairments caused by walking, Kane et al. [78] developed *walking user interfaces*. Walking user interfaces are interfaces that automatically adapt to user movement; when a user of a mobile device begins walking, the interface presented to them is automatically updated to improve the legibility of text and the size of targets. These adaptations serve to address the motor and visual impairments induced by walking with a mobile phone. This approach incorporates the situationally-induced impairment nature of SSM, but does not model the user's unique abilities – each user is presented with the identical walking user interface.

3.3.3 Recolouring Tools Similar to SSM-Based ICD

As discussed in Chapter 1, recolouring tools require, as input, the type and severity of CVD to be accommodated. This problem has long been recognized as a shortcoming of many recolouring tools, but only a few systems have attempted to address it. These systems are discussed now, with critical reflection on the similarities and differences between these and the modelling approach presented in this dissertation.

Yang et al. began developing a recolouring tool that utilizes a computerized Farnsworth-Munsell 100-Hue Test (FM-100) [31] to determine the type and severity of the user's CVD [157], but only present initial findings that FM-100 scores are correlated with severity of CVD (which was a design goal of the original FM-100 test). No recolouring results are presented, the FM-100 is not used to determine type of CVD (only severity), and the path forward from this work is not discussed. The FM-100 is also a time-consuming test (taking approximately 15 minutes [31]), that does not accommodate environmental variations which can lead to situationally-induced CVD.

Yang et al. continue this work by developing a new colour vision test that is similar to the FM-100 test, but based on the HSV (hue, saturation, value) colour space [158] and a colour amplification recolouring tool that uses the results from this colour test as input [159]. The author's choice of the HSV colour space is not well suited to a colour vision test because it is not perceptually uniform, nor does it offer properties such as easy identification of isoluminant colour sets. These difficulties arising from HSV challenge the validity of their colour vision test, which in turn reduces the validity of the recolouring tool. Like the system presented above [157], the colour vision test is also not designed to be completed rapidly to facilitate frequent calibration in dynamic environments.

The modelling approach most similar to the final model presented in this dissertation is that developed by Mochizuki et al. [104, 105, 106]. In this model, the discrimination ellipsoids [114] in CIE XYZ colour space are measured for people with anomalous trichromacy, and then used to perform colour amplification recolouring. A discrimination ellipsoid is the 3D boundary around a colour in a colour space that defines the point at which an individual is able to just notice a difference (i.e., colours inside the ellipsoid are not differentiable from the

central colour; colours outside the ellipsoid are differentiable – a more extensive description of discrimination ellipsoids is given in Chapter 5).

By measuring discrimination ellipsoids, Mochizuki et al. capture the type and severity of CVD of an individual, and even are able to capture simultaneous CVDs (i.e., an individual having two or more types of CVD). This is very close to the goals of the SSM-based ICD presented in this dissertation. However, to measure discrimination ellipsoids, Mochizuki et al. chose to use CIE XYZ colour space, which is not a perceptually-uniform colour space. As a result, discrimination ellipsoids are required for a number of sample colours in order to properly assess an individual's colour differentiation abilities. Multiple samples are required because the discrimination ellipsoid for each colour in CIE XYZ colour space is unique. In a perceptually-uniform colour space, the discrimination ellipsoid for any colour can be simply rotated to find the discrimination ellipsoid for any other colour (see Chapter 6). As a result of this, Mochizuki et al. collected discrimination ellipsoids for 13 colours by linearly walking out along 16 vectors from each of the 13 colours. This linear walk involves presenting two colours to a user and asking if they are differentiable or not, and a very fine step was used for the linear walk (between 1×10^{-5} and 9×10^{-5} in a colour space with dimensions ranging from 0.0 to 1.0). Assuming a modest discrimination threshold of 0.01 CIE XYZ units, this would involve 13 colours $\times$ 16 vectors $\times$ 100 steps = 20800 colour pairs presented to the user. If responding to a colour pair took one second (approximately what I found in Chapter 4), then the calibration step would take about 20,800 seconds (5.8 hours). By comparison, the SSM-based ICD presented in this dissertation takes approximately 2 minutes to calibrate (see Chapter 5).

3.4 Using SSM to Construct an Individualized Model of Colour Differentiation

As discussed above, Situation-Specific Modelling (SSM) consists of the in-situ measurement of a particular user's abilities (calibration component) and the extension of the calibration measurements to a general representation of the user's abilities (modelling component). Through the measurement of the user's abilities in situ, the influence of most internal and external

factors that affect colour differentiation can be incorporated into the calibration process, and therefore the modelling process. A general overview of calibration and modelling components of the SSM-based Individualized Model of Colour Differentiation (ICD) are presented below.

3.4.1 ICD Calibration

As discussed above, SSM calibration can be an active process (explicit calibration step) or a passive process (implicit calibration). Some user abilities (e.g., pointing performance) can be measured passively because a prediction can be made regarding the user's intentions while performing a pointing action (e.g., he/she is attempting to click on the nearest target when the user presses the mouse button). Colour differentiation tasks however, typically do not have a feedback path to the computer. For example, website text and background colours could be manipulated to test colour differentiation abilities; but if a user navigates to a website and then leaves, it is very difficult to determine whether he/she was able to read the text or not. For this reason, the ICD presented in this dissertation utilizes an active calibration procedure. Of course, a passive calibration procedure is less intrusive for the user, so techniques for reducing the reliance upon active calibration are discussed in Chapter 8.

A naive approach to calibration would ask the user whether he/she perceives a difference between every possible pair of colours. These responses could be stored in a database, which would be consulted whenever a differentiability prediction was requested to simply return how the user responded during the calibration. This would essentially eliminate the modelling component of the SSM, as no generalization from the calibration data set is necessary – the calibration provides a complete data set.

However, RGB colour is the primary display and image colour space, and it represents colours using eight bits per channel, so a single colour is represented with 24 bits, giving 2^{24} different colours. All possible pairs gives $2^{24} \times 2^{24}$ pairings. If it takes the user one second to respond whether a pair contains differentiable colours or not, it will take 2^{48} seconds (approximately 8.9 million years) to complete a calibration. Even if a display only uses 16 bits to represent each colour (five for red, six for green, five for blue), this still gives $2^{16} \times 2^{16}$ pairings, resulting in a 136-year calibration. Reducing this even further, if only 10 bits are used, this still would result in a calibration that would take over 12 days to complete. The

naive approach is obviously not feasible.

Instead of collecting colour differentiability measurements for every possible pair of colours, ICD calibration will gather colour differentiation measurements for a sample of colours. Colour differentiation is an example of a *psychophysical* task, a class of activities that involve relationships between stimuli in the world (here, it is colours) and sensation in a human (here, perception of 'differentness') [45]. Differentiation is specifically related to the idea of signal discrimination, which considers the minimum difference between two stimuli such that the difference can be perceived (also called the 'just-noticeable difference' or JND) [14]. As with most discrimination tasks, colour differentiation is an easy task if the two colours are very different, but becomes more difficult as the two colours become similar.

Certain colour spaces (e.g., RGB, CIE L*u*v*) incorporate a distance property that allows the idea of 'similarity' for colours to be quantified, and to better characterize a user's ability to perform colour differentiation. Considering colour spaces as physical spaces (in which each colour occupies a unique location) allows the metaphor of movement through a colour space. As one moves away from any particular colour, increasingly different colours are encountered. While moving on any path away from a colour, a point will eventually be reached where the colours that are being encountered switch from not differentiable to differentiable, in reference to the starting colour. This transition is called a psychometric function, which can be described mathematically using a sigmoid curve. The sigmoid must be approximated in a model, and this approximation can range from a simple step function, which transitions from not differentiable to differentiable in a single step (Figure 3.2), to more complex approximations. The point where the single step takes place or the 50% level for the sigmoid function is called the *differentiation limit* and represents the limit of the colours that are not differentiable from the reference colour.

The calibration procedures presented in this dissertation will find differentiation limits for a set of colours. Stimuli are repeatedly shown to the user and responses about the stimuli are collected from the user. For measuring colour differentiation abilities of a user, the stimuli will be pairs of colours, and the user's response will indicate whether the user can tell a difference between the colours or not. One colour will be the colour for which a differentiation limit is being measured, and the other colour will be a colour that is some distance away from

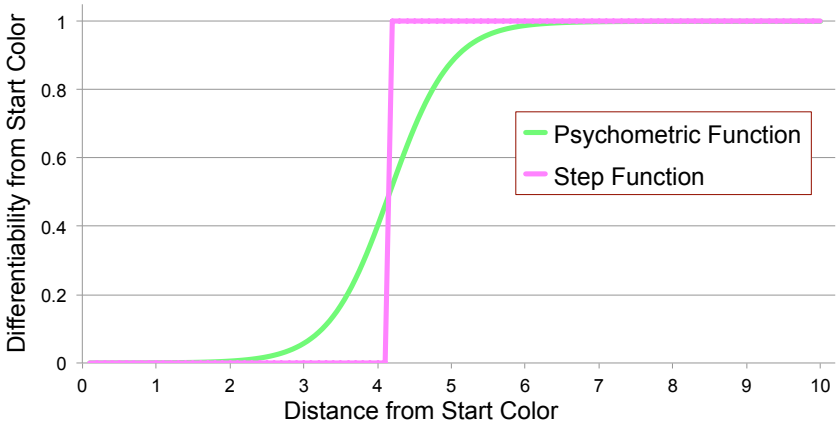


Figure 3.2: Approximation of a normalized psychometric sigmoid function with a simple step function.

the first colour. By controlling the difference between these two colours in each stimuli, the point at which the two colours become differentiable (differentiation limit) can be found.

As discussed above, calibration can consist of performance or judgement-based tasks. The first ICD calibration technique (presented in Chapter 4) utilizes a judgement-based calibration task. During the evaluation for this ICD version, I recognized that the calibration process was generating many false positives – the user was mistakenly reporting that the colours were differentiable when they in fact were not because they were identical. To address this, the calibration presented in Chapter 5 incorporates a performance-based calibration task.

Calibration techniques for the ICD are described in detail in Chapters 4 and 5.

3.4.2 ICD Modelling

As discussed above, the modelling component of SSM extends the measurements obtained during the calibration procedure to a full general representation of the user's ability. For the SSM-based ICD, the modelling component takes the differentiation limits for the restricted set of colours used in the calibration procedure, and generalizes them so that differentiation limits for any colour in the working colour space can be identified.

With the ability to identify the discrimination limits for any colour, the differentiation limits can be used to define a discrimination volume around a colour. All colours within

this volume are predicted to be not differentiable from the starting colour, and all colours outside the volume are predicted to be differentiable. Once the model is able to construct a discrimination volume for any colour, a prediction about the differentiability of any two colours can be made. In computational terms, this could be accomplished through a simple API that provides a single function `areDifferentiable(Colour C1, Colour C2)`. This function accepts two colours and returns `true` if the model predicts that they are differentiable and `false` otherwise. To accomplish this, the function determines the discrimination volume for C1, and then checks to see whether C2's colour coordinates are located within this volume. If it is not in the volume, then the colours are differentiable. This approach is the general approach utilized in the model described in this dissertation.

By providing a single `areDifferentiable(Colour C1, Colour C2)` function, the SSM-based ICD is able to meet the fundamental needs of recolouring tools. As discussed in Chapter 1, recolouring tools need to identify problem colours and find suitably-differentiable replacement colours. If a recolouring tool is provided with an ICD that encodes the colour differentiation abilities of a user with typical colour vision and another ICD that encodes the colour differentiation abilities of the user with CVD, the recolouring tool can compare responses between the two models for different pairs of representative colours in an input image to identify problem colours. If two colours are differentiable for the user with typical colour vision, but not differentiable for the user with CVD, then these colours are problem colours. To find suitably-differentiable replacement colours, the differentiability of potential replacement colours for the user with CVD can be compared to the differentiability of the original problem colours for the user with typical colour vision. If they are both differentiable, then suitably-differentiable replacement colours have been identified; if they are not differentiable, then the recolouring tool needs to continue to search for replacement colours.

Chapter 4 presents a feasibility implementation of the ICD with an initial calibration technique and a method for generalizing the calibration measurements to a full model of colour differentiation for a user. Chapter 5 presents a more advanced modelling technique that is more closely aligned with human colour perception, and an accompanying calibration technique.

3.5 Additional ICD Constraints for Recolouring

Two additional requirements for the Individualized Model of Colour Differentiation (ICD) are low prediction time and high prediction power.

As presented in Chapter 2, recolouring tools typically perform a computationally-demanding constraint optimization to find a recolour mapping. This can result in hundreds of thousands of calls to `areDifferentiable`, so making this function fast is essential to enable fast recolouring performance.

Additionally, recolouring tools generally do not seek to find replacement colours that are simply differentiable for the user with CVD, but rather seek to find suitably-differentiable replacement colours – colours that have the same degree of differentiability for the user with CVD as the original colours for someone with typical colour vision.

To reduce differentiation prediction time, the ICD can be improved to build a single discrimination volume for a calibration and then modify this volume using affine transforms to answer a prediction query. To improve the prediction power, the ICD can be extended to provide a `howDifferentiable` function that reports how differentiable the two colour parameters are for the user. Chapter 6 presents modifications to the ICD that reduce the time to perform a prediction and increase the prediction power.

Finally, the problem presented in this dissertation is *existing recolouring tools do not help most people with CVD*. The SSM-based ICD provides a way to model the colour differentiation abilities of a particular user in a particular situation, which should address this problem. To show that the SSM-based ICD is an effective solution, Chapter 7 describes an ICD-based recolouring tool and its evaluation in a user study. This study conclusively shows that the SSM-based ICD approach extends recolouring tools to helping many people with CVD, including those with inherited, situationally-induced and simultaneous CVD.

I begin by describing a feasibility implementation of the ICD in the next chapter.

Chapter 4

Feasibility of Individualized Models of Colour Differentiation

This chapter presents an initial implementation of the individualized model of colour differentiation (ICD-1) that tests the feasibility of utilizing the situation-specific modelling approach presented in the previous chapter to model individual differences in colour differentiation abilities. The motivation for a feasibility study is given first, followed by the technical details of the implementation of ICD-1. Following this, an evaluation of ICD-1 is presented, followed by a short discussion of the results of the evaluation that identifies where ICD-1 succeeds as well as areas for improvement. [1]

4.1 Feasibility of ICDs to Meet Design Goals

The solution presented in this dissertation (Section 1.3) is to construct a model of colour differentiation that captures the effect of every internal and external factor that influences colour vision. Situation-specific modelling (SSM) was presented as a means to achieve this in Chapter 3, and this approach entails a calibration step for measuring a user's colour differentiation abilities in situ, and a modelling step that takes these measurements and constructs a model that is able to answer the query `areDifferentiable(Colour c1, Colour c2)`.

The primary goal of such a model is that it should be a more accurate model of a user's colour perception – in particular, a more accurate model of their ability to differentiate colours on a computer display. Secondary goals are that the model should be easy and cheap

[1] Portions of this chapter were published as: Flatla, D.R., Gutwin, C. (2010) Individual Models of Color Differentiation to Improve Interpretability of Information Visualization. In *CHI '10: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2563–2572.

to obtain, and should be compatible with existing approaches for recolouring.

The purpose of the feasibility implementation presented in this chapter is to evaluate the efficacy of the SSM approach for developing individualized models of colour differentiation by determining how well ICD-1 achieves the goals just described. To accomplish this, an initial implementation of an ICD is described below (both the calibration and the modelling components), followed by an evaluation of this model in a user study that explores the properties of this model in terms of these goals.

4.2 ICD-1: Individualized Models of Colour Differentiation Using RGB

RGB colour space will be used in the implementation of ICD-1. Almost all images are encoded using some form of RGB colour space, and almost every display uses RGB colour space; it is the dominant colour space in digital environments. Additionally, conversions from RGB to other colour spaces (e.g., to CIE L*u*v*) make assumptions about the gamma response curve, white point, and spectral properties of the colour producing elements of a display (e.g., the peak wavelength of the spectral distribution of each red, green, and blue sub-pixel in a CRT display). By avoiding these conversions and relying on only RGB colour space, the negative effects of these potentially-incorrect assumptions will be minimized.

I recognize that RGB is a simplistic colour space choice for the feasibility implementation presented in this chapter, i.e., there may be costs associated with using RGB colour space that outweigh the costs associated with potentially imprecise assumptions when converting from RGB to other colour spaces. Indeed, future implementations of ICD (Chapters 5–6) are based on more sophisticated colour spaces (in particular CIE L*u*v*), and integrate significant improvements to the calibration and modelling components while also improving the prediction accuracy of the ICD approach. That said, the improvements presented in later chapters would not have come to light without a feasibility or baseline implementation to improve upon and compare against.

4.2.1 RGB Colour Space

The RGB colour space used in ICD-1 uses 24 bits to represent any colour in that space. These 24 bits are evenly split into three 8-bit chunks. The first 8-bit chunk (most significant 8 bits) is used to represent the amount of red in a colour, the middle 8-bit chunk represents the amount of green in a colour, and the last 8-bit chunk represents the amount of blue. RGB is an *additive* colour space in that the individual components of a colour (red, green, and blue) are composed to create the final colour. Eight bits per channel allow 256 discrete values for each channel, giving 2^{24} possible colours in the RGB colour space used here.

When a channel has a value of 0, none of that channel (red, green, or blue) is in the final colour, and when a channel has a value of 255 (maximum), the channel is maximized for the final colour. For example, the colour (0,0,0) represents black (none of any channel), and the colour (255,0,0) represents a bright red (with no green or blue). When each channel contributes an equal amount to the final colour (i.e., each channel has the same value), a greyscale colour results (including black, as above, and white which is represented by (255,255,255). The RGB colour space can be simply visualized using a cube in which each channel occupies one dimension of the cube (see Figure 4.1).

4.2.2 Calibration: What information will build the model?

The general approach to constructing the first individualized model of colour differentiation (ICD-1) based on the situation-specific modelling approach will be to test the user's colour differentiation capabilities at different parts of the RGB colour space, and use these empirical values to configure a model that provides the `areDifferentiable` function described in Chapter 3. The term 'the user's colour differentiation capabilities' means the smallest RGB colour space difference between two colours, both visible on screen, that the user sees as different (the just-noticeable difference, or JND [14, 45], as discussed in Chapter 3).

The most accurate model possible would test the user's colour differentiation ability with each possible colour combination (2^{48} possible combinations when eight bits are used to represent each RGB channel), but sampling must be used to reduce the amount of testing – that is, ICD-1 will test fewer points in the colour space, and interpolate between these

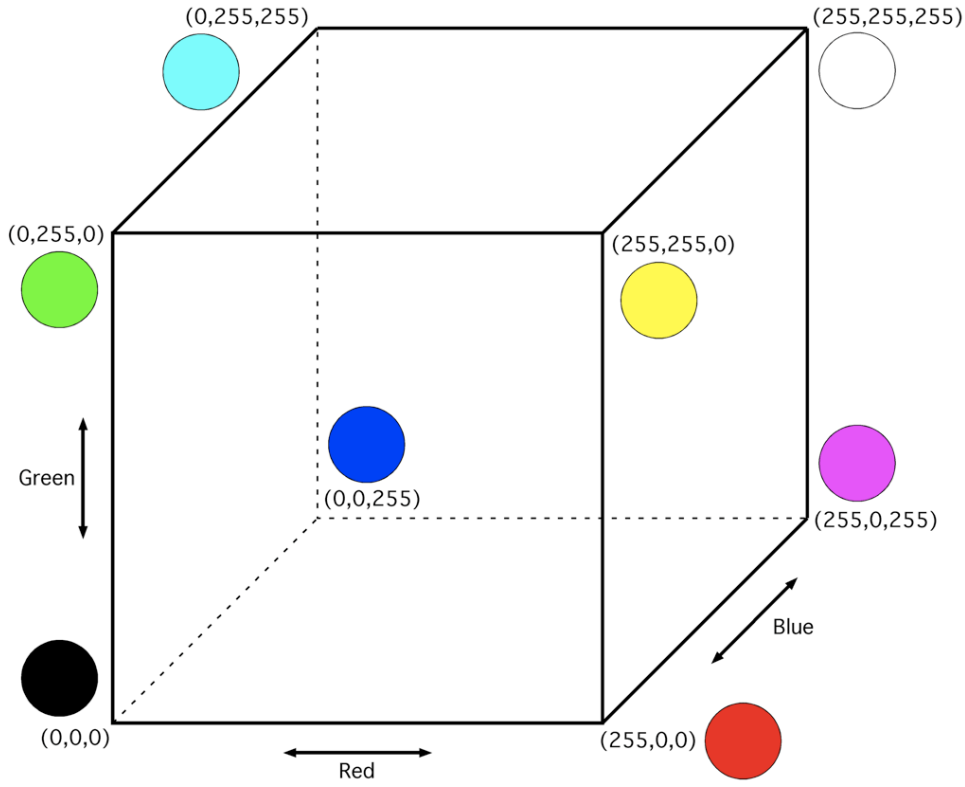


Figure 4.1: RGB colour cube with the colour at each corner shown. Red channel runs left to right, green channel runs up and down, blue channel runs front to back. The line from black (0,0,0) to white (255,255,255) contains the achromatic (greyscale) colours.

points when needed. There are many possible interpolation functions, but based on previous research into perception of sensory stimuli (e.g., [14, 45]), linear interpolation between samples is used for ICD-1.

To measure a user's colour differentiation abilities, a calibration technique was developed that gathers information for calibrating the ICD-1 model to a specific user in a specific situation. To gather this information, a simple judgment-based task is used in which two colours are presented to the user, and the user provides a response indicating whether the two colours are differentiable or not. To measure a single just-noticeable difference for a colour, the user is presented with an 8x6 grid of circles in which half of the circles are one colour and the other half of circles are set to a colour that is a certain distance away from the first colour in RGB colour space (see Figure 4.2). For each presentation, the arrangement of the two colours is randomized to eliminate the visual strain caused by repeated exposure to identical visual patterns. If the user responds that the colours are 'differentiable', then the

distance is reduced and the screen redisplayed. If the user responds that the colours are 'not differentiable', then the distance is increased and the screen redisplayed. This is continued in a binary search fashion until the just-noticeable difference point is identified. To help reduce the errors introduced by the user having 'momentum' in their responses (e.g., providing many 'differentiable' responses in a row, then accidentally labelling 'not differentiable' colours as 'differentiable'), the presentations for each step of every differentiation limit binary search are interlaced and randomized. As the maximum range of RGB channel values that need to be explored via binary search is 256 for 8-bits-per-channel colour, the user needs to provide about ten responses to find a single differentiation limit.

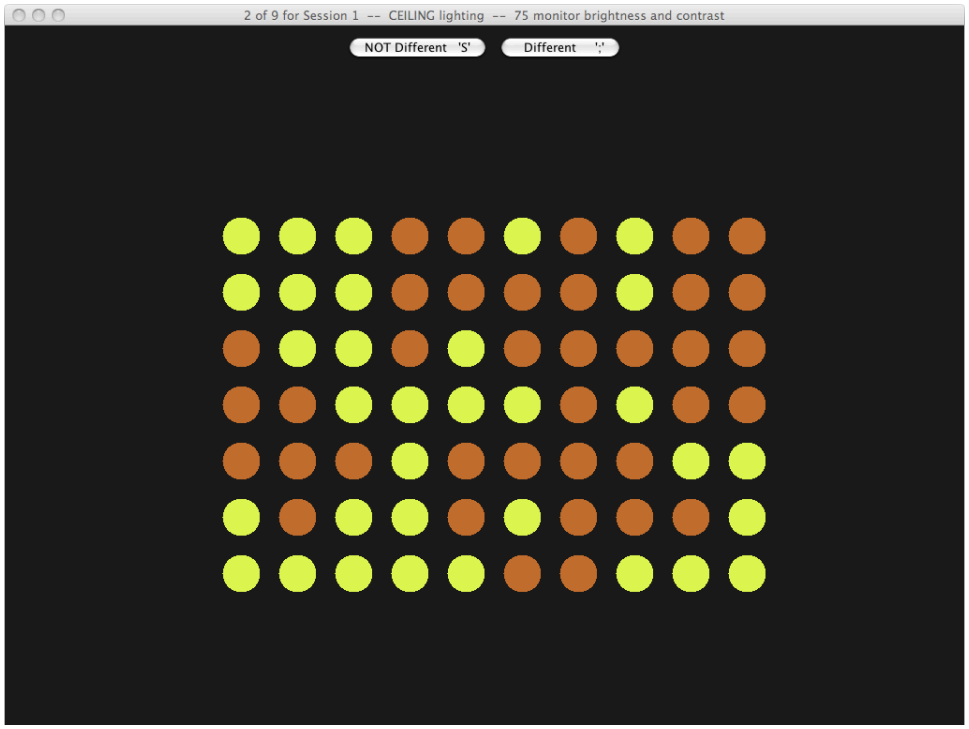


Figure 4.2: Presentation of differentiation task for calibration of ICD-1.

In preliminary studies, it was found that a simple two-swatch (e.g., two circles or semicircles) interface did not work well because it resulted in substantial retinal after-image effects. As a result, I employed a grid of coloured circles in which two colours were randomly assigned to the circles. This prevented repeating identical patterns of colours, thereby offsetting the after-images experienced by users.

4.2.3 ICD-1: What will the model predict? Differentiation limits

ICD-1 will make predictions on the differentiation limits for a given colour. A differentiation limit is defined as the colour space distance between two colours at which the two colours just become differentiable. For ICD-1, there will be six differentiation limits for any colour, two for each of the three colour channels (R, G, and B). The two limits are the upper limit (how much of a channel needs to be added to make the colours differentiable), and the lower limit (how much of a channel needs to be subtracted to make the colours differentiable).

For any colour in the RGB colour cube, these six limits define a *discrimination box* within the colour cube that contains all colours that cannot be differentiated from the input colour (see Figure 4.3). As a simple example, consider the process for a single channel (e.g., red) for the colour shown in Figure 4.3 (RGB = (140,140,80)). Along the red channel, the upper limit for a user is the amount of red channel that needs to be added to the original colour so that a differentiable colour is produced (i.e, the colour (140+upperlimit,140,80) is just differentiable from (140,140,80)). The lower limit is similar, but for values below 140 (i.e, the colour (140-lowerlimit,140,80) is just differentiable from (140,140,80)). If the lower limit is 18 and the upper limit is 22, then the colours in the range (123,140,80) – (161,140,80) are not differentiable from the original colour (140,140,80) for this user in this situation.

The discrimination box can then be used to determine whether colours in a display are safe to use. The discrimination box is determined for one of the two colours passed to `areDifferentiable`. If the other colour parameter lies outside of the first colour's discrimination box, then the colours are predicted to be differentiable (i.e., `true` is returned), but if the other colour parameter lies within the first colour's differentiation box, then the colours are predicted to be not differentiable (`false` is returned). In the above example, if two colours with RGB values of (128,140,80) and (140,140,80) were passed to `areDifferentiable`, the model would predict that those colours are not differentiable for the user.

Recognizing that human colour differentiation limits are probably not well-described by a rectilinear box, the next chapter (Chapter 5) explores the use of discrimination ellipsoids to represent colour differentiation abilities. Discrimination ellipsoids are a more accurate representation of human colour differentiation than the boxes used in this chapter [114, 153],

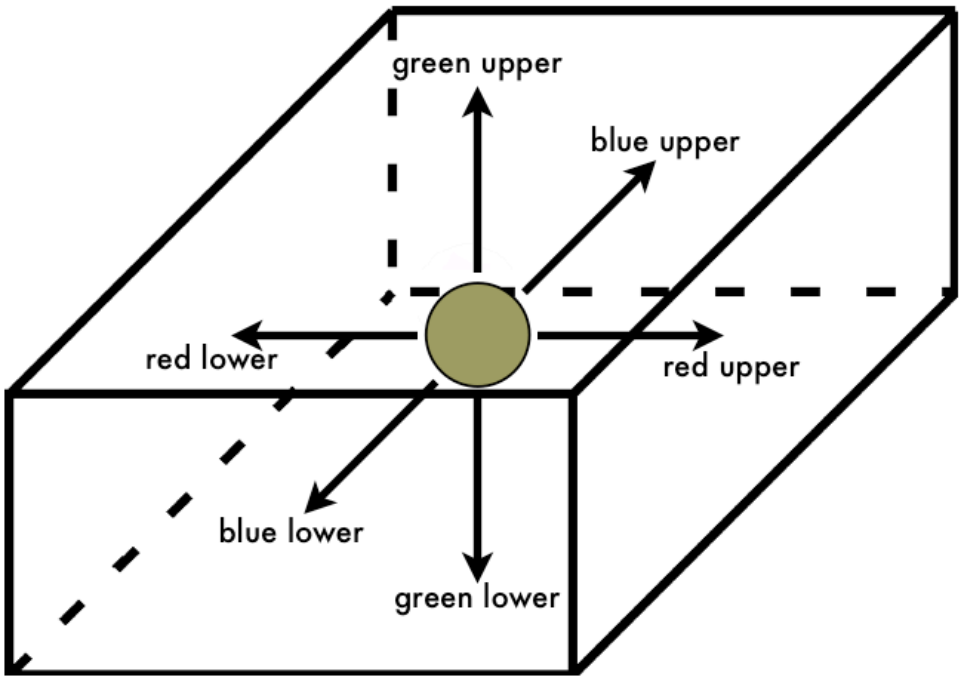


Figure 4.3: Example RGB colour space ICD-1 discrimination box for a given colour.

however discrimination boxes are much simpler, thereby allowing the feasibility of ICD to be assessed before further model improvements are made.

4.2.4 ICD-1 Example

As stated above, sampling is used to reduce the calibration measurements required for the model, and interpolation is used to estimate differentiation limits for colours not directly measured during calibration. As a simple example of how interpolation will be used, consider the case of differentiation along a single channel as introduced above. A model of the user's differentiation ability will be built by testing the user at different points on this channel, and then linearly interpolating between these known points.

The simplest model would be built from two samples, such as the two endpoints: that is, ICD-1 would test the user's upper differentiability limit when R=0, and his/her lower limit when R=255, and then use these to interpolate all other limits. Assume that these limits are determined to be 10 and 35, respectively, and that the other two channels are held constant for now. A key observation is that it is not necessary to measure both upper and lower differentiation limits for each colour because one colour's upper differentiation limit is another

colour's lower differentiation limit (e.g., in the example given in Section 4.2.3, the upper red channel limit of 22 for (140,140,80) is also the lower red channel limit for (162,140,80)). As a result, these two differentiation limits (10 and 35) can be used to determine two pairs of limits, one upper and one lower, as shown in Figure 4.4 (upper limits: R_0=10, R_{220}=35; lower limits: R_{10}=10, R_{255}=35). As an upper limit cannot be measured for R=255 (ceiling), and a lower limit cannot be measured for R=0 (floor), the lines for the limits do not cover the entire scale, but can be extrapolated if necessary. Using the equations for the lines defined by these two pairs of limits, the upper and lower differentiation limits for all red-channel values can now be determined by simply solving each equation for a given red-channel value.

The simple one-channel example described above does not take into account the possible influence of other channels in differentiability. That is, differentiability with R=0 may be different when G=50 and B=50 than when G=200 and B=200. Therefore, additional samples will be needed for the R channel to account for the influence of different G and B values. Linear interpolation is used again to predict for colours in between our input samples. This means that the example model from above now requires the user's differentiation limits for the R channel with four combinations of G and B: 0,0; 0,255; 255,0; 255,255, resulting in eight differentiation limits for each channel in a 2-sample model of RGB colour (i.e., three differentiation limits are needed for each outside corner of the RGB colour cube – upper limits when the value of a channel is 0 at a corner, and lower limits when the value of a channel is 255 at a corner).

Using the eight colours represented by the corners of the RGB colour cube, this simple model can be extended to a full model. To begin, upper or lower differentiation limits for each channel must be collected at these eight colours, giving 24 limits to measure, one for each channel at each corner of the RGB cube, using the calibration technique described above. Also described above is the key observation that the upper limit for one colour is the lower limit for another colour, allowing ICD-1 to extend the 24 limits to 48 limits (24 upper and 24 lower).

Using these 48 limits, linear functions that describe the limits for a channel as one 'moves' along an outside edge of the RGB cube can be generated using interpolation. Continuing the red channel example, the edge defined by the corners (0,0,0) and (255,0,0) will use endpoint

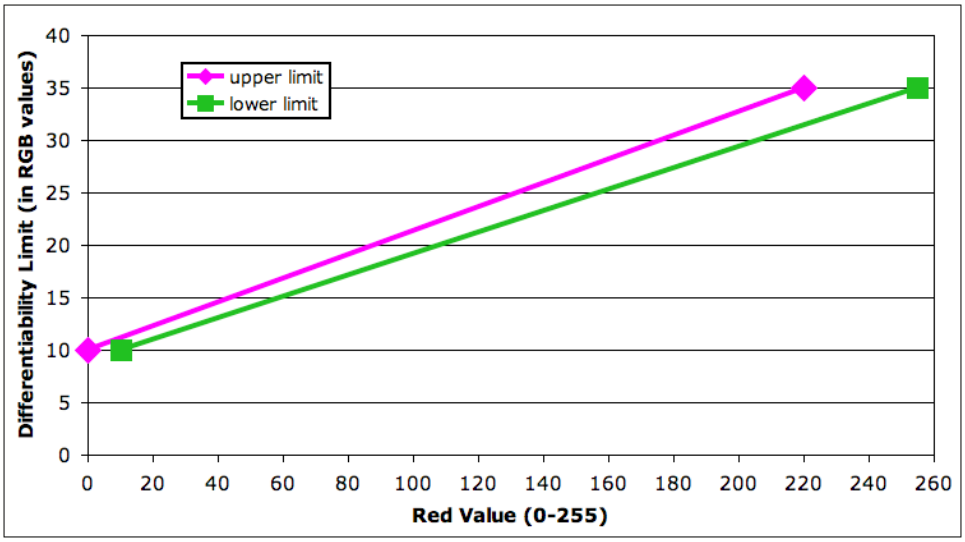


Figure 4.4: Linear interpolation in a simplified 1-channel 2-sample ICD-1 model.

differentiation limits to define two functions – one for upper limits and one for lower limits. There are twelve of these edges: four parallel to the red channel dimension, four parallel to the green dimension, and four parallel to the blue dimension. Each edge holds two functions – one defining upper differentiation limits and one defining lower differentiation limits – for the RGB channel that varies with movement along the RGB cube edge.

When the discrimination box for a colour is requested, each channel is processed independently and the differentiation limits for each channel are used to calculate the final box. To process the red channel, the red channel upper and lower differentiation limit linear functions are used. First, the red value for the colour is used to determine the upper and lower differentiation limits along each edge parallel to the red channel dimension of the RGB colour cube. The red channel value for the colour is used to solve each linear function, giving an interpolated differentiation limit. This gives four upper and four lower differentiation limits for the red channel, but as discussed above, these limits need to be placed into the context of the green and blue channel values for the colour. As such, the four upper and four lower differentiation limits are used to define two upper and two lower differentiation limit linear functions across the green or blue dimension (arbitrarily picking the green channel here). The green channel value for the input colour is used to solve these functions, resulting in two upper and two lower differentiability limits. Finally, these four differentiability limits are used to define two (one upper and one lower) differentiation limit linear interpolation

functions spanning the blue channel. The blue channel value for the input colour is used to solve each of these functions, giving a single upper and lower differentiation limit. These final values represent the red channel differentiation limits for the input colour.

To determine the green differentiation limits, the green differentiation limit functions are used with the green channel value to establish the four upper and four lower limits (two functions), then the red value to determine two upper and two lower limits (one function), and then the blue value to determine the final upper and lower limits. Blue limits are found by using blue, then red, then green colour values to repeatedly interpolate.

It may be noted that the order after the initial selection of four points is irrelevant (red $\rightarrow$ green $\rightarrow$ blue is equivalent to red $\rightarrow$ blue $\rightarrow$ green for red channel limits), and in the implementation of ICD-1, every limit is calculated using both approaches and cross-validated to ensure correctness.

4.3 Evaluation of ICD-1: Model Accuracy and Robustness

To explore the effects of different numbers of samples on model quality, and to evaluate the accuracy and effectiveness of the models, a set of empirical differentiation data was gathered. The study collected data from 16 participants, and investigated three questions:

- Do models with more input samples perform better than those with fewer samples?
- Do the models perform differently for users with CVD compared to those with typical colour vision?
- How robust is the model to moderate changes in environmental factors such as light or background?

4.3.1 Methods

This section describes the experimental design used to evaluate ICD-1 in terms of the three research questions listed above.

Participants, Apparatus, and Task

Sixteen volunteers (all male, mean age 26.0 years) were recruited from the local community. Eight participants had CVD to some degree (based on the Ishihara plate test [71]), and eight had no indication of CVD. The Ishihara test was performed by the authors in a non-clinical setting, and was used simply to identify the presence or absence of CVD. These tests revealed a mix of protan and deutan (both forms of inherited red-green CVD) effects in the participants with CVD.

The study was carried out in a room with controlled lighting, and used a custom-built Java application to measure the participant's colour differentiation ability. The application presented participants with a series of differentiation tests identical to the technique described in Section 4.2.2 and shown in Figure 4.2; the participant's job was to state for each test whether the two colours on the screen were the same or different, by pressing one of two keyboard keys. The participant's responses to these questions were used to empirically determine their differentiability limits. To save time, only upper differentiability limits were gathered for this study.

Participant instructions are shown in Figure E.1. Each field of circles was presented to the participant for two seconds. This time limit was chosen because initial pilot studies suggested that allowing more time did not substantially improve the accuracy of the results, but did result in longer study completion times. Participants were instructed to make their selections as quickly and as accurately as possible.

Models with Different Numbers of Samples

One of the goals of the study was to determine how accuracy is affected by the number of samples in the model. As described above, 24 differentiation limits represents the minimum number of limits to gather, so this would result in a shorter calibration, but also could mean reduced accuracy. To assess the tradeoff between model accuracy and calibration time, I gathered calibration data from 125 colours – that is, from five evenly-spaced points on each of the RGB channels (values of 0, 55, 110, 165, and 220). This resulted in 125 colours x 3 channels per colour x 1 limit per channel = 375 differentiation limits collected per participant.

From this data, I built four models: a 5-sample-per-channel model using all of the data, a 4-sample-per-channel model using the lower four points from the list above, a 3-sample-per-channel model using the endpoints and middle value, and a 2-sample-per-channel model using only the endpoints. (I shorten the names of the models from here on to 'the n-sample model').

Robustness

The goal of the empirical modelling approach is to increase the specificity of the model – to be able to automatically include the specific characteristics of the user and their environment. One potential risk of this approach, however, is an overly-specific model that is too specific to the details of the current environment and thus not effective when any of those details change. Although there are unlikely to be dramatic changes to the context (e.g., a user's CVD will not switch from one type to another, and a user's monitor will not suddenly switch calibration), there are several ways in which minor changes can occur (e.g., lighting conditions will change over a work day).

To determine how quickly the model's accuracy degrades as the environment changes, another set of limits were gathered. In addition to the 375 calibration limits measurements described above, a set of eight additional colours was randomly generated, and used to collect participant upper differentiation limits for each channel of these eight colours in four different situations. This data allowed the user's 'ground truth' responses in each situation to be compared to the predictions made by ICD-1 to evaluate the model's robustness to situational variation. I used the same 'differentiable'/'not differentiable' task described above. The additional conditions were:

- *Lighting.* In addition to the normal ceiling lighting that was used for the standard tests, I gathered test sets with low lighting (all lights off), and lamp light (a lamp shining on the screen at approximately a 45° angle).

- *Background grey.* As background colours influence human colour perception (and therefore colour differentiation) [130], I included two additional background colours (grey levels 128 and 230) in addition to the dark grey background used for the standard tests

(grey level 25).

- *Monitor colour adjustment.* The display monitor allows the adjustment of RGB gains; in addition to the normal value of 75 for each channel, I also collected data with the RGB gains each set to 50 and 100.

- *User fatigue.* I compared differentiation limits gathered at the beginning and the end of the experiment to determine whether users' responses change as they become more tired.

Each of these four situations were tested independently, by holding the remaining situational conditions at their calibration levels. As a result, the fatigue data acted as the baseline measurement with no situational variation (except, of course, user fatigue). This resulted in 2 fatigue repetitions x 8 colours x 3 channels per colour x 1 limit per channel = 48 + (8 colours x 3 channels per colour) x (2 lighting + 2 background + 2 monitor) = 144 = 192 differentiation limits gathered per participant for this part of the study.

Procedure

There were three parts to the study: collection of calibration data, collection of standard test data, and collection of test data under different environmental conditions. Due to the large number of differentiation limits gathered, limits were only gathered once (i.e., no repetitions).

Part 1: Calibration data. The model was calibrated based on empirical samples. In this phase of the study, the 375 differentiation limits described above were gathered, from which the different models (5-sample, 4-sample, 3-sample, 2-sample) were built.

Part 2: Empirical data for model testing. To test the accuracy of the various model configurations, empirical data was gathered about the user's actual differentiation limits for the eight colours used for the robustness testing described above. Each colour was tested eight times in random order to better estimate the user's true differentiation limits. This resulted in an additional 8 colours x 3 channels per colour x 8 repetitions = 192 differentiation limits being gathered.

Part 3: Empirical data for robustness tests. To test the model's accuracy when environmental conditions change, I also gathered the 192 differentiation limits described above for

the eight test colours in each of the four situations described above.

As described in Section 4.2.2, it takes approximately ten coloured circle grid presentations to gather a single differentiation limit. Each presentation takes approximately one second to complete, resulting in the study taking approximately 10 presentations x (375 + 192 + 192) differentiation limits = 7590 presentations x 1 second/presentation = 7590 seconds = just over 2 hours to complete. Due to the extreme length of this study, participants were allowed frequent breaks to rest their eyes.

Two analyses were carried out to evaluate the model: first, the accuracy of the model was tested using the empirical differentiation data gathered in Part 2 of the study; second, the robustness of the model was tested using empirical differentiation data gathered in Part 3.

4.3.2 Validating the Model: Accuracy

The first analysis tested the accuracy of ICD-1 by comparing the predictions made by the model with the empirical data (which was not used in the formation of the models). As described above, the model predicts differentiation limits for a given colour, so the model predictions were compared to the empirically-determined limits gathered for the eight test colours in Part 2 of the study. Four models were tested, each with a different granularity (2, 3, 4, and 5 samples per channel).

For each of the eight colours in our test set, the model predicted the differentiability limits for the R, G, and B channels. Part 2 of the study empirically determined these same differentiability limits, and these are used as the 'ground truth' values against which the model's predictions were compared.

To assess the predictor's accuracy, the types of errors that a predictor can make were considered. The ICD-1 predictor can make two types of error – either over or under the true limit value. When the predictor exactly matches the ground truth value, it is not an error. It was noted that the two types of error are not equal in real-world terms, since an over-estimation will result in false negatives (where the model says that two colours are not differentiable, and the empirical data shows that they are), and an under-estimation results in false positives (where the model predicts that two colours are differentiable but the empirical data shows that they are not). Over-estimation therefore presents much less of a problem

in terms of an ICD-based recolouring tool: it avoids mistakenly allowing non-differentiable (problem) colours to remain in the original image, but it does reduce the number of colours that can be used.

Based on this analysis, the term 'safe accuracy' is used to represent the proportion of predictions that are 'safe' – that is, that will not result in a false positive error. This measure (A_{safe}) is the ratio of exact predictions plus over-estimations to the total number of cases.

It is possible to intentionally increase the predicted limits by a constant (called the *limit offset*), in order to increase the over-estimation and reduce the number of false positives. This moves the distribution of errors towards the 'safe' side of the mean. The limit offset can be included as part of the model, and allows the model to be tuned so that predictions trade off between false positives and false negatives. The limit offset was therefore used as the measure by which the models are compared – in the tests below, the limit offset needed to achieve a safe accuracy of 0.95 is reported (i.e., such that 95% of predictions are either correct or over-estimates).

Figure 4.5 shows the histogram of prediction errors around the empirically-derived value (4-sample model); errors are somewhat normally distributed, but skewed toward over-estimation, with a long-tail of under-estimations. The model makes approximately the same number of over- and under-estimates. Table 4.1 and Figure 4.6 show the accuracy results. Accuracy was tested separately for the two groups of participants (CVD and non-CVD), to determine whether CVD had an effect on the model's accuracy. This showed that higher offsets were needed for the participants with CVD to maintain the same safe accuracy. It can be seen from Figure 4.6 that four samples per channel (4x4x4 = 64 sample colours) appears to be the appropriate number of samples, as it exhibits improved accuracy over the 2 and 3-sample models, and little to no advantage comes from the increased sample rate of the 5-sample model.

The explanation for the reduced accuracy with CVD users is that this group generally has much larger values for their limits, meaning that the linear interpolation functions hit a ceiling (255) more quickly than for non-CVD users. The model therefore has fewer values with which to build accurate interpolation functions, resulting in a less accurate model with a higher required offset.

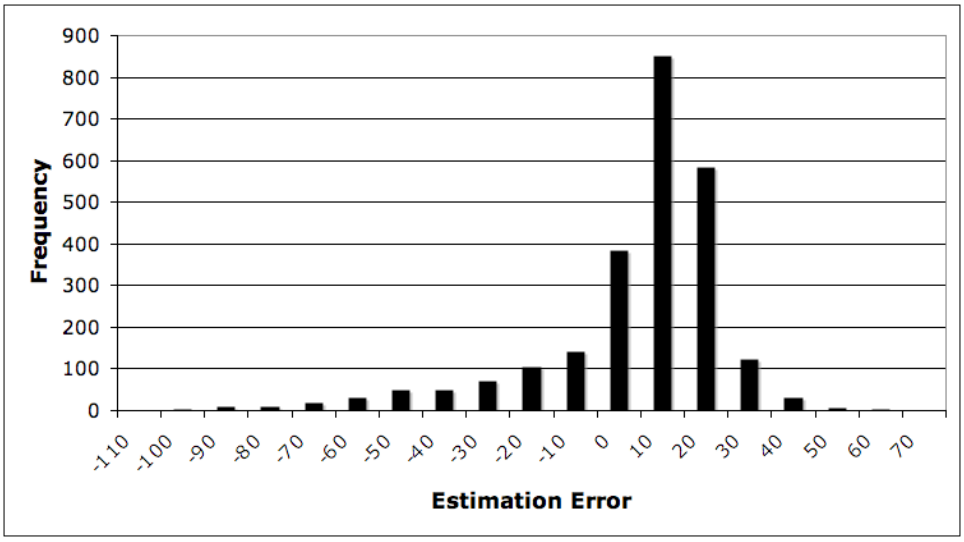


Figure 4.5: Distribution of prediction errors around the true value for the 4-sample model (model predicted differentiation limit minus 'ground truth' differentiation limit).

	5-sample		4-sample		3-sample		2-sample	
Group	Offset	S.D.	Offset	S.D.	Offset	S.D.	Offset	S.D.
CVD	52	24.3	52	23.5	67	27.5	81	24.8
nonCVD	41	18.0	43	18.7	46	19.0	58	15.7
All	47	21.1	48	21.1	57	23.2	69	20.2

Table 4.1: Mean accuracy results by model granularity: Offset indicates the value needed to reach 0.95 A_{safe}; S.D. = standard deviation of estimation errors.

Two further issues to be addressed in this accuracy analysis are the relationship between the limit offset and the safe accuracy, and the degree to which over-estimation will reduce the number of colours available to an ICD-based recolouring tool.

Relationship Between Limit Offset and Safe Accuracy

Figure 4.7 shows how addition of different limit offsets affect safe accuracy. It is clear that covering the last remaining cases of under-estimation is expensive in terms of the amount of colour space that is used; however, it is possible to reach any safe-accuracy level, including 100%.

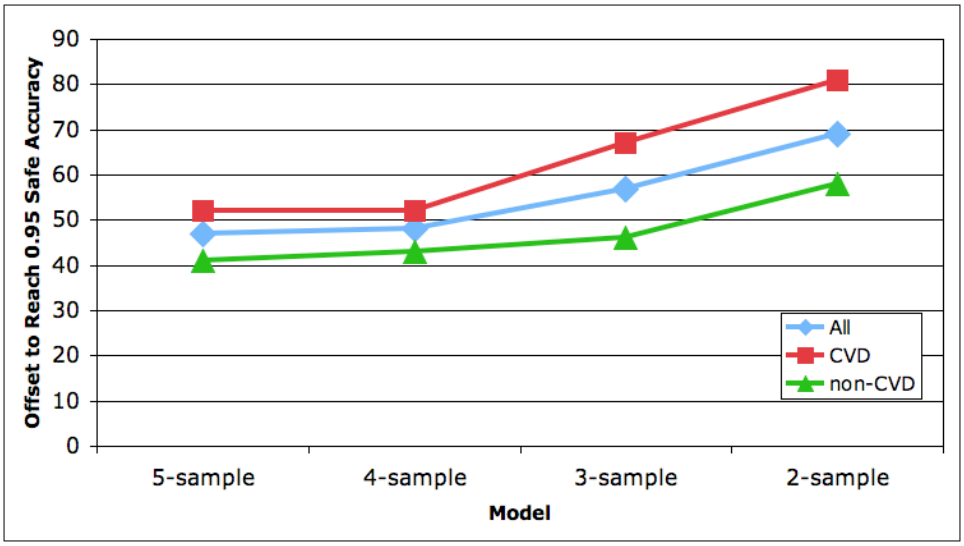


Figure 4.6: Accuracy (mean offset needed to reach .95 A_{safe}), by model granularity and participant group.

Reduction in Available Colour Space Through Overestimation

In order to find a set of differentiable colours, discrimination boxes for the colour set can overlap but must not contain any other colour in the set (i.e., two boxes can have overlapping corners without eliminating the differentiability of the two colours represented by the boxes. Over-estimation 'uses up' more colours than necessary in order to avoid false positive errors, but the number of colours available to a recolouring tool depends on several factors. First, in a two-colour situation, there is little problem, since even with a limit offset of 100 on each channel, there should still be a large colour space remaining within the cube after avoiding the input colour's limit box.

If we had a perfect predictor (i.e., no over-estimation required), 0.6% of all possible RGB colours would be removed on average. Using the 4-sample model, with limit offset to maintain 95% safe accuracy, 14.5% of all possible RGB colours are eliminated. Even though the model eliminates many more colours than necessarily required, it still leaves a large set of possible colours to choose from.

The number of available colours even with a very conservative model is large enough to deal with most colour tasks. For example, even if only three values per channel can be used, this still provides 27 colours, more than enough for the seven maximum that is suggested for

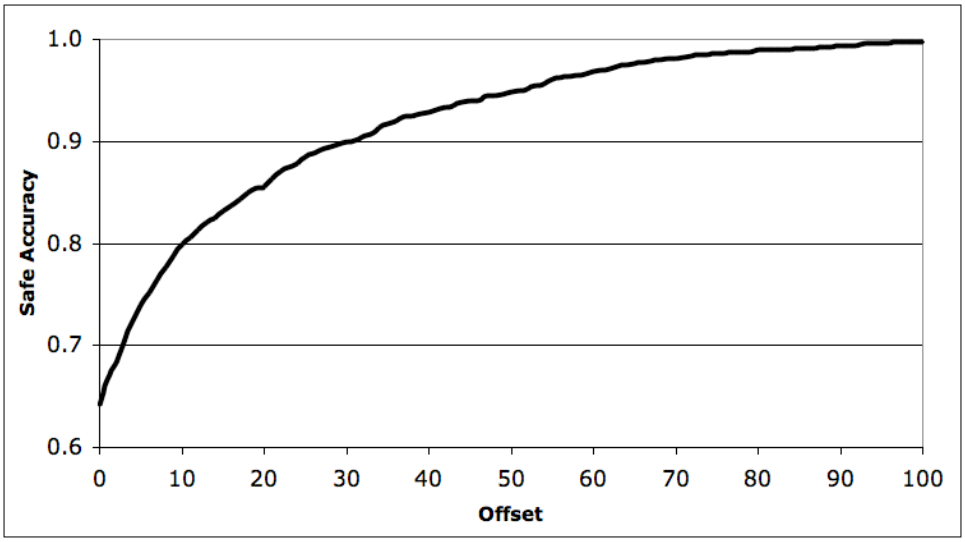


Figure 4.7: Offset amount versus safe accuracy (4-sample ICD-1 model).

categorical encoding [53].

If needed, it is also possible to reduce the degree of over-estimation – this allows more false positives, but preserves a larger colour space for a recolouring algorithm. However, this is unlikely to be a problem except in extreme cases, since even large over-estimations of limits still leaves a very large number of available colours.

An extension to this question is the issue of co-differentiability for larger sets of colours, which is discussed in more detail later in this chapter. Co-differentiability is a constraint on sets of colours that requires that every colour is differentiable from every other colour in the set.

Robustness of the Model

For the evaluation above, the limit offset needed to bring the model's predictions to 95% safe accuracy was calculated. The evaluation compared this offset value with the offset calculated using the standard test set. By testing that model against the robustness data (Part 3 of the study), it can be determined whether the model's accuracy degrades when environmental conditions change by moderate amounts. Figure 4.8 shows the results of this analysis, using the 4-sample model.

The different environmental factors had different effects on the model's accuracy, and CVD and non-CVD groups also had different results. There are both increases and decreases

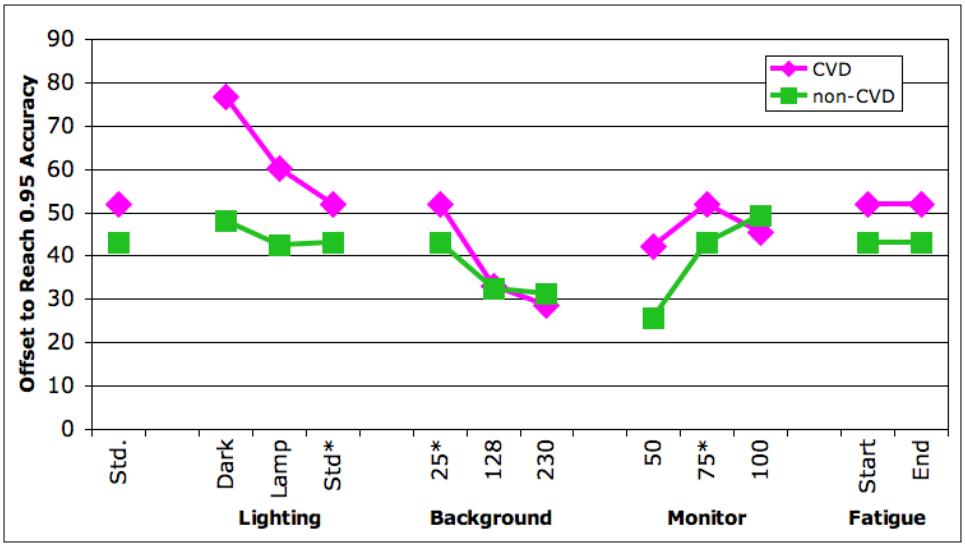


Figure 4.8: Comparison of offset under different environmental conditions (* indicates the value used in the standard test).

to the offset value: increases mean that the model is less accurate in these situations and decreases mean that the model is more accurate. The size of the changes is not dramatic, meaning that the model is not overly sensitive to small changes in the environment. Adding 30 to the limit offset (for non-CVD users) would handle all of the environmental changes that were tested.

Results for users with CVD differ from non-CVD users. There are several potential reasons for this, including the fact that the participants with CVD had several types and severities of colour deficiency.

4.4 Discussion of ICD-1 Evaluation

The evaluation showed that ICD-1 automatically incorporates specific characteristics of the user and his/her local environment, and provides good performance both for individuals with CVD and those with typical colour vision (in addition, the tuneable limit offset allows different balances between false positive rate and colour availability). The study also demonstrated that ICD-1 is robust in the face of moderate environmental change (such as ambient lighting level or background brightness).

ICD-1 can be used with existing recolouring tools without requiring major changes to

those systems. Where a recolouring tool would previously query a simulation module to determine whether colours are likely to be confused, the system could now use the predictive functionality of our model instead. To make a prediction, the recolouring tool would provide the model with two colours, and the model would provide a prediction about their differentiability for the user represented by the model (as per the `areDifferentiable` function described in Chapter 3. These capabilities should also allow the ICD model to be used with a wide variety of other systems, providing them with the benefit of individualized modelling.

Some tasks (such as matching the colours in a bar chart with colours in the chart legend) require the entire colour set to be *co-differentiable* – every colour is differentiable from every other colour. ICD-1 can also be used for this situation, using a process that 'packs' the limit boxes for successive colours into the colour space. The general algorithm, presented in Algorithm 2 below, specifies the process:

1. Choose one colour as the starting colour
2. Add the starting colour's limit box to a do-not-use region
3. For each additional colour that must be co-differentiable:
   (a) Choose a colour just outside the do-not-use region
   (b) If the do-not-use region equals the colour cube, then fail
   (c) Add that colour's limit box to the do-not-use region

Algorithm 2: Algorithm for determine a set of co-differentiable colours.

This process can also be used to determine the maximum number of colours that are available for a single task. For example, if a user's upper differentiability limit is always a value of 50, then 125 different (and co-differentiable) colours can be used for a particular task. Also note that different tasks can re-use the same colour space. For example, if two colours are used to represent 'visited' and 'unvisited' links in a web browser, the same colours could be reused for a bar chart in the same website, since users will be able to separate the colours based on their context.

When the model is used in a recolouring tool, the tuneable offset becomes particularly valuable, as it can be used to choose colours that will maximize the likelihood of differentiability. For example, if a chart image requires three co-differentiable colours, the model can inform an algorithm that maximizes the distance between these colours based on the user's perception. In addition, the model can report exactly how accurate its choices are, since the distances between the colours can be used to determine the probability of false positive (exactly as the limit offset is used in Figure 4.7). In situations where more colours are required than what can be provided with the specified safe accuracy, the model can report the actual probability of false positive errors based on its attempt to maximize distance between the colours.

4.4.1 Limitations of ICD-1

During the development and evaluation of ICD-1, two major limitations were identified: length of calibration, and after-image effects of the calibration task.

Length of Calibration. Utilizing the 4-samples-per-channel model determined above, the calibration time for ICD-1 can be calculated. To gather four samples per channel, 64 colours distributed throughout the RGB colour cube are needed. Three differentiation limits are needed for each colour, resulting in the calibration collecting 192 differentiation limits. Using the estimates of ten presentations to gather a single differentiation limit and one second per presentation used above, this results in a calibration procedure that requires 1920 seconds (32 minutes) to perform. If only a single calibration was ever required by the ICD, then a one-time 32 minute calibration may be tolerable for a user. Unfortunately, calibrations are required whenever the situation or internal state of the user changes (although the above results showed that fatigue had little effect on colour differentiation abilities). As a result, calibrations are both long and possibly frequent, limiting the real-world applicability of the ICD to recolouring tools.

Calibration Task. Many participants in the user study commented on eye strain and fatigue related to the calibration task of analyzing the rectangular grid of circles and reporting whether the colours presented were differentiable or not. The repetitive presentation of colours in the grid of circles has the effect of inducing an after-image for the user (the

previous presentation would temporarily influence how a user sees the current presentation). As a result, it was observed that some participants reported that colours were 'differentiable', when they were not. Upon examination of the study data, it was found that this was indeed the case; sometimes participants reported identical colours as being differentiable. This results in the calibration procedure not truly identifying a user's colour differentiation abilities, but incorporating some level of error. This is partly due to the repetitive (and long) nature of the ICD-1 calibration, but it is also partly due to the judgement nature of the calibration. Participants simply reported whether they could see two different colours or not. When after-images occurred, the participant was truthfully reporting that they 'saw' two colours, even though the two colours were identical.

In the next chapter, these two limitations of ICD-1 will be addressed. First, the RGB colour space is abandoned in favour of a perceptually-uniform colour space which results in a substantial reduction in the number of measurements needed during calibration (thereby shortening the calibration process). Second, the calibration task itself is reinvented to incorporate a performance-based task, helping alleviate the problems associated with after-images.

Chapter 5

Reducing ICD Calibration Time

In this chapter, revisions to the feasibility implementation of the ICD (Chapter 4), are presented. These revisions are focussed on the main challenge real-world use presents for ICD-based recolouring – the lengthy calibration procedure of ICD-1. Considering that real-world use of an ICD-based recolouring tool will typically require frequent calibrations, reducing the time to perform a calibration is key to the real-world applicability of the solution presented in this dissertation. The chapter begins with an introduction to the problem, then presents revisions to the modelling and calibration components resulting in a new model version (ICD-2). ICD-2 is then compared to ICD-1 in a comparative user study that examined calibration time and prediction accuracy. The chapter concludes with a brief discussion of the implications of the described revisions. [1]

5.1 Problem: Time-Consuming ICD-1 Calibration

Although ICD-1 produced good results in the above evaluation, it has some limitations that prevent it from being useful in real-world use. The most substantial limit is the amount of time needed for calibration – this issue is considered here, and a way to dramatically reduce the calibration time requirement is presented.

One main reason for the long calibration time of ICD-1 is the use of the RGB colour space. RGB is not a perceptually-uniform colour space [130] – that is, the perception of difference for one pair of colours that has a particular distance in RGB space will not necessarily be the same as for two other colours that have exactly the same distance. This means

[1] Portions of this chapter were published as: Flatla, D.R., Gutwin, C. (2011) Improving Calibration Time and Accuracy for Situation-Specific Models of Color Differentiation. In *ASSETS '11: The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, 195–202.

that the differentiation limits gathered by ICD-1 do not generalize well from the sample colour; as a result, ICD-1 needs to measure differentiation limits for many colours (64) uniformly spread throughout the RGB colour cube. The binary search method of determining a differentiation limit (see Section 4.2.2 above) requires approximately ten presentations before a single differentiation limit is identified. ICD-1 needs to determine one differentiation limit per channel per calibration colour; for 64 sample colours, this requires 10x3x64=1920 presentations. If each decision takes one second, calibration requires a total of 32 minutes.

The issue of calibration time goes beyond the simple total time for the procedure, however, since the situation-specific approach requires that a new calibration be carried out for each new colour-perception situation. Although the evaluation of ICD-1 showed that the model is reasonably robust when environmental factors undergo small changes, there will still be several large-scale shifts that will require recalibration (e.g., different models might be required for indoor versus outdoor viewing, or for different monitors that have very different characteristics). In addition, some users may simply want the model to be as accurate as possible even for smaller environmental changes. The need for multiple models means that the time needed for each calibration must be multiplied by the number of calibrations performed. If calibration time is above a threshold (e.g., a few minutes), users will be reluctant to disrupt their existing tasks to re-calibrate.

5.2 Model and Calibration Revisions to Reduce Calibration Time

To address the problem described above (lengthy calibration), ICD-2 moves from relying on the RGB colour space to a perceptually-uniform colour space: CIE L*u*v*. This colour space will allow calibration measurements (differentiation limits) for a single sample colour to generalize to other colours in this space. This generalizability of calibration data will enable a sizeable reduction in calibration time as the number of colours to measure differentiation limits for is reduced from the 64 sample colours of ICD-1 to a single sample colour for ICD-2.

5.2.1 Calibrating in a Perceptually-Uniform Colour Space

By using a perceptually-uniform colour space, the number of sample colours tested during calibration can be reduced, because the limits for the sample colours will better generalize to other colours. To accomplish this, ICD-2 uses CIE L*u*v* [112, 153, 130], a colour space that is perceptually uniform and that separates the description of a colour into a luminance axis (L* – ranging from 0 (black) to 100 (white)) and two chromaticity coordinates (u* and v* – centred at 0, range of ±100). This allows colours of equal luminance to be found simply by holding L* constant and varying u* and v*. When u* and v* equal 0, the colour is achromatic (black, grey, or white). Five equal luminance (isoluminant) slices of L*u*v* colour space are shown in Figure 5.1.

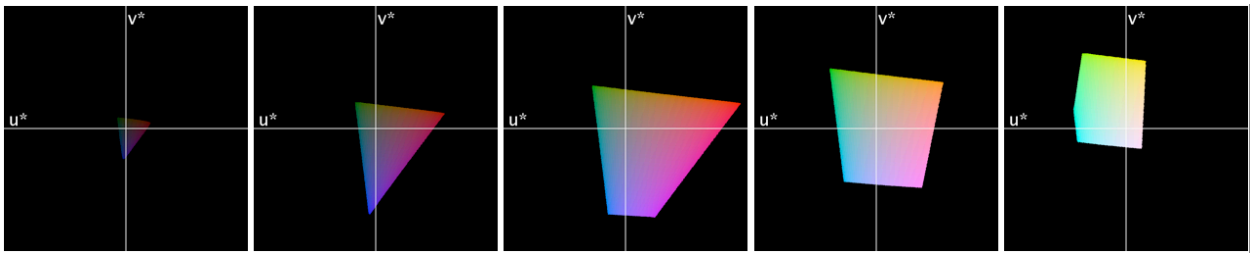


Figure 5.1: Five isoluminant slices of CIE L*u*v* colour space showing the u* and v* axes. From left to right: L*=10, L*=30, L*=50, L*=70, L*=90.

In the L*u*v* colour space, the differentiation abilities of an individual can be described well using a discrimination ellipse when only colours of identical luminance are considered [118]. For a given colour, a discrimination ellipse can be found that surrounds the colour. Those colours outside the ellipse are differentiable from the original colour and those inside are not differentiable from the original colour. To extend this to a three-dimensional shape, discrimination ellipsoids [114] are used in the ICD-2 model to replace the discrimination boxes from ICD-1. The discrimination ellipsoid is defined using the major and minor axes of the discrimination ellipse and two points above and below the equal luminance plane to form the third axis of the ellipsoid.

Since images are represented using the RGB space in digital environments, the use of L*u*v* requires translation between these two colour spaces. ICD-2 uses the sRGB transform provided by `http://www.brucelindbloom.com/` assuming a gamma of 2.2 and a D65 white

point to move between RGB and CIE XYZ, and then to move between XYZ and L*u*v*. As a result, transformations from RGB to L*u*v* (and back again) can be easily accomplished. Details for these conversions are given in Appendix A.

5.2.2 Finding Discrimination Limits (Ellipses) in ICD-2

To find the discrimination ellipse for a particular colour, six discrimination limits are measured via a calibration procedure. Instead of finding these limits along RGB channel dimensions as in ICD-1, these limits are measured along three dichromatic colour confusion lines [153] within an isoluminant plane. A colour confusion line is defined by a base colour (grey is used here) and a copunctal point. In CIE L*u*v* colour space, these confusion lines are straight lines between the base colour and the copunctal point, but in CIE L*a*b* colour space, I determined empirically that these confusion lines are curves. As a result, the geometry for finding discrimination limits along each confusion line, and for determining the resulting discrimination ellipse, is simpler in CIE L*u*v* than in CIE L*a*b*. For this reason, CIE L*u*v* was selected as the perceptually-uniform colour space for ICD-2.

Copunctal points are the colour space representation of the contribution of a single cone type to colour perception. When a person is entirely missing a cone type (as in dichromacy), he/she is unable to distinguish colours along any line that intersects his/her respective copunctal point. As a result, a colour confusion line represents the set of colours that are not differentiable for someone missing the cone type represented by the copunctal point (see Figures 5.2 and 5.3).

Each confusion line gives two differentiation limits, one moving from ICD-2's grey base colour to the copunctal point, and one moving from the base colour away from the copunctal point (Figure 5.4), resulting in six differentiation limits (two for each copunctal point). Each differentiation limit measures the user's ability to differentiate colours along each confusion line from a central achromatic grey. These six limits are then used to generate the best-fit ellipse using approaches outlined in earlier work [33, 49]. Although only four points are required to define the major and minor axes of an ellipse, six points are necessary here to determine the angular orientation of the ellipse. The major axis of this ellipse will be aligned with the protan, deutan, or tritan confusion line, and the six points are used to determine

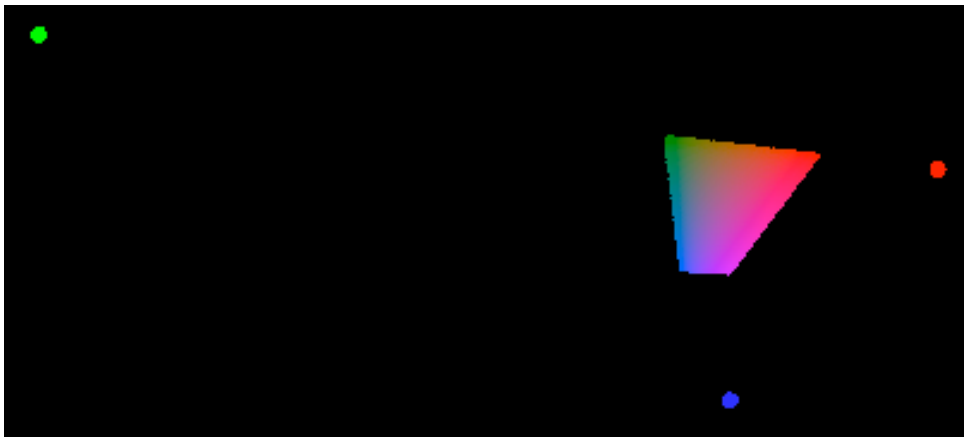


Figure 5.2: Copunctal points for missing long-wavelength sensitive cone (red), medium-wavelength sensitive cone (green), and short-wavelength sensitive cone (blue) for the L*=50 isoluminant plane.

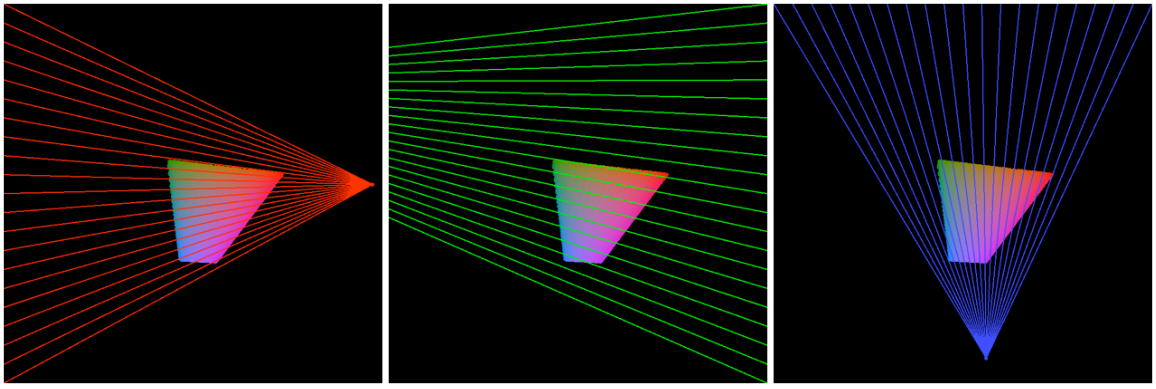


Figure 5.3: Confusion lines for missing long-wavelength sensitive cone (red), medium-wavelength sensitive cone (green), and short-wavelength sensitive cone (blue) for the L*=50 isoluminant plane.

the correct confusion line. Once the best-fit ellipse if generated, its major and minor axes half lengths are used to define the horizontal cross-section of the ellipsoid.

To find the two points above and below the luminance plane, two additional discrimination limits are measured via the calibration procedure. These limits correspond to the amount of luminance that needs to be added and subtracted from the base colour in order for the user to perceive a difference. These values are used to define the third axis of the discrimination ellipsoid.

An ellipsoid can be described using the following formula:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{5.1}$$

Figure 5.4: Confusion line colours for individuals with tritanopia (blue-yellow dichromatic CVD) using the base colour for the ICD-2 model: a) confusion colours toward the tritan copunctal point, b) tritan simulation of (a), c) confusion colours away from the tritan copunctal point, d) tritan simulation of (c).

where a is the best-fit ellipse major axis half length, b is the best-fit ellipse minor axis half length, and c is the amount of luminance that is added or subtracted, as described above. This formula is used in ICD-2 to internally represent the discrimination ellipsoid.

5.2.3 Calibration Procedure for ICD-2

In spite of the precautions to prevent after-image effects in the calibration technique for ICD-1 (see Section 4.2.2), many participants reported trouble telling when two colours were genuinely differentiable or not. Their difficulties were also reflected in the empirical data when participants reported that two identical colours were differentiable. Partly responsible is the length of calibration, which is being addressed in this section, but also partly responsible is the judgment nature of the calibration task (simply reporting whether the colours are differentiable or not). To address this, a performance-based calibration task was developed for ICD-2.

To calibrate the ICD-2 model, eight discrimination limits are needed for a single base colour. These limits are found in a similar binary search manner as the old model, but the technique has been modified so that the user no longer provides a judgment about the differentiability of the two colours presented, but rather performs a task. If the user can perform the task, then it is interpreted that the user can see the difference between the two colours. If they cannot do the task, then it is interpreted as the user not being able to

differentiate between the two colours. For ICD-2, I use a neutral grey as the base colour (L*u*v* = 50,0,0). These L*u*v* coordinates map to RGB colour (118,118,118) using the sRGB transform described above and in Appendix A.

This performance task involves the user identifying the orientation of a hollow circle with 1/8th of its perimeter missing (Figure 5.5). The circle is presented to the user, and if they can identify the location of the gap, they press a correspondingly labeled key on the numeric keypad. If they see no gap, the user presses the space bar. A binary search approach is used to find the differentiation limit, where the difference between the circle and the background is increased when the user sees no difference, and decreased when they see a difference.

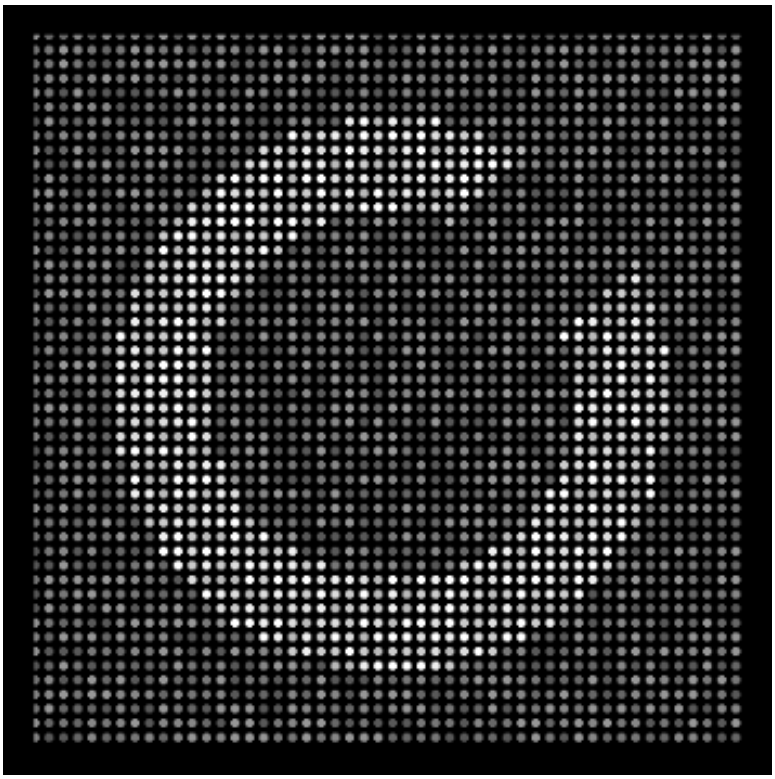


Figure 5.5: Screen presented during ICD-2 model calibration. User presses numeric keypad key that matches the rotational orientation of the gapped circle, or space if he/she sees nothing.

To facilitate the performance task, the presentation of the colours to the user has been modified to approximate the approach used in [11, 118] to determine discrimination ellipses, which utilizes a similar gapped-circle performance task and random luminance noise (described below). A 400x400 pixel region on a black screen is presented to the user. This

region is filled with a regular pattern of small (4-pixel diameter) circles, with black between the circles. The gapped circle introduced above is superimposed on this background of small circles, such that the background is the base colour, and the gapped circle is a colour along the confusion line (or luminance line) for which a discrimination limit is desired (Figure 5.5). The numeric keypad of the keyboard was modified with labels such that the labels matched the possible orientations of the gapped circle.

When two colours are placed directly adjacent to each other, any differences in luminance between the two colours results in the user seeing a difference between the two colours, even though this difference may go away as soon as a small gap is introduced between the two colours. To offset the effect of luminance contrast, temporal random luminance noise [11] was applied to the entire presentation (background and gapped circle). This noise produces colours with identical u*v* coordinates (chromaticity), but varying L* values (luminance). The black space between the small circles further reduces the effects of luminance contrast.

When a differentiation limit has been identified, its Euclidean distance from the base colour in L*u*v* space is recorded. This gives a distance along each confusion line (6 measures) and distance in luminance above and below the base colour (2 measures).

5.2.4 Making Predictions with ICD-2

To answer an `areDifferentiable(Colour c1, Colour c2)` query, the ICD-2 model first converts both colours to L*u*v* colour space (using the sRGB transform mentioned above), and determines which is closer to the base colour using Euclidean distance. The closest colour is chosen because the discrimination limits for the base colour are known, and if the L*u*v* colour space is not perceptually uniform for certain types of CVD, then the discrimination limits for the colour that is closest to the base colour should be more similar to the base colour differentiation limits. The colour parameter closest to the base colour is called the primary colour, the other colour parameter is called the secondary colour. To determine an appropriate discrimination ellipsoid, the differentiation ellipsoid defined by the calibration data needs to be adapted to the primary colour.

To save computation, before the discrimination ellipsoid is adapted, a luminance comparison is performed. Using the luminance thresholds determined during the calibration for

the base colour, luminance bounds for the primary colour are determined. If the luminance of the secondary colour is outside of these bounds, the query returns `true` (the colours are differentiable). If the luminance of the secondary colour falls within these bounds, then the adaptation is performed.

To adapt the discrimination ellipsoid of the base colour to the primary colour, the confusion lines for the primary colour are found (between the primary colour and each copunctal point). As confusion lines are defined by the copunctal point and another point, they are not rotationally invariant as different 'other' points are selected. Once the colour confusion lines for the primary colour are found, the algorithm walks along each confusion line away from the primary colour until the L*u*v* distance just exceeds the differentiation limit L*u*v* distance found during the calibration. The L*u*v* coordinates for the colour along each direction on each confusion line are then used to specify six points in the u*v* plane for the primary colour. These six points are used to find the best-fit ellipse. This process of walking out from the a colour is illustrated in Figure 5.6 for two different colours.

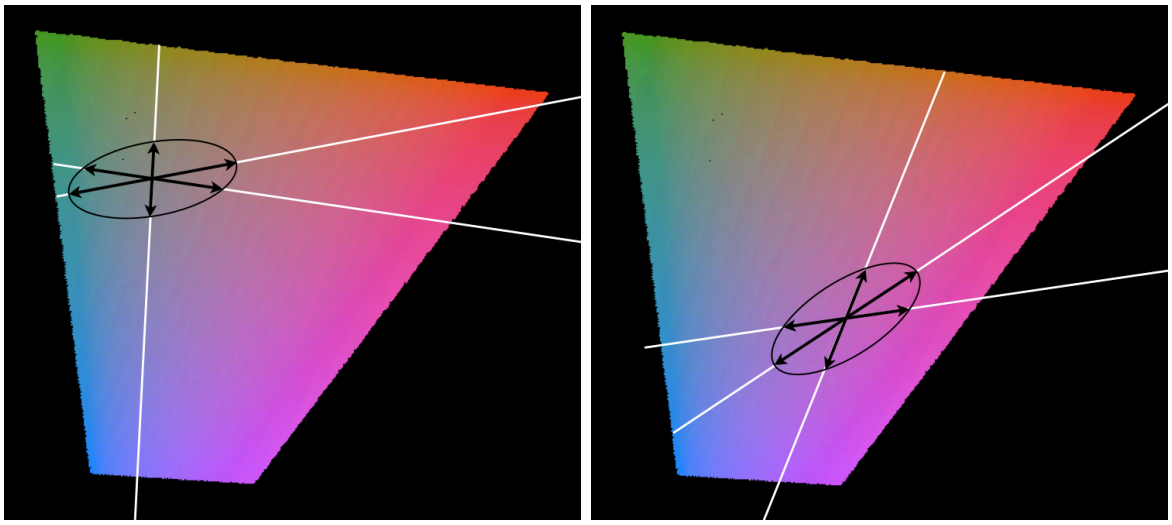


Figure 5.6: Left: six 'walks' along three confusion lines to find the best-fit ellipse for the base grey. Right: six 'walks' along the three confusion lines to find the best-fit ellipse for a purple colour.

Using the luminance of the secondary colour, the best-fit ellipse is resized to be the discrimination ellipse for the primary colour, but at the secondary colour's luminance, by using modifications of the formula for an ellipsoid given above to find the adjusted half

major and half minor axis lengths for the resized ellipse:

$$a' = \sqrt{a^2 - \frac{a^2 \times (L_S - L_P)^2}{c^2}} \tag{5.2}$$

$$b' = \sqrt{b^2 - \frac{b^2 \times (L_S - L_P)^2}{c^2}} \tag{5.3}$$

where a' is the resized major axis half length, b' is the resized minor axis half length, L_P and L_S are the luminance values for the primary and secondary colours, respectively.

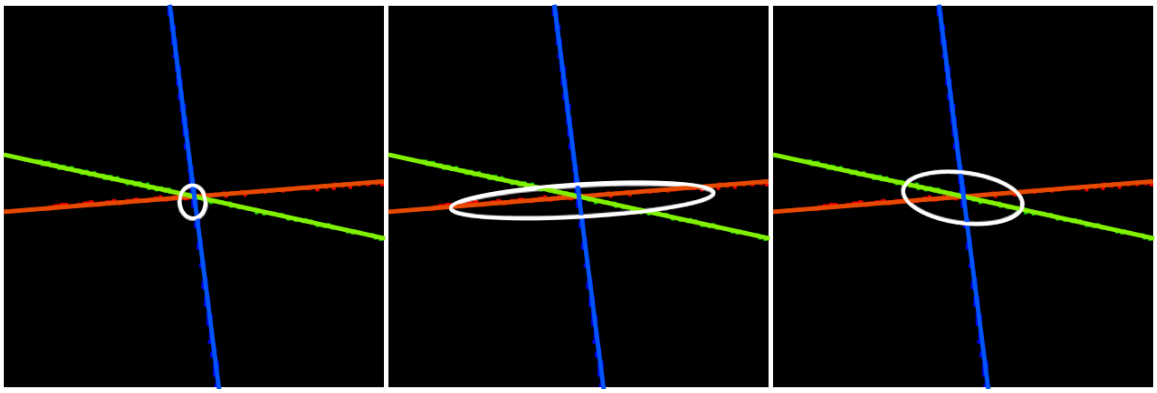


Figure 5.7: Three examples of the ellipses found by ICD-2: typical colour vision (left), protan CVD (middle), deutan CVD (right). Coloured lines indicate colour confusion lines (protan = red; deutan = green; tritan = blue).

Once the resized ellipse is found, its center, dimensions, and orientation are used to transform a unit Java `Ellipse2D.Double` object. When transformed, this object is defined by a `Path2D.Double` object, which provides a `contains(Point2D.Double)` method, which determines if a given point is within the ellipse. The u*v* coordinates for the secondary colour are packaged into a `Point2D.Double` object and passed as a parameter to the 'contains' method. If the point is in the ellipse, then the query returns `false` (not differentiable), otherwise it returns `true` (differentiable).

5.3 Evaluation of ICD-2: Calibration Time and Model Accuracy

To evaluate the changes implemented in ICD-2, the ICD-2 model was compared to ICD-1 in an empirical user study. There were two main goals of the evaluation – first, to confirm that

calibration of ICD-2 is in fact faster than the feasibility model (and to determine the actual reduction in time), and second, to determine ICD-2's accuracy compared with the existing approach.

5.3.1 Study Methods

To compare the ICD-2 model with ICD-1, a user study with 16 male participants (mean age 33.8 years) was conducted. Of these 16 participants, eight self-identified as having CVD (mean 39.0 years) and eight self-identified as not having CVD (mean 28.7 years). As both ICD-1 and ICD-2 are general models of colour differentiation, no tests were performed to assess the type or severity of participant CVD. I constructed a Java application using the Processing libraries for displaying visual content to the screen (`www.processing.org`). The study ran in a single location on a Windows 7 machine using a 20-inch 1600x1200 Dell 2001fp monitor.

Participant instructions are shown in Figure E.2. Each rotated-circle stimulus was shown for an unlimited period of time, which resulted in some pilot study participants taking too long to provide responses, with no clear improvement in response accuracy. To counteract this, the instructions included a suggestion that users spend no longer than three to four seconds on each stimulus. Participants were also encouraged to be careful while entering their choices to help balance speed and accuracy.

During the study, participants performed two tasks. The first task collected calibration data for generating both models. In the second task, 'ground truth' responses were collected from the participant to evaluate each generated model.

Calibration Task

As the study was designed to compare the models to each other, identical calibration procedures were used for ICD-1 and ICD-2. Due to the performance-based nature of the ICD-2 gapped circle calibration technique described above, this was chosen as the calibration method for both models. The procedure was modified for ICD-1 in order to gather increasing and decreasing differentiation limits on RGB channels.

As ICD-1 calibration is time consuming, a reduced set of ICD-1 model calibration points was gathered to reduce study run time. Nine points were chosen in RGB space to approximate the uniform spread of the 64 calibration points in the true calibration of ICD-1. These were (118,118,118), and eight additional colours, one halfway along each ray from this start colour to the eight corners of the RGB colour cube. This gave the following nine colours: grey, black, green, yellow, red, blue, purple, cyan, and white. For each of these, six differentiation limits (three upper and three lower) were collected from the user, for a total of 54 differentiation limits. To calibrate ICD-2, the eight differentiation limits around (118,118,118) described above were collected. This gave a total of 62 differentiation limits.

As described in Section 4.2.4, ICD-1 utilizes linear interpolation of the differentiation limits for a large sample (64) of RGB colours to perform predictions. As was described in Section 4.3.2, fewer samples of RGB colours causes ICD-1 to be less accurate in its predictions. The calibration for ICD-1 in this study uses only nine RGB colours, so to ensure that the reduction in the number of calibration colours did not put ICD-1 at a disadvantage, the evaluation task for this study was designed to not request differentiability predictions for any colour besides the nine calibration colours (i.e., the parameters for `areDifferentiable` were one of the nine colours and any other RGB colour). As a result, no linear interpolation was necessary because the ICD-1 discrimination box for each of the nine calibration colours was determined by the calibration procedure. This has the effect of slightly favouring ICD-1, providing it the best possible case of not having to perform interpolation to estimate any differentiation limits.

The order of the 62 limits was randomized and presented to the participant sequentially. When the participant supplied a response (either the space bar for 'no circle visible,' or the appropriate numeric keypad key) the difference between the background and the circle was adjusted accordingly and the limit was reinserted into the sequence. If an incorrect numeric key was pressed, it was interpreted the same as pressing the space bar. Once the participant had given a response for each of the 62 limits, the order was shuffled and presented sequentially to the participant again. This was repeated until the binary search for each limit converged on a single value. For ICD-1, this value was reported as a raw RGB channel difference. For ICD-2, this value was converted into its equivalent L*u*v* Euclidean

distance from the base colour. The entire calibration required about 400 presentations to the participant, taking approximately 30 minutes. The participant could take a break at any time, and was encouraged to take at most 3-4 seconds per presentation.

The total time to collect each differentiation limit was recorded as well. With this time data, the total time to gather the 54 ICD-1 model differentiation limits, and the total time to gather the eight ICD-2 calibration differentiation limits was measured. As the original old model requires 192 calibration limits, the actual time to collect the 54 limits was scaled up ($scaledTime = 192 \times \frac{measuredTime}{54}$) to reflect the ICD-1 calibration time using 64 sample RGB colours.

Evaluation Task

Once the calibration data was collected, the participant took a five minute break to rest their eyes. Once finished, the participant then performed the evaluation test. At the beginning of this task, the calibration data from the first session was used to generate both ICD-1 and ICD-2 models. These models were then used to generate evaluation trials as described below.

Any two randomly-selected colours have a high probability of being differentiable. It was desirable to use evaluation data that would provide a more uniform chance of each model predicting that the colours would be differentiable or not. To accomplish this, the models based on the calibration data for each participant were used to generate the evaluation trials for that participant. Using the nine RGB colours mentioned above (grey, black, green, yellow, red, blue, purple, cyan, white), each model was used to generate two sets of 15 colours – one that the model predicts as being differentiable from the supplied colour and one that the model predicts as being not differentiable from the supplied colour. To accomplish this, colours were uniformly randomly selected from a volume twice as large as the ICD-1 discrimination box or the ICD-2 discrimination ellipsoid. Each randomly-selected colour was then predicted as differentiable or not differentiable by the corresponding model and added to the appropriate set (differentiable or not differentiable). The sets were returned when they were both full. This resulted in 15 colours in a set x 2 sets x 2 models x 9 base colours = 540 trials for the evaluation session. These trials were randomly presented to the participant using the same test procedure (with the gapped circle) as the ICD-2 calibration.

If the participant correctly identified the orientation of the gapped circle, then the colours were recorded as 'differentiable'; otherwise 'not differentiable' was recorded. This was used to establish a 'ground truth' set of colour comparisons. For each of these responses, each models' differentiability prediction was also recorded.

5.3.2 Study Design

The evaluation used a repeated-measures factorial design with two factors: model type (ICD-1 or ICD-2) and CVD presence (typical colour vision or CVD). The CVD-presence factor was used only to check for interactions in the other analyses.

Four dependent variables were recorded by the system. Calibration time was gathered from the calibration task as described above. The remaining three variables were calculated from the raw correct/incorrect data gathered from comparing the predictions to ground truth: overall accuracy (number of correct predictions over total trials), false positive rate (proportion of predictions that incorrectly suggested that colours were differentiable), and false negative rate (proportion of predictions that incorrectly suggested that colours were not differentiable).

The analysis used repeated-measures ANOVAs to test the effects of model type on these four dependent variables, and to look for interactions with CVD presence.

5.3.3 Results: Calibration Time

The time needed to carry out the entire calibration for both ICD-1 and ICD-2 was recorded, and the ICD-1 time was scaled to reflect the true calibration procedure for the feasibility model (from 54 to 192 differentiation limits). As shown in Figure 5.8, calibration time for ICD-2 (mean of 2.14 minutes) is dramatically lower than for ICD-1 (mean 51.97 minutes). Not surprisingly, the effect of model type is significant ($F_{1,14}$=123.81, $p<0.001$); there was no interaction with presence of CVD ($F_{1,14}$=0.077, $p=0.79$). The 24-fold improvement is proportional to the reduction in the number of calibration trials (from 192 to 8).

The reason for the reduced time to calibrate ICD-2 is simple – using the perceptually-uniform L*u*v* colour space and basing the discrimination volume on known colour confusion

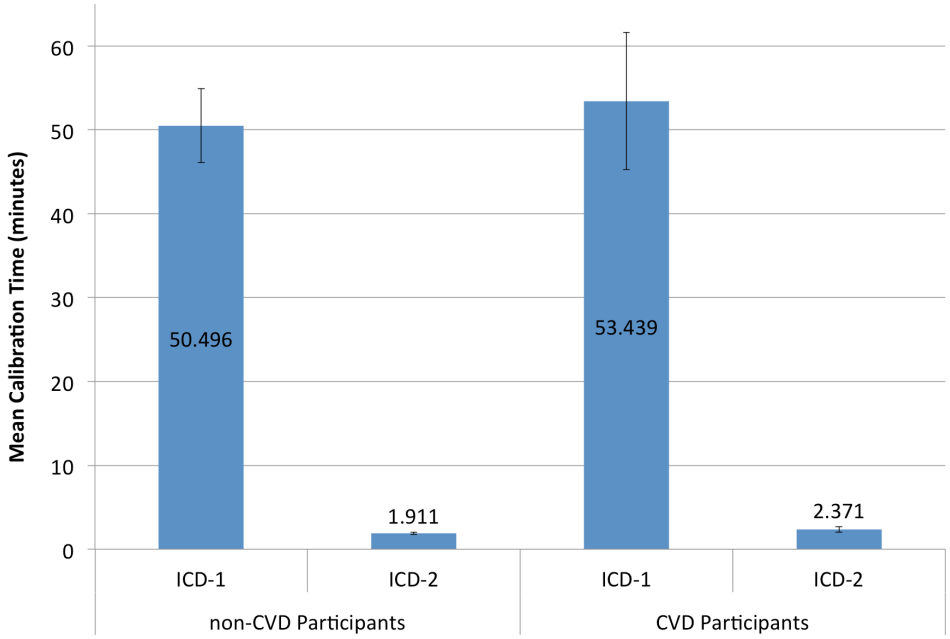


Figure 5.8: ICD-1 and ICD-2 calibration times ±s.e., by CVD presence.

lines means that ICD-2 requires far fewer calibration samples than the old model. These changes show the value of building the model on principles that more completely characterize human colour perception. The short calibration time of ICD-2 (just over two minutes) means that even taking additional samples to further improve accuracy (as discussed below) can be considered.

5.3.4 Results: Model Accuracy

I tested the accuracy of the models' predictions by comparing them to the ground truth of the 540 evaluation trials collected from participants. The overall mean accuracy for ICD-1 was 75.8%, and for ICD-2 was 79.2% (see Figure 5.9). ANOVA showed that model type had a significant main effect on accuracy ($F_{1,14}$=9.38, $p<0.01$), with ICD-2 at approximately 3.4% higher accuracy. There was no interaction with CVD presence ($F_{1,14}$=0.28, $p=0.60$).

The increase in accuracy for ICD-2 compared with ICD-1 was modest (3.4%), so it is difficult to conclusively determine the source of the improvement. However, I believe that the change from a discrimination box (in ICD-1) to a discrimination ellipsoid (in ICD-2) is the main reason for the better performance: previous research has shown that an ellipsoid

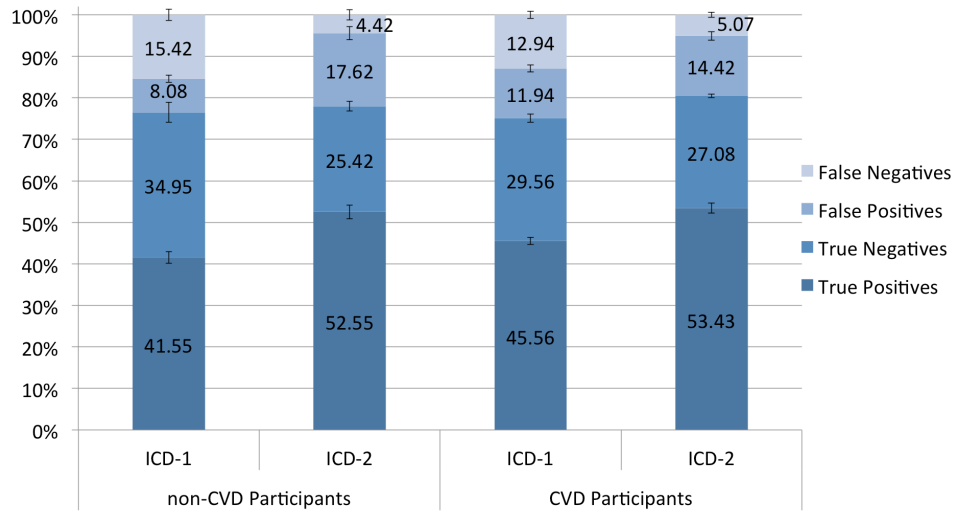


Figure 5.9: Model accuracy ±s.e., by CVD status and model (true positive rate + true negative rate = overall accuracy).

better matches the way that humans perceive colour [114] and the way that individuals with CVD have difficulty with differentiation [16].

5.3.5 Results: Error Rates

The mean false positive rate for ICD-1 was 10.01%, and for ICD-2 was 16.02%. ANOVA shows that this difference is significant ($F_{1,14}$=64.19, p<0.001); ICD-2 had approximately 6% more false-positive errors. In addition, there was a significant interaction between model type and CVD presence ($F_{1,14}$=22.16, p<0.001). As shown in Figure 5.9, the false-positive difference between the old and new models is larger for participants with typical colour vision than for participants with CVD. I discuss the implications of these differences below. Mean false negative rates were 14.18% for ICD-1 and 4.75% for ICD-2. This difference is also significant ($F_{1,14}$=154.17, p<0.001). There was no interaction with CVD presence ($F_{1,14}$=4.23, p=0.059), although the false-negative difference between ICD-1 and ICD-2 is larger for participants with typical colour vision than for participants with CVD.

False positive errors are more serious for recolouring tools than false negatives, since false positives lead to situations where the recolouring algorithm is not able to identify problem colours (because two colours that are actually not differentiable are predicted to be). The seriousness of these situations is compounded further if the recolouring tool also

finds replacement colours that are not in fact differentiable by the user.

The higher false positive rate of ICD-2 is directly caused by the model ellipsoid being smaller than it should be; the reasons why the model chose too-small ellipsoids, however, are not clear. One possibility is that the step between colours along the colour confusion lines is too large. As these colours were pre-computed (to save processing time), it is possible that the chosen step was too large. This could result in the calibration returning a differentiation limit that was on the 'not differentiable' side of a step, even though the true differentiation limit was somewhere in the middle of the step. This would result in unnecessarily small differentiation limits, leading to a small ellipsoid.

A step size that is too large would also explain the difference in false positives between CVD and non-CVD participants. The error introduced by this problem would have an additive (not multiplicative) effect on the volume of the resulting model. Smaller volumes (e.g., for non-CVD users) would be more greatly affected by reducing their differentiation limits by a fixed amount than larger volumes (e.g., for CVD users).

5.4 Discussion of ICD-2 evaluation

The study provides three main results about the performance of ICD-2:

1. Calibration of ICD-2 is dramatically faster than the feasibility ICD-1 model, requiring 1/24 the time of the old calibration;

2. ICD-2 is significantly more accurate than the old model, with a 3.4% improvement in overall accuracy;

3. ICD-2 does show a higher rate of false positives (although primarily for participants with normal colour vision).

Here I consider three issues that were raised in the study: how the false positives can be addressed in a real recolouring tool, how the ellipsoid approach will generalize, how the ellipsoid models can be improved, and the time to make a colour differentiability prediction in ICD-2.

ICD-2 Limit Offset. The problem of false positives can be dealt with through a 'limit offset' factor that arbitrarily increases the size of the ellipsoid (similar to ICD-1). The size of the limit offset will vary from user to user, but a fixed value can be determined and applied to all tuned models when each model is created. An exploration of appropriate limit offset values for a revised version of the ICD (ICD-3) is given in Section 6.2.5. The overall accuracy of the situation-specific modelling approach used in the individualized model of colour differentiation is such that even a liberal offset value will not greatly reduce the number of colours available to a recolouring algorithm (e.g., the discrimination ellipsoid can be doubled in size without overly restricting the set of possible replacement colours for recolouring tools). Limit offsets for ICD-2 can be implemented in a very similar manner as they were for ICD-1. As the differentiation limits are encoded as L*u*v* Euclidean distance between the base colour and the differentiation point along each colour confusion line, simply adding or subtracting a fixed value (the limit offset) from each distance value will result in a larger or smaller discrimination ellipsoid, achieving the desired limit offset effect.

Generalizing to Other Situations and Users. ICD-2 is able to generate different ellipses for different types of users; for example, Figure 5.7 shows ellipses for a user with typical colour vision, a user with protan CVD, and a user with deutan CVD. This variation in generated models (and the associated accuracy results) present a strong argument that this modelling approach is applicable to many individuals with a variety of colour differentiation abilities. To evaluate this further Chapter 7 presents an ICD-based recolouring tool, and compares it to two existing recolouring tools for a variety of situations and user colour differentiation abilities. The evaluation presented above (Section 5.3) did include two individuals with CVD who were older (63 and 67 years old, one diagnosed with cataracts), who both experienced gains in accuracy from ICD-1 to ICD-2 (3% and 2%, respectively). These results suggest that the model will generalize well, at least with regard to internal variations such as age and illness.

Increasing the Number of Differentiation Limits. The three confusion lines introduced above give rise to six differentiation limits defining the discrimination ellipse. To improve the shape, location, and size of the ellipse, additional differentiation limits at different points can be collected. These would be along lines of a different rotational orientation. Each new

line would introduce two additional discrimination limits, so the accuracy of the model can be balanced against calibration time.

Time to Perform Predictions. Although the ICD-2 model introduces a dramatically reduced calibration time and a slight gain in prediction accuracy, these gains come at the cost of time to make a prediction. As discussed above, the ICD-2 model incorporates some modifications to reduce the processing time required to make a colour differentiability prediction (e.g., the luminance bounds checking described in Section 5.2.4). These modifications were implemented because it was recognized early on that the confusion line walking technique to find the discrimination ellipsoid is computationally demanding. As a new ellipsoid is determined with every call to `areDifferentiable`, predictions using ICD-2 are slower than ICD-1 (see Section 6.1). To address this limitation, revisions to the internal representation of the model used in ICD-2 must be made, and are motivated and presented in the next chapter.

Chapter 6

Improving ICD Prediction Time and Power

This chapter presents further improvements to the individualized model of colour differentiation (ICD). As ICD-2 (Chapter 5) significantly improved the calibration technique, the improvements presented in this chapter will focus on the model generated from the existing calibration technique. The model revisions presented in this chapter are:

1. ICD-3: Reduces the time to perform a prediction to improve the responsiveness of recolouring tools.

2. ICD-4: Expands the power of each prediction from the boolean `areDifferentiable` function to a real-valued `howDifferentiable` function to enable recolouring tools to perform constraint-optimization to find replacement colour sets.

This chapter begins with a description of how ICD prediction times are reduced in ICD-3 as well as a justification for why this reduction is valuable in terms of real world recolouring tool performance. The prediction time, accuracy, and limit offset performance of ICD-3 is then evaluated. Following this, the modifications to obtain the real-valued `howDifferentiable` function of ICD-4 are described. The value of the `howDifferentiable` function over the basic `areDifferentiable` function is then illustrated visually. This chapter concludes with a brief overview of the refinements and study results for ICD-3 and ICD-4, briefly discussing their implications. [1]

[1]Portions of this chapter were published as: Flatla, D.R., Gutwin, C. (2012) Situation-Specific Models of Color Differentiation. In *TACCESS: ACM Transactions on Accessible Computing*, vol. 4 no. 3, 13:1–13:44

6.1 ICD-3: Improving ICD Prediction Time

As described in Chapter 5, ICD-2 addresses one of the real-world limitations of the original approach (i.e., that calibrations took a long time). To accomplish this, the calibration technique and internal modelling component were substantially revised to incorporate *discrimination ellipsoids*, and used these ellipsoids to perform predictions. These revisions provided a 24-fold reduction in calibration time, enabled a performance-based calibration technique, and improved the prediction accuracy of the model.

However, the ICD-2 revisions also introduced a substantial time penalty for performing predictions. The feasibility implementation of ICD (ICD-1) uses trilinear interpolation to predict the boolean differentiability of any two colours. This prediction requires calculating a discrimination box in RGB colour space for one colour, then testing whether the other colour was inside or outside of this box. Calculating the discrimination box and testing the other colour are constant-time calculations, independent of the magnitude of the discrimination limits used to define the box. As a result, predictions in ICD-1 take a constant amount of time and are fast (the issue of speed is explored in depth below).

Similar to ICD-1, ICD-2 calculates a discrimination volume (ICD-2 uses an ellipsoid) for one of the prediction colours, and then tests whether the other colour is inside or outside this ellipsoid. However, to determine the discrimination ellipsoid, ICD-2 utilizes a programmatic 'walk' along each of six colour confusion lines to define the bounds of the ellipsoid. As a result, the time to calculate the discrimination ellipsoid is proportional to the severity of the user's CVD; the more severe the CVD, the longer it takes to calculate the discrimination ellipsoid, thereby increasing the time to perform a prediction.

To illustrate the differences in prediction time between ICD-1 and ICD-2, a small study was conducted using the calibration data from the evaluation for ICD-2 (Section 5.3). Average prediction times across one million pairs of random RGB colours for ICD-1 and ICD-2 are presented in Figure 6.1. The mean prediction times for ICD-1 are 0.00028 ms for non-CVD participants and 0.00026 ms for participants with CVD. For ICD-2, the times are 0.06097 ms (non-CVD) and 0.31582 ms (CVD). These differences reveal over a 200-fold increase for non-CVD participants and over a 1200-fold increase for participants with CVD. Note that the

times for CVD participants are substantially greater than non-CVD participants, illustrating the influence of the longer ICD-2 confusion line 'walks' for participants with CVD.

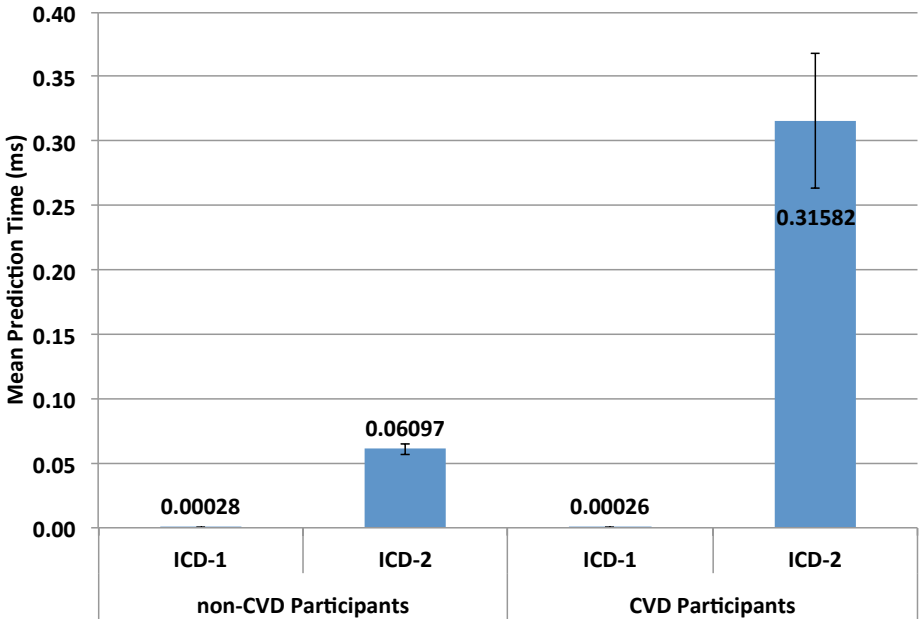


Figure 6.1: ICD-1 and ICD-2 mean prediction time ± s.e., by CVD presence.

The time required to execute a call to the `areDifferentiable` function is important because most recolouring algorithms make a large number of these comparisons in order to find a suitable set of replacement colours. For example, an image with 128 quantized colours could take more than seven minutes to process – this is clearly longer than most users would be willing to wait for a recolouring; therefore, improving the speed of the model's predictions is critical for real-world applicability. The ICD-3 model, which addresses this problem, is discussed in the following sections. ICD-3 uses the same calibration technique and calibration data as ICD-2, but substantially changes the core prediction functionality in order to make predictions much faster.

The key improvement of ICD-3 over ICD-2 is that it eliminates the necessity of recalculating the discrimination ellipsoid for each prediction by adapting a single global ellipsoid to each prediction. Constructing a new discrimination ellipsoid for each prediction requires a substantial amount of processing in ICD-2 (as shown in Figure 6.1. By replacing this with a global ellipsoid that is generated once, the prediction time of ICD-3 is greatly reduced.

This improvement is next explored in greater depth, beginning with an exploration of the

implications of these changes for recolouring tools.

6.1.1 Why is prediction speed important for recolouring?

In Section 2.5, the basic four-step process for recolouring was described. The third step of this algorithm involves the fundamental aspect of recolouring – that of determining a map from the representative colours to replacement colours. To find this replacement set of colours, recolouring tools typically use iterative constraint optimization techniques. These generate an initial set of replacement colours, and then iteratively improve the set until an optimal or near-optimal set of replacement colours is found.

To improve a potential set of replacement colours, the colours need to be evaluated in some manner to determine whether one set is better than another. To perform this evaluation, colour differentiation models are used. The goal of the recolouring tool is to make the differentiability between the colours in the replacement set (for the CVD user) as close to the differentiability of the colours in the original representative set (for the non-CVD user). The necessity of this goal can be seen in the situation of recolouring natural images – it is desirable to maintain the 'naturalness' of the image by recreating for the individual with CVD the colour differences apparent to someone without CVD. Alternatively, if the goal is to maximally enhance colour differences, then natural images (e.g., landscapes, portraits) would take on a strong and possibly disturbing sense of artificiality.

To evaluate a set of replacement colours, the differentiability of all pairs of colours in the replacement set is determined. These pairwise differentiability values are then compared to the pairwise differentiability values for the original set of representative colours. If the pairwise differentiability of one set of replacement colours is closer to the representative set of colours than another set of replacement colours, then the first replacement set is better than the second.

To evaluate a single set of replacement colours, all pairs of colours in the replacement set must be compared using the colour differentiation model (requiring $\frac{n\times(n-1)}{2}$ or $O(n^2)$ comparisons). Constraint optimizations (e.g. simulated annealing, genetic algorithms) use many iterations to determine an optimal or near-optimal solution. As a result, the $O(n^2)$ predictions are performed a number of times in the process of finding the set of replacement

colours. Assuming the number of iterations and the number of representative colours for an input image are fixed, then the time to find the replacement set of colours is directly proportional to the number of predictions performed. As an example, Kuhn et al.'s recolouring tool [90] defaults to finding 128 representative colours. Assuming 100 iterations are made in the constraint optimization algorithm (Kuhn et al. use a mass-spring system to achieve this), the number of differentiability predictions will be:

$$\frac{128 \times 127}{2} \times 100 = 812800$$

If a prediction takes 0.31582 ms to perform on average (ICD-2's mean prediction time for CVD participants in Figure 6.1), then the constraint optimization process will take 256 seconds (almost 4.5 minutes) to complete. This example illustrates that substantially reducing the time to perform a prediction will have direct implications for the time to perform a recolouring. For this reason, this is the main goal of ICD-3.

6.1.2 Improving ICD-2 Model Prediction Time

To reduce the time to perform a prediction, the core algorithm used to make predictions in ICD was examined to identify which steps take too much processing time. I found that generating a new discrimination ellipse for each prediction was the most time-consuming step.

As previously discussed in Section 5.2.4, when the `areDifferentiable(C1,C2)` function of ICD-2 is called, `C1` and `C2` are compared to the base colour to determine which is closest in L*u*v* colour space. The closest colour is called the *primary* colour, and the more distant colour is called the *secondary* colour. A discrimination ellipsoid is then generated for the primary colour, and the secondary colour is compared to this discrimination ellipsoid. If the secondary colour is within the discrimination ellipsoid, then the colours are not differentiable (return `false`), otherwise they are differentiable (return `true`).

To determine the discrimination ellipsoid for the primary colour in ICD-2, the ellipse that best fits the discrimination limits recorded in the calibration file is found. This best-fit ellipse is then coupled with the luminance discrimination limits (also recorded in the calibration file) to find the discrimination ellipsoid. To determine the six points necessary to find the best-fit

ellipse, ICD-2 performs six iterative 'walks' out from the primary colour's u*v* coordinates, one along each confusion line. To identify the most precise points, very small steps are taken in this iterative walk. Although this results in precise points from which to define the best-fit ellipse, the walking algorithm is computationally intensive. To compound the problem, this six-fold walking process is repeated for each call to `areDifferentiable` (i.e., each prediction).

The improvement of ICD-3 is to determine a single ellipsoid with the first invocation of `areDifferentiable` and then to reuse this ellipsoid for all future predictions. This 'global ellipsoid' is centred on the base colour (L*u*v* = [50.0, 0.0, 0.0]), and has its major axis aligned with the dominant colour confusion line for the user represented by the calibration file. For a given call to `areDifferentiable`, the global ellipsoid is translated so that it is centred on the L*u*v* coordinates of the primary colour, and is rotated in the u*v* plane so that the ellipsoid's major axis aligns with the dominant confusion line for the user. This is illustrated in Figure 6.2, where the location, size, and orientation of the discrimination ellipsoid as determined by the calibration points is shown on the left; a translated and rotated discrimination ellipsoid is shown on the right.

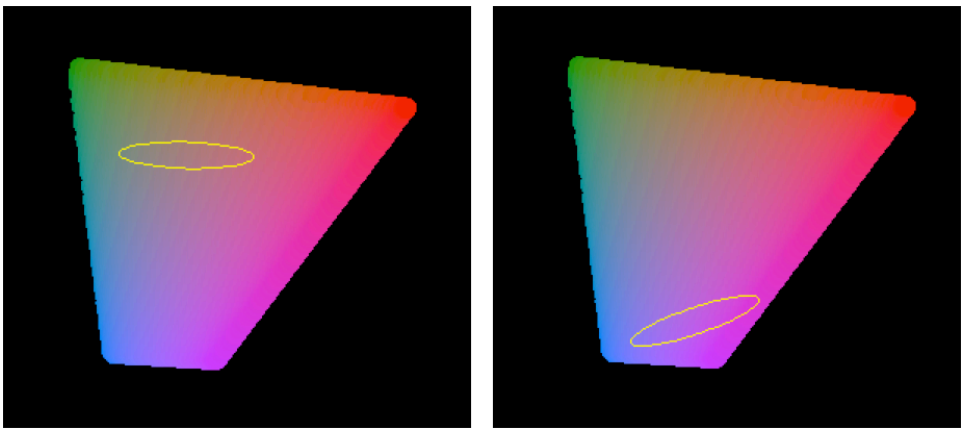


Figure 6.2: Discrimination ellipse (u*v* slice through ellipsoid) is shown for the calibration file (left) and after translation and rotation (right).

To generate the global ellipsoid, each differentiation limit in the calibration file is converted to L*u*v* coordinates by defining a polar coordinate representation of a colour in L*u*v* space, then converting these polar coordinates to Cartesian L*u*v* to find a point for the best-fit ellipse algorithm. As described above in Section 5.2.1, the CIE L*u*v* colour space

is a colour space defined by one luminance (relative perceived lightness - L*) axis, and two chromatic axes – u* and v* – that range in value from -100 to +100. Holding u* and v* both equal to 0.0 but varying L* gives the achromatic axis, which contains all greyscale colours from black (L*=0) to white (L*=100.0). Any straight line of constant L* radiating out from the achromatic axis defines a hue. The distance along this line is the chroma (colourfulness) of a colour. Each sample in the calibration file contains both a hue angle and a distance. The hue angle specifies the hue of a colour, and the distance specifies the chroma of a colour. These values for hue and chroma for each calibration sample are converted to L*u*v* Cartesian coordinates by:

$$u* = chroma \times cos(hue) \tag{6.1}$$

$$v* = chroma \times sin(hue) \tag{6.2}$$

$$L* = luminance\ of\ base\ colour\ (50.0) \tag{6.3}$$

L*u*v* Cartesian coordinates are generated for each calibration sample, and the entries that specify chromatic entries (not luminance ones) are used to find the best-fit ellipse using the same library as used in ICD-2. The major axis radius, minor axis radius, and orientation angle of this best-fit ellipse are used to define the same values for the global ellipsoid (major axis = a, minor axis = b). The luminance data entries in the calibration file are used to define the c-axis radius for the ellipsoid. With a, b, c, and angular orientation data, a complete ellipsoid is thereby defined.

To make a prediction, the global ellipsoid must be translated to the primary colour's L*u*v* coordinates and then rotated to align with the user's confusion line at those coordinates. To translate the global ellipsoid to the primary colour's location in L*u*v* colour space, the global ellipsoid's L*u*v* coordinates are simply changed to match the primary colour's coordinates. To rotate the global ellipsoid when making a prediction, an individual copunctal point is used. Ideally, the major axis of the global discrimination ellipsoid should be collinear with the confusion line for one of the three types of dichromacy (protanopia, deuteranopia, or tritanopia), thereby intersecting the copunctal point for that particular type of dichromacy (see Section 5.2.2, above). In practice, I found that the major axis rarely intersects the copunctal point, but typically passes near it. To accommodate this,

I developed the idea of an individual copunctal point in which a copunctal point for a particular user is created. To find this, it was identified that the three dichromatic copunctal points lie just outside the L*u*v* colour space, and are somewhat uniformly spread around it. We used these three points to define a circle, and found the point of intersection between the major axis of the global ellipsoid and this circle using the `JavaGeom` library (`www.geom-java.sourceforge.net`). This intersection point defines the individual copunctal point. This is illustrated in Figure 6.3. When the global ellipsoid is rotated, the rotation is performed about the center of the ellipsoid, and oriented so that the major axis intersects this individual copunctal point. As the individual copunctal point does not change once the ellipsoid is generated, it is calculated once when the global ellipsoid is generated to minimize processor load.

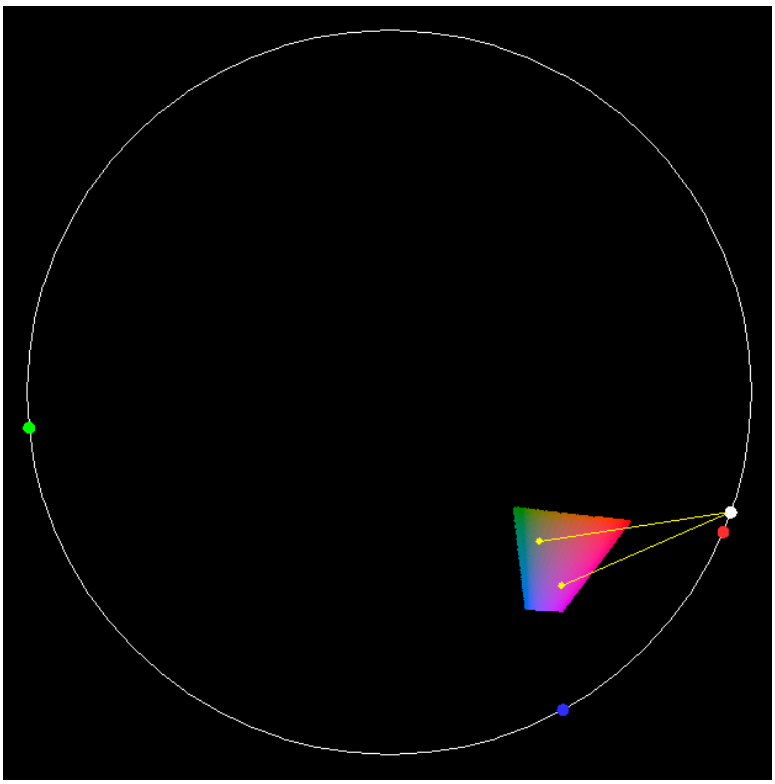


Figure 6.3: Illustration of finding an individual copunctal point. Protan (red), deutan (green), and tritan (blue) copunctal points. The large white circle is the circle used to find the individual copunctal point (smaller filled white circle). Yellow points are two colours (grey and purple) with their associated confusion lines intersecting the individual copunctal point.

6.1.3 ICD-3 Limit Offset

One additional small change from ICD-2 to ICD-3 is how the limit offset is implemented. In Section 4.3.2, I described a manner of tuning the colour differentiation model to make it more or less likely to make type-1 errors (false positives). To achieve this, I proposed adding a constant value to each differentiation limit that defines the discrimination ellipsoid (or the discrimination 'box/volume' used in ICD-1). This results in a larger discrimination ellipsoid, which will predict more colours as being 'not differentiable' thereby reducing the occurrence of false positives. Upon examination, I found that adding a fixed value to each ellipsoid dimension does not result in the ellipsoid scaling proportionally as it grows or shrinks. To address this, I modified limit offset to be a multiplicative scale factor. As such, the default limit offset is no longer 0.0 in ICD-3, but 1.0 (to represent no discrimination ellipsoid scaling). To implement this, the a, b, and c values for the global discrimination ellipsoid are multiplied by the limit offset when these properties are determined, as well as whenever a new limit offset is set.

6.2 ICD-3 Evaluation: Prediction Time, Prediction Accuracy, and Limit Offset

I conducted three separate studies to evaluate ICD-3. All three studies used either the calibration or the ground-truth data gathered for ICD-2 (see Section 5.3). First, I compared the prediction times of ICD-3 and ICD-2 using the calibration data; second, I compared the models' prediction accuracy using the ground-truth data; third, I explored the behaviour of the new limit offset implemented in ICD-3 using the ICD-2 evaluation calibration data.

6.2.1 ICD-3 Prediction Time Study Design

This evaluation used a repeated-measures factorial design with two factors: model type (ICD-2 or ICD-3) and CVD presence (non-CVD or CVD). The CVD-presence factor was included to determine if the models have different prediction times depending on the type of user. Average prediction time was the dependent measure.

To measure prediction time for each model, I used the eight CVD and eight non-CVD calibrations gathered during the evaluation of ICD-2. For each participant, I generated an ICD-3 model and an ICD-2 model according to the data specified in the participant's calibration files. Once the models were generated, one million pairs of random RGB colours were generated. Each model was then used to execute `areDifferentiable` on the entire set of pairs and the total time was recorded, then averaged. This relatively large set of random RGB samples was chosen because the ICD-3 model is very fast, requiring a large sample size to get representative timing results.

6.2.2 ICD-3 Prediction Time Results

On average, ICD-3 was 56.4 times faster at making predictions than ICD-2. RM-ANOVA, showed main effects of model type ($F_{1,14}$=52.69, $p<0.001$) and CVD presence ($F_{1,14}$=25.96, $p<0.001$), as well as an interaction effect between these factors ($F_{1,14}$=25.96, $p<0.001$). Post-hoc Bonferroni-corrected (four comparisons; adjusted alpha=.0125) paired *t*-tests showed significant differences between ICD-2 and ICD-3 for non-CVD ($p<0.001$) and CVD participants ($p<0.001$), as well as significant differences between CVD and non-CVD participants for ICD-2 ($p<0.001$), but not for ICD-3 ($p=.036$). These results are summarized in Figure 6.4, below. These prediction times were measured on a 2.0 GHz Intel Core 2 Duo MacBook with 4 GB of RAM, running OS X Snow Leopard (10.6.8).

It is clear that the average prediction time for ICD-2 presented in Figure 6.4 is greater than the ICD-2 prediction time shown in Figure 6.1. The reason for this is that this later study used a slightly revised version of ICD-2. Recognizing that ICD-2 has an overly lengthy prediction time, a caching mechanism was implemented in the revised ICD-2 version to reduce prediction times by retaining calculated ellipses to use in future predictions. This approach makes sense for image recolouring, where colour differentiation predictions are made repeatedly on a relatively fixed set of colours (the image's representative colours). However, in this evaluation, pairs of random RGB colours were utilized, substantially reducing the effectiveness of the caching mechanism, but still incurring the computational overhead of performing (and updating) the cache. As a result, the average prediction time for ICD-2 in this study is longer than the average ICD-2 prediction time presented in Figure 6.1.

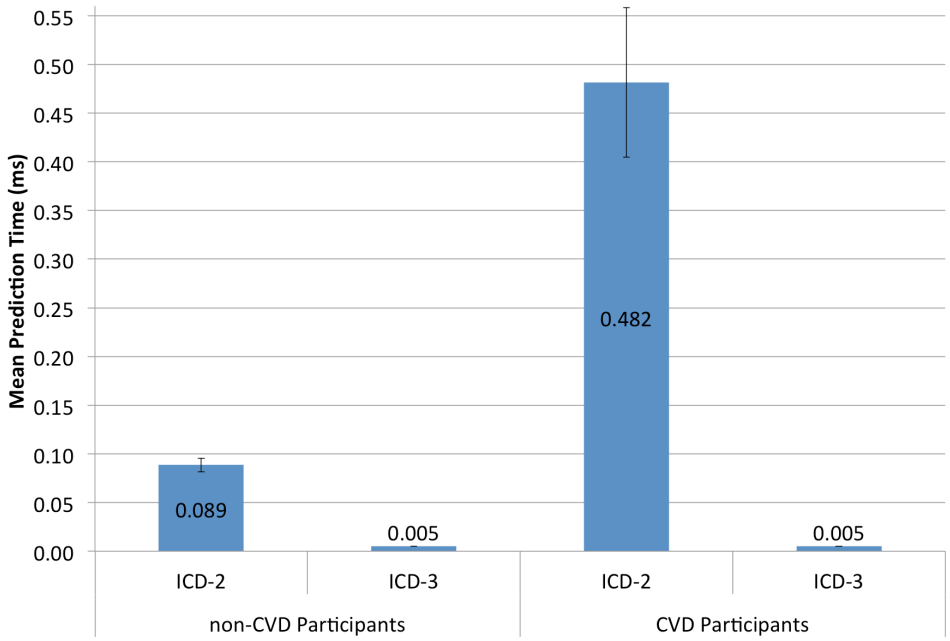


Figure 6.4: ICD-2 and ICD-3 mean prediction time ± s.e., by CVD presence.

The results show that not only is ICD-3 faster at making predictions than ICD-2, but it is also more consistent in the amount of time taken to make a prediction, regardless of whether the participant has CVD or not. This result is intuitive because ICD-2 performs six confusion line 'walks' to find the best-fit ellipse data points. The farther the model has to walk down a confusion line (as is the case for participants with CVD), the longer the model will take to perform a prediction. As ICD-3 performs the same translations, rotations and calculations of intersections independent of the size of the global ellipsoid, the prediction time for ICD-3 is stable across all participants.

Returning to our motivation for improving prediction speed presented in Section 6.1.1, the example given there demonstrated that recolouring may require millions of differentiability predictions (812800, in that particular example). With the prediction timing results presented in Figure 6.4, performing a recolouring using ICD-2 for a user with CVD will take 392 seconds (over 6.5 minutes) on average. The same recolouring using ICD-3 will take just over four seconds, which is much closer to near real time performance requirements for recolouring tools.

Although ICD-3 makes significant gains in the time to perform a prediction, these gains may come at the cost of decreased prediction accuracy. To address this, the prediction

accuracy of ICD-2 and ICD-3 are compared next.

6.2.3 ICD-3 Accuracy Study Design

This evaluation used a repeated-measures factorial design with two factors: model type (ICD-2 or ICD-3) and CVD presence (present or not); CVD presence was included in the analysis only to check for interactions with model type. I collected error rates for the two models and compared them on five dependent measures: true positive rate (proportion of predictions that correctly suggested that colours were differentiable), true negative rate (proportion of predictions that correctly suggested that colours were not differentiable), false positive rate (proportion of predictions that incorrectly suggested that colours were differentiable), false negative rate (proportion of predictions that incorrectly suggested that colours were not differentiable), and overall accuracy (number of correct predictions over total trials).

To measure the raw values used to determine our dependent measures, I used the 540 ground-truth colour-differentiation responses gathered from each participant during the evaluation of ICD-2 presented in Section 5.3. For each participant, ICD-2 and ICD-3 models were generated using the participant's respective calibration data file. For each entry in the participant's 'ground truth' set of responses (540 colour differentiation responses), each model performed a prediction of the differentiability of the entry's constituent colours. These predictions were compared to the ground truth response from the participant, and the result (true positive, true negative, false positive, false negative) was recorded for each model.

6.2.4 ICD-3 Accuracy Results

RM-ANOVA found no significant main effect of model type on any of the dependent measures, and no significant interactions between model type and presence of CVD (all $p>0.05$). The error rates are shown in Figure 6.5 below.

The results show that the speed improvement of ICD-3 does not come at a cost in prediction accuracy. These results are not surprising considering that both models incorporate essentially identical internal ellipsoidal representations to represent a user's colour differentiation abilities. The difference is in how each model constructs discrimination ellipsoids from

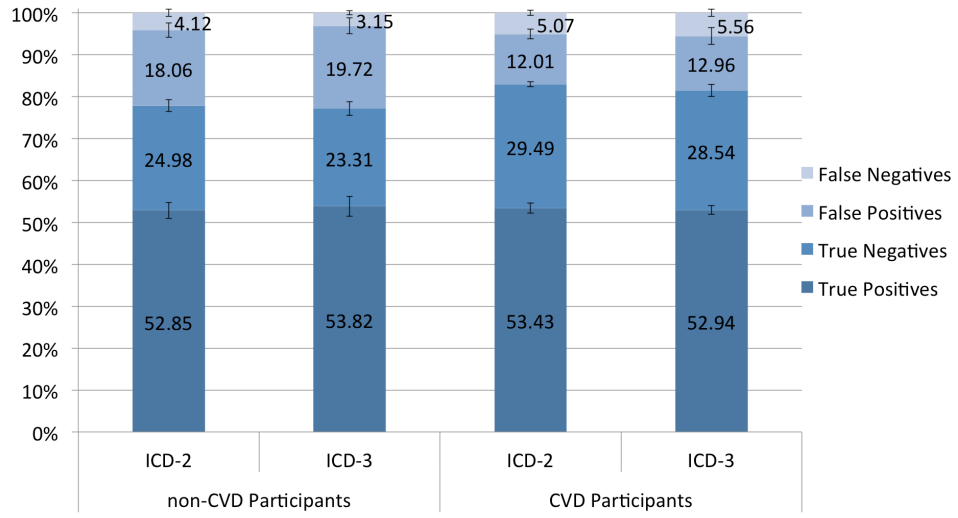


Figure 6.5: True positive, true negative, false positive, and false negative rates ± s.e. for ICD-2 and ICD-3 by CVD presence. Accuracy is true positive rate plus true negative rate.

calibration data and prediction parameters, not the resulting ellipsoids. As a result, their respective prediction accuracies are not significantly different.

6.2.5 ICD-3 Limit Offset Analysis

Instead of performing a comparative analysis between ICD-2 and ICD-3 regarding the effect of limit offset on either prediction performance or prediction accuracy, I simply present data that describes the prediction accuracy of ICD-3 as the limit offset varies. This is for two reasons:

1. The limit offset implemented in ICD-1 and ICD-2 was additive in nature (a constant value was added to each dimension of the discrimination volume). The result of this type of limit offset is that as the limit offset is increased, the discrimination volume does not scale uniformly. To correct this, I implemented the limit offset as a multiplicative scaling factor in ICD-3. An outcome of this change is that the effect of limit offset is not directly comparable between the older versions of the model and ICD-3.

2. As shown in the prediction time analysis above, the size of the ellipsoid does not influence the prediction time for ICD-3 (CVD ellipsoids are much larger than non-

CVD ellipsoids, yet prediction time was not significantly different). ICD-2, however, does exhibit a dependency between ellipsoid size and prediction time, as shown above. Because these differences in performance are already known, repeating the prediction time performance analysis in terms of limit offset changes offers little additional value.

For these two reasons, I opted to present descriptive accuracy statistics to illustrate how the new limit offset affects prediction accuracy for ICD-3. These data are presented in Figures 6.6 and 6.7 and can be used to select a limit offset for use in ICD-3. If someone is to utilize ICD-3 in an application (not necessarily recolouring) that favours false negatives over false positives, then a small limit offset can be chosen. If, on the other hand, the application favours false positives over false negatives, then a larger limit offset can be chosen. These charts also identify the limit offset that optimizes the model prediction accuracy, if the application requires maximum accuracy. In this way, these charts can be used to select limit offsets when utilizing ICD-3 to make colour differentiation predictions.

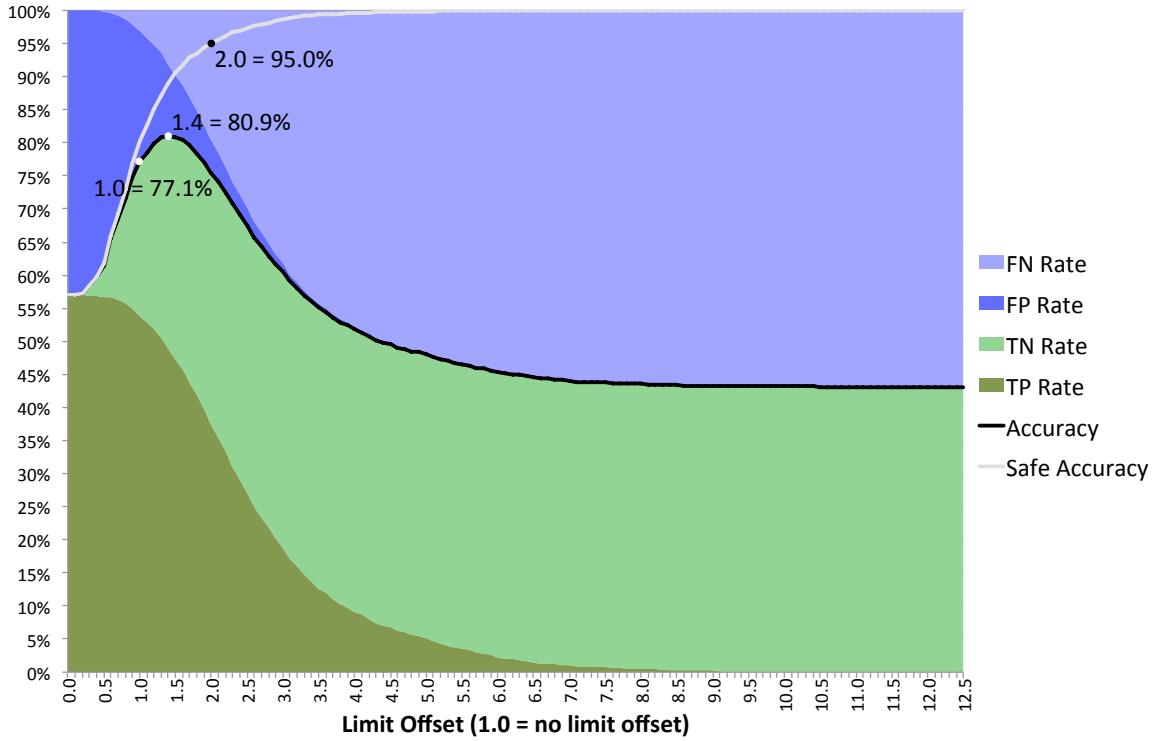


Figure 6.6: True positive, true negative, false positive, false negative, accuracy, and safe accuracy rates for ICD-3 for non-CVD participants. Maximum accuracy (1.4 = 80.9%), 95% safe accuracy (2.0 = 95.0%), and default limit offset accuracy (1.0 = 77.1%) are shown.

As can be seen in Figures 6.6 and 6.7, a limit offset of 0.0 produces only true positives and false positives. This is because a limit offset of 0.0 causes the discrimination ellipsoid to have a volume of 0, so every prediction is predicted as completely differentiable (positives). It can also be seen that as the limit offset increases, the model makes increasingly more 'not differentiable' (negatives) predictions. This manifests as a long tail reduction in true positives and false positives, which stabilize to 0 at the right end of each chart (12.5 for non-CVD participants, and 9.0 for CVD participants). Moving from left to right (increasing limit offset), the peak accuracy can be found (non-CVD: 1.4 = 80.9%; CVD: 1.2 = 82.6%), which suggests that to achieve maximum accuracy when using ICD-3, a small limit offset should be used. Following the analysis of ICD-1(Section 4.3.2), I also present the 'safe offset' value at which the model makes less than 5% false positive predictions (non-CVD: 2.0 = 95.0%; CVD: 1.5 = 95.8%).

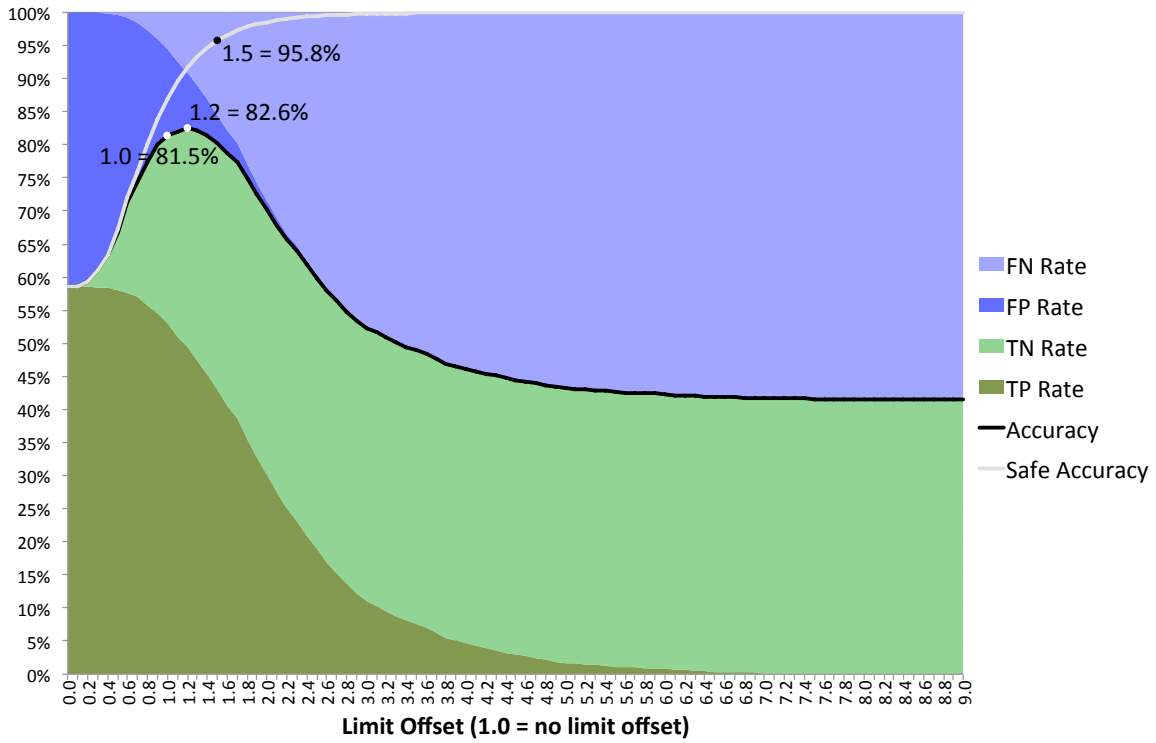


Figure 6.7: True positive, true negative, false positive, false negative, accuracy, and safe accuracy rates for ICD-3 for CVD participants. Maximum accuracy (1.2 = 82.6%), 95% safe accuracy (1.5 = 95.8%), and default limit offset accuracy (1.0 = 81.5%) are shown.

As stated above, the non-CVD participant limit offsets stabilize at 12.5, but CVD participant limit offsets stabilize at 9.0. This is because CVD participants begin with a larger discrimination ellipsoid, so the ellipsoid grows more quickly with limit offset increases for CVD participants. As a result, the secondary colour for every prediction falls within the discrimination ellipsoid at a lower limit offset for CVD participants, leading to earlier stabilization of all negative predictions.

6.3 ICD-4: Predicting Degree of Differentiability

The final problem found in the earlier ICD models is that the model's predictions are binary – that is, they only provide a yes/no prediction of one colour's differentiability from another, corresponding to a simple step function that approximates the user's actual psychometric function (see Chapter 3). In this section, the final version of the model, ICD-4 is presented. ICD-4 uses a closer approximation of the actual psychometric function (a sigmoid function) and is therefore able to provide the specific degree of differentiability between two colours when making a prediction.

Although the step-function approximation works well in many settings (as shown in the above evaluations), it has two limitations. First, in situations where the recolourer has few colours left to work with (e.g., when many colours need to be changed), the model will be forced to consider colours that are closer to the boundary of user's discrimination ellipsoids; in these cases, it will be useful to have more information about the actual differentiability of the candidate replacement colours. Second, there are situations where degree of differentiation is a necessary part of the use of colour – for example, gradients in false-colour visualizations require that colours for the changing variable move smoothly away from a starting colour. To recolour these visualizations, the system must be able to predict the specific degree of differentiability.

To accomplish this, ICD-4 extends the boolean `areDifferentiable` function used in the previous models to a real-valued `howDifferentiable` function that returns values between 0 and 1. This range is mapped to the psychometric function described above, in which 0 represents the situation where the user will never see the two colours as different, 1 represents

the situation where they will always see them as different, and values in between represent the proportion of cases where the two colours will be seen as differentiable. A predicted value of 0.5 represents colours that are on the boundaries of the discrimination ellipse, and should be seen as different from the starting colour about half of the time (this is the threshold definition of 'just noticeably different' [14]). To provide the maximum benefit for constraint optimization based recolouring tools, the `howDifferentiable` function should never actually return exactly 0 or 1, as these would represent plateaus in the solution space. Plateaus result in the recolouring tool having insufficient information available in order to determine where to look for more optimal solutions. As a result, ICD-4's `howDifferentiable` function returns values exclusively between 0 and 1.

6.3.1 Implementing `howDifferentiable` in ICD-4

The new version of the prediction function utilizes a sigmoid function at the boundary of the discrimination ellipsoid, instead of the step function used in previous versions. The value returned by `howDifferentiable` follows the psychometric function described in Section 3.4.1. This change is still an approximation to the true sigmoid curve for a particular user (i.e., ICD-4 does not gather new calibration data to determine the exact curve), but is a much closer approximation to the shape of the psychometric function than a simple step function.

Sigmoid functions map any input value x to a value between 0.0 and 1.0 with the following equation:

$$sigmoid(x, p_{50}) = 1.0/(1.0 + e^{-0.1\times(x-p_{50})}) \tag{6.4}$$

where p_{50} is the value of x where the sigmoid should return 0.5 (the half-way point). To determine a `howDifferentiable` value for any given pair of colours, a value for x and a value for p_{50} are needed. These can be found by generating a vector from the primary colour's colour space coordinates that passes through the colour space coordinates of the secondary colour. The point at which this vector intersects the primary colour's ellipsoid represents the p_{50} value. The center of the secondary colour represents the x value. To move these inherently three-dimensional values to single-dimension values to be used in the sigmoid function given above, distance from the origin along this vector can be used. The

intersection will be at one distance, and the center of the secondary colour will be at another (possible identical) distance.

To determine the real-valued `howDifferentiable` value in ICD-4, the primary and secondary colours within L*u*v* space are translated such that the primary colour is centred on the origin with its ellipsoid's major axis collinear with the u* axis. This allows ICD-4 to define the vector described above, which starts at the origin (0,0,0) and passes through the coordinates of the secondary colour (x_2,y_2,z_2). This vector can be defined parametrically by:

$$x = t \times x_2 \tag{6.5}$$

$$y = t \times y_2 \tag{6.6}$$

$$z = t \times z_2 \tag{6.7}$$

The point at which this vector intersects the ellipsoid for the primary colour can be found by substituting this parametric definition into the equation for the primary colour's ellipsoid, giving:

$$1 = \frac{(t \times x_2)^2}{a^2} + \frac{(t \times y_2)^2}{b^2} + \frac{(t \times z_2)^2}{c^2} \tag{6.8}$$

Solving for t gives:

$$t = \frac{1}{\sqrt{\frac{x_2^2}{a^2} + \frac{y_2^2}{b^2} + \frac{z_2^2}{c^2}}} \tag{6.9}$$

Once t has been calculated, the coordinates for the intersection can be found by substituting the value for t into Equation 6.8, and solving for x, y, and z. This gives a three-dimensional point that represents the intersection of the surface of the primary colour's ellipsoid with the vector from the primary colour's center (0,0,0) to the secondary colour's center (x_2,y_2,z_2). The Euclidean distance from the origin to the intersection point is calculated and this serves as p_{50} in Equation 6.4. The Euclidean distance from the origin to the secondary colour's translated and rotated center is calculated, and this serves as x in Equation 6.4.

6.4 Evaluation of ICD-4: Gradients

As a simple demonstration of how real-valued predictions provide more power than binary predictions, ICD-3 and ICD-4 were used to recolour an image of a 12-step gradient colour

scheme that could be used in false-colour images using a modified version of the recolouring tool presented in Chapter 7. The gradient (see Figure 6.8, top left) contains a bipolar gradient which begins with blue, transitions through grey, and ends with a brownish green, and could be used to colour code elevation above and below sea level in a topographic map. This gradient lies very near the tritanopic colour confusion line, hence presents difficulties to those with inherited tritanopic CVD (commonly called 'blue-yellow colour blindness'), as illustrated in Figure 6.8, top right.

ICDRecolour (presented in Chapter 7) was modified and used to recolour the gradient image using ICD-3, which uses the binary predictor, and ICD-4, which provides degree-of-differentiability values. For ICD-3, ICDRecolour's default behaviour was used – recolouring was allowed to continue in an unrestricted fashion using unconstrained replacement colours until 12 mutually-differentiable replacement colours were identified using the user's colour differentiation model. The ICD model provided to ICDRecolour was calibrated using hypothetical tritanopic responses (typical protan, deutan, and luminance limits, large tritan limits).

For ICD-4 recolouring, ICDRecolour was extended to perform contraint optimization using simulated annealing. To accomplish this, two ICD-4 model instantiations were used – one representing the hypothetical tritanopic individual used in the ICD-3 recolouring, and one representing a typical individual. Using simulated annealing, the solution space of possible replacement colour sets was searched to find a minimal-cost replacement colour set. The cost function was composed of a component that favoured retaining the original colours (maintaining naturalness) and a component that favoured making the replacement colour set as mutually differentiable as the original colours (restoring differentiability). These cost components were added together to obtain the final cost function. If this modified recolouring tool is able to produce a replacement colour set that has the same internal differentiability for the user with CVD as the original representative colour set has for the users without CVD, then the `howDifferentiable` function operates as intended.

The results of the recolouring are shown in Figure 6.8, middle and bottom left, with tritanopic simulations on the right. ICD-4 provides a much smoother gradient in the replacement colour set than ICD-3, since ICD-3 can only produce a set of colours that are all

differentiable (rather than differentiable by specific amounts). Note that these recolouring results must be assessed in terms of how the individual with tritanopia will perceive them (simulated images on the right of Figure 6.8), as this is the type of CVD being accommodated through the recolouring.

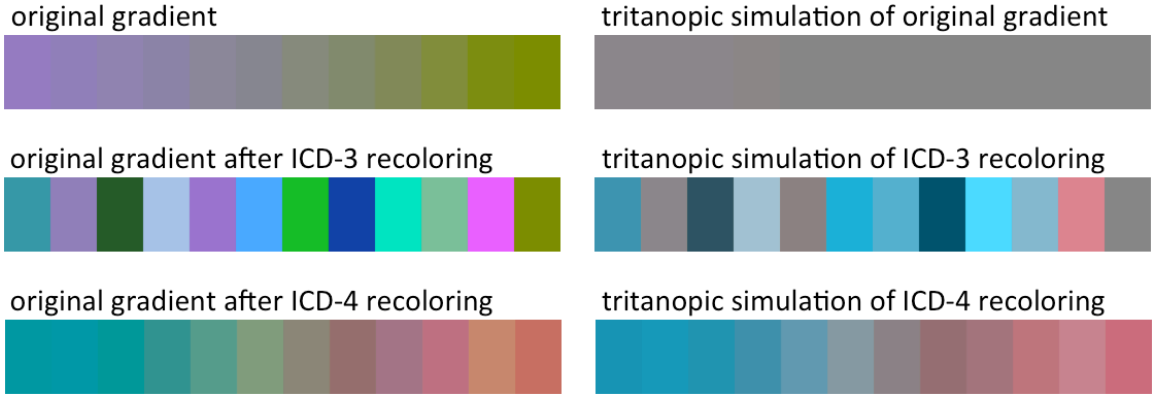


Figure 6.8: Left: Original 12-step gradient; ICD-3 recolouring; ICD-4 recolouring. Right: simulation of left images by user with tritanopia. Both recoloured versions use identical calibration data for a hypothetical user with tritanopia.

6.4.1 ICD-4 Prediction Speed and Accuracy

In terms of prediction speed and performance, the extensions proposed here were actually implemented before the prediction speed and accuracy evaluation of ICD-3 was competed. As a result, the results presented in Figures 6.4 and 6.5 hold for ICD-4. Incorporating the sigmoid function at the ellipsoid boundary to allow `howDifferentiable` predictions does not result in slow prediction speeds or reduced accuracy over previous versions of the model.

6.5 Discussion

The evaluations above show the main benefits of the ICD-3 and ICD-4 individualized models of colour differentiation:

- ICD-3's current model prediction time (0.005 ms per prediction) is low enough to make real world constraint-optimization recolouring possible without undue delays for the user;

- ICD-4's revised prediction mechanism provides degree-of-differentiation information that increases the range of visualizations that can be recoloured.

Chapter 4 showed that the situation-specific modelling (SSM) approach can be applied to individualized models of colour differentiation (ICD) to enable colour differentiation models that are sensitive to differences between users as well as to the influence of environmental variations. Chapter 5 introduced a refined calibration technique that is 24 times faster than the feasibility implementation's calibration. The current chapter reduced the ICD prediction time and increased the ICD prediction power to enable the application of ICDs to recolouring tools. In the next chapter, an ICD-based recolouring tool is presented, along with a comparative evaluation of this recolouring tool to examine the full impact of the model developed in this dissertation as a solution to the problem that existing recolouring tools do not help most people with CVD.

Chapter 7

ICD-Based Recolouring

This chapter presents the final evaluation of the individualized model of colour differentiation (ICD) presented in previous chapters. In this chapter, the implementation of a recolouring tool based on the ICD is described, along with a comparative evaluation of this recolouring tool. [1]

7.1 Applying the ICD to Recolouring

In Section 1.1, it was identified that existing recolouring tools do not help most people with CVD because they rely on models of colour differentiation that only represent a small proportion of the population of people with CVD. Chapters 3–6 describe the solution presented in this dissertation and show that the proposed Individualized model of Colour Differentiation (ICD) based on Situation-Specific Modelling (SSM) represents a viable technique for modelling the colour differentiation abilities of particular users in particular situations.

However, none of the evaluations presented in Chapters 3–6 directly address the problem presented in Section 1.1, namely that *existing recolouring tools do not help most people with CVD*. In this chapter, the ICD developed in previous chapters (in particular, Chapter 6's ICD-4), is used to construct an ICD-based recolouring tool called ICDRecolour. With ICDRecolour, the effectiveness of SSM-based ICDs as a solution to this dissertation's problem can be evaluated.

To conduct this evaluation, the colour matching performance of participants with and without CVD were analyzed. Their matching performance was measured under a number

[1] Portions of this chapter were published as: Flatla, D.R., Gutwin, C. (2012) SSMRecolor: Improving Recoloring Tools with Situation-Specific Models of Color Differentiation. In *CHI '12: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2297-2306.

of different situations using no recolouring tool, ICDRecolour, and two publicly-available recolouring tools. Both matching accuracy and match-selection response time were analyzed, showing that ICDRecolour resulted in a 20% increase in matching accuracy to an overall average of 90%, and allowed participants to make matching selections approximately two seconds faster. These results were largely consistent across participant type and situation, and demonstrate the value of the Situation-Specific Modelling approach used to develop the ICD.

The remainder of this chapter is organized into two major sections. The first section describes the design and implementation of ICDRecolour. The second section describes the experimental design and results for the comparative evaluation. The chapter concludes with short discussion of the implications of these findings.

7.2 ICDRecolour: An ICD-based Recolouring Tool

ICDRecolour is an image recolouring tool that reduces images to sets of representative colours, determines a mapping from representative colours to replacement colours, and applies this map to the original image to produce a recoloured image. The majority of this chapter is focussed on how ICDRecolour finds the mapping from original colours to replacement colours, as this is where the ICD model is utilized. This step incorporates two smaller steps – identifying problem colours and finding suitably-differentiable replacement colours. More formally, recolouring tools follow a four-step process to perform recolouring:

1. Reduce the colours in the input image to a representative set of colours (usually through quantization or clustering). As the number of colours in an image can be quite large, representative colours are used to reduce recolouring time.

2. Within the representative colours, identify 'problem colours' – colours that the user has difficulty differentiating.

3. Find suitably-differentiable replacement colours for the problem colours, thereby generating a mapping from representative colours to replacement colours.

4. Replace the original input image colours with the recolour mapping colours to generate a recoloured image.

It should be noted that Step 3 does not require that every representative colour maps to a new replacement colour. Representative colours can also map to themselves, thereby maintaining the original colour during recolouring.

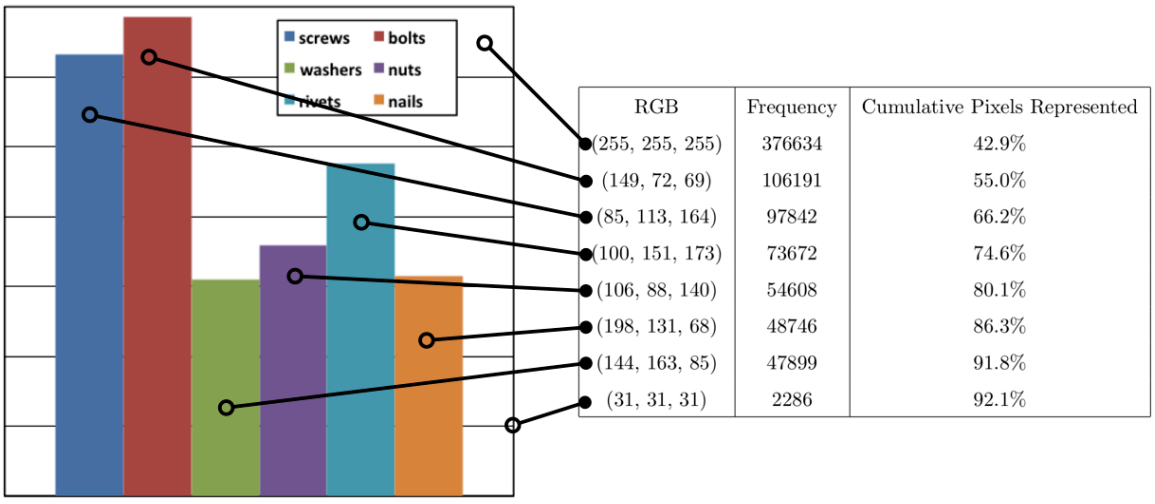
ICDRecolour is implemented in Java, utilizing the colour conversion functions described in Appendix A.

7.2.1 Identifying Representative Colours

After loading the image file to be recoloured, the first task of ICDRecolour is to identify the representative colours in the input image. This has been explored extensively in previous work (e.g., [60, 74, 90] – see Section 2.5 for more details), so ICDRecolour employs a simple approach to achieve this step. In addition, ICDRecolour's central contribution is the use of the ICD to improve recolouring, so the focus for ICDRecolour is on Steps 2 and 3, rather than Steps 1 and 4.

To find the representative colours for an image, ICDRecolour utilizes a simple image analysis method. This method starts with the user specifying the number of representative colours to identify in the input image (similar to specifying the number of clusters in k-means clustering [79]). The input image is then scanned pixel-by-pixel, creating a histogram of the frequency of each RGB colour contained in the image. This histogram is implemented as a `Java.util.TreeMap<Integer,Integer>` object with integer RGB colours as keys, and the frequency or count for that RGB colour as its value. Once the frequency histogram is found, it is sorted into frequency-descending order (most frequent colours first). The first N colours serve as the representative colours, where N is the number of representative colours specified by the user. This process is demonstrated in Figure 7.1 using the chart from Figure 1.1.

Although this approach works well for an image that contains a clear number of representative colours (e.g., a bar chart that utilizes a unique colour for each bar, plus the background and text colours), there are cases where the user will have difficulty specifying the number of representative colours. An alternative is to choose the highest-frequency colours until



RGB	Frequency	Cumulative Pixels Represented
(255, 255, 255)	376634	42.9%
(149, 72, 69)	106191	55.0%
(85, 113, 164)	97842	66.2%
(100, 151, 173)	73672	74.6%
(106, 88, 140)	54608	80.1%
(198, 131, 68)	48746	86.3%
(144, 163, 85)	47899	91.8%
(31, 31, 31)	2286	92.1%

Figure 7.1: Sample input image with the first eight entries in its respective colour frequency histogram. It can be seen that using eight representative colours allows over 92% of the image pixels to be represented by this set of representative colours.

some pre-set percentage of original image pixels is represented. In the example shown in Figure 7.1, the eight representative colours represent just over 92% of the image's pixels, so a value between 90% and 95% could be used to automatically generate an approximate set of representative colours using this method.

7.2.2 Finding Representative Colour to Replacement Colour Mapping

The main contribution for ICDRecolour is using the ICD developed in this dissertation to identify the problem colours in the set of representative colours and to find a mapping from the set of representative colours to replacement colours, where problem colours are replaced with suitably-differentiable replacement colours.

Continuing with the example in Figure 7.1, the original (left) and simulated (right) appearance of this chart to someone with protanopia is shown in Figure 7.2. It can be seen that someone with protanopia will have difficulty differentiating the first and fourth columns ('screws' and 'nuts'), the third and sixth columns ('washers' and 'nails'), as well as possibly the first and fifth column ('screws' and 'rivets'). These pairs of colours represent problem colours for this user.

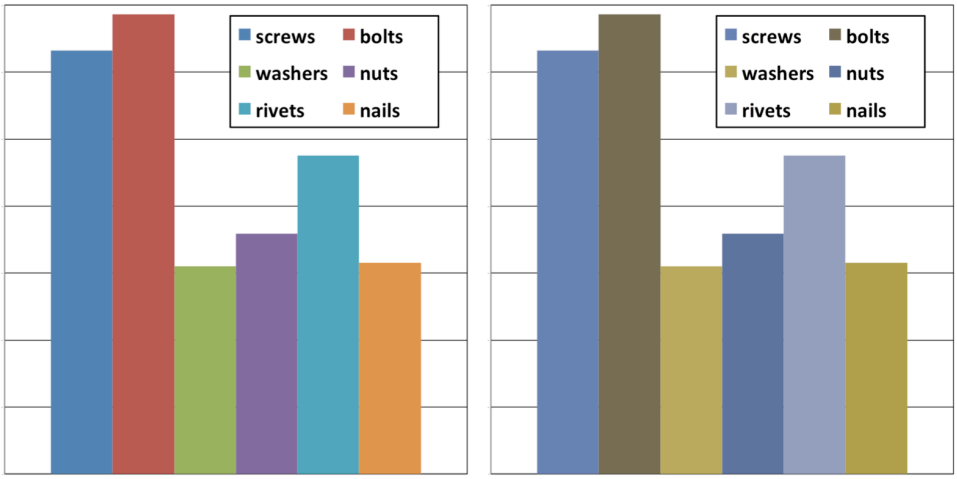


Figure 7.2: Left: a simple Microsoft Excel for Mac 2011 chart with six categories. Right: this chart as seen by someone with protanopia, illustrating the colours that present differentiation problems for someone with this type of inherited CVD. Simulation courtesy `www.vischeck.com`

ICDRecolour builds a mapping from the set of representative colours (Figure 7.3 to a set of replacement colours. This two-step process is iterative, first identifying a non-differentiable problem colour from the representative colour set, then assigning a replacement colour for the identified colour. The iteration continues greedily until no problem colours are identified (i.e., all the resulting colours are differentiable). This process is fundamentally different from previous approaches because any colours that are differentiable from all other colours are typically maintained, resulting in only problem colours being modified. As mentioned in Section 1.1, many existing recolouring tools modify all of the colours in an image, resulting in the unnecessary introduction of false colours.

As discussed in Section 2.6.3, two approaches to naturalness preservation during recolouring have been developed. In the first approach, most colours in the original image are modified by moving them all to the set of colours perceived identically by people with dichromacy and typical colour vision. This allows the solution search space to be substantially reduced, helping limit the time to recolour an image. The second approach identifies single problem colours, and replaces them with more differentiable colours, but leaves colours that do not cause problems unchanged. Both approaches result in replacing problem colours with colours that are more differentiable for users with CVD, and both approaches preserve naturalness for the user with CVD. The latter approach however, also preserves naturalness for users

with typical colour vision. In cases where users with and without CVD may be working with original and recoloured images (e.g., in a collaborative work environment such as co-authoring a paper), this latter approach provides naturalness preservation for every user. For these reasons, ICDRecolour employs the second approach to naturalness preservation; only problem colours are replaced, every other colour is maintained during recolouring.

Figure 7.3: The set of representative colours (including text and background colours) for the left chart in Figure 7.2, as determined using the procedure described in Section 7.2.1.

Colour Differentiation Model

The in-situ calibration technique described in Chapter 5 is performed to generate an individualized model of colour differentiation (ICD) for the user. This calibration captures the effect of any factor (inherited, acquired, or environmental) on the colour differentiation abilities of a user, and allows predictions to be made about these abilities. The model provides the Boolean function `areDifferentiable(C1,C2)`, which predicts whether colours `C1` and `C2` are differentiable or not. For illustration purposes, it is assumed that the user has protanopia and has performed the calibration procedure already.

Replacement Colours

ICDRecolour draws replacement colours from one of two sets: unconstrained and luminance-maintained. The unconstrained set is simply a random RGB colour generator; it is unconstrained because nothing restricts the suggested replacement colour. The luminance-maintained set contains RGB colours of the same CIE L*u*v* luminance as a single representative colour. One set for each representative colour is loaded at runtime from a pre-generated database when ICDRecolour begins the recolouring procedure. Replacement colours for each representative colour are drawn from the representative colour's isoluminant replacement colour set, ensuring the replacement colour is the same perceived luminance as the original

representative colour. For the illustration given here, the unconstrained set of replacement colours is used, but the unconstrained and luminance-maintaining sets are used during the evaluation of ICDRecolour.

Recolour Mapping Generation Algorithm

To generate the recolour mapping, the `areDifferentiable` function of the ICD is used to compare every possible pair of representative colours. Each comparison results in a prediction from the model indicating whether the pair is differentiable or not. The results from these comparisons are represented as a network where representative colours are nodes, and an edge between two nodes indicates that the two colours are not differentiable for the user. Figure 7.4 shows the colour differentiation network for the original chart image.

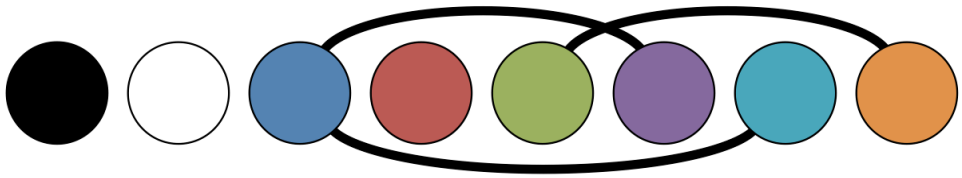


Figure 7.4: Colour differentiation network for Figure 7.2, left for someone with inherited protanopic CVD. The network shows the five problem colours linked into three pairs for the original chart image.

To identify the first replacement colour, the representative colour with the highest degree in the network (i.e., is confused with the most other colours) is replaced with a colour from the replacement set (the unconstrained set for this example), and the colour differentiation network is regenerated using the ICD. In this example, the blue (third from left) is replaced with a more vibrant blue. Once the network is regenerated (Figure 7.5), the edges representing confusion with the original colour are eliminated.

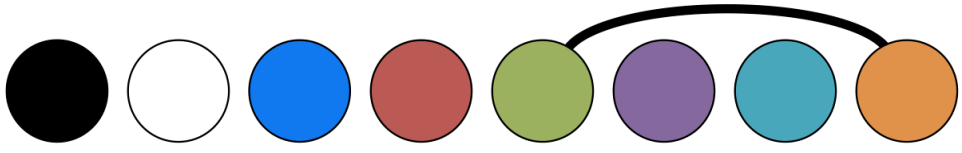


Figure 7.5: Colour differentiation network after the blue (third from left) has been replaced with a more vibrant (and differentiable) blue. Now only two problem colours (with one edge between them) remain.

The algorithm again selects the highest degree node (as there are two nodes with degree of one, the rightmost orange colour is arbitrarily chosen). Replacement colours are selected until this node's edge is eliminated, switching the problematic orange to a bright green (Figure 7.6).

Figure 7.6: Colour differentiation network after the orange (rightmost) representative colour has been replaced with a bright yellow. No edges remain, indicating this set of replacement colours is entirely mutually differentiable.

The colour differentiation network shown in Figure 7.6 contains no edges, indicating that the network contains no more problem colours. This set of colours is now used to define a mapping from representative colours to replacement colours. In ICDRecolour, this mapping is generated as a `Java.util.TreeMap<Integer,Integer>`, and is shown in Figure 7.7.

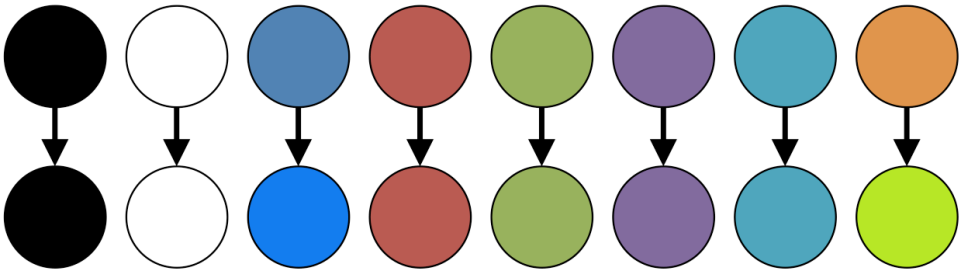


Figure 7.7: Recolour mapping from the representative colours to the replacement colours that give no non-differentiable colour pairs (no problem colours). The third from the left and rightmost colours have been replaced – all other representative colours remain.

7.2.3 Generating the Recoloured Image

Now that the representative colour to replacement colour mapping has been generated, this mapping can be used to recolour the original image to generate the recoloured output image. To accomplish this, a duplicate `Java.awt.image.BufferedImage` is created using the same size and colour profile as the original image. The original image is then analyzed pixel-by-pixel. For each pixel, the mapping is checked to see if the pixel's RGB colour is contained

in the representative set. If it is part of the representative set, then the replacement colour for that representative colour is used to set that pixel's colour in the duplicate image. If the pixel's colour is not part of the representative set of colours, then the original pixel colour is simply copied into the duplicate image. Once all of the pixels have been processed, the duplicate image is output as the recoloured image. Results from the illustration presented here are shown in (Figure 7.8, left). Figure 7.8, right shows the protanopia-simulated appearance of the recoloured chart to show that the original colour differentiation problems have been resolved.

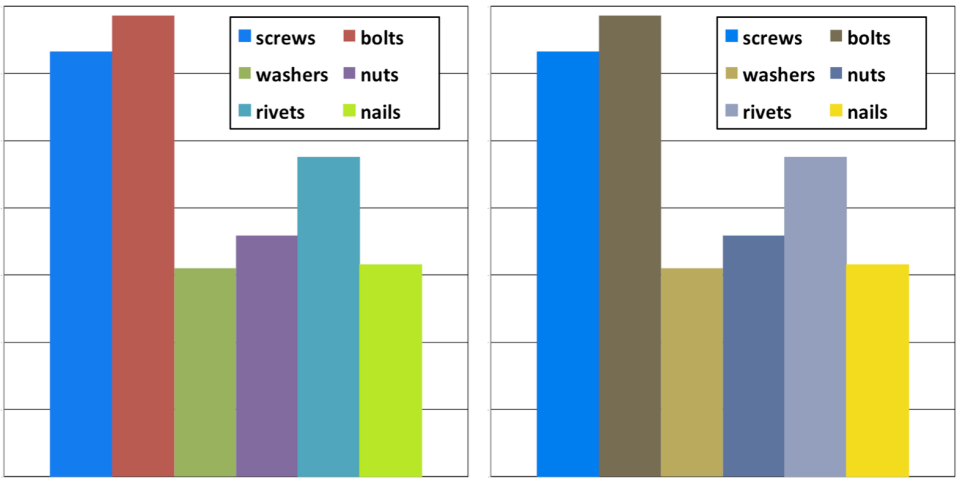


Figure 7.8: Left: the recoloured version of the original image. Right: the recoloured chart as seen by someone with protanopia. Simulation courtesy `www.vischeck.com`

Although the results presented in Figure 7.8 show improved differentiability for someone with protanopia, it is clear that this technique for generating the recoloured image is limited. One issue that may arise is how colours along edges are handled, in particular varying shades of a dominant colour introduced through anti-aliasing. These colours are unlikely to be identified as representative colours, so will remain in the recoloured image, even though the original representative colour may have been changed (sometimes dramatically). This results in anti-aliasing colours that do not match the original intent of the anti-aliasing.

To address this, the simple direct copying of the current solution can be extended to modify these colours during copying. For each pixel colour that is not a representative colour, the most perceptually-similar representative colour can be found. If this nearest representative colour was modified (has a new replacement colour), then the CIE L*u*v*

colour space 3D vector between the representative colour and the replacement colour can be determined. The pixel under consideration can then have its colour translated in CIE L*u*v* colour space the same direction and magnitude of this vector (but from the original pixel's colour), resulting in the non-representative colours being shifted appropriately.

7.3 Evaluation of ICDRecolour

To evaluate ICDRecolour, a user study was conducted to compare the performance of ICDRecolour with two existing dichromatic recolouring tools: the Daltonize system available on the WWW (`vischeck.com/daltonize`) [30], and three versions of Kuhn's recolouring tool [90]. Kuhn's tool was chosen because it represents the state of the art in recolouring, and also attempts to preserve naturalness as much as possible. ICDRecolour also maintains the maximum number of original colours, making Kuhn's a reasonable comparator. These systems were compared using three different CVD-inducing situations, with two colour tasks, for individuals with and without inherited CVD.

The study was designed to determine whether the SSM+ICD approach can improve on the state of the art in recolouring – that is, whether our new recolouring tool allows people both with and without inherited CVD to better differentiate colours in a variety of CVD-inducing situations (situationally-induced CVD).

7.3.1 Methods

Participants

I recruited 21 volunteers (15 male, mean age 31.0 years) from the community – nine with inherited CVD (8 male, 32.2 years), and twelve with no CVD (7 male, 30.1 years). All participants were screened for CVD using the HRR Polychromatic Plates [50]; of those with inherited CVD, five participants had deuteranopic/deuteranomalous vision, three had protanopic/protanomalous vision, and one had unclassified red-green CVD.

CVD-Inducing Situations

A control condition and three types of situationally-induced CVD were used in the study. I chose situations that were likely to be commonly encountered in the real world, and that were representative of different colour vision problems.

- *Normal (control) view.* Colours were presented to the participant with no added situationally-induced CVD.
- *Tinted glasses.* Participants wore yellow-tinted glasses while performing the matching task; this had the effect of altering the hues of the colours on the screen. These glasses are worn for certain sports and by people who experience glare when driving at night.
- *Darkened monitor.* For this situation, the monitor's settings were adjusted to reduce the contrast and brightness of the colours produced. Brightness was set to 50%, contrast to 0%, and the red, green, and blue primary gains were set to zero. This condition reduces luminance differences between colours (common in bright sunlight).
- *Broken monitor (no red).* This situation simulated a monitor that cannot display red. All colours displayed on the screen were programmatically altered to have their red channel values set to zero.

Recolouring Tools

I compared six recolouring schemes: a control condition with no recolouring, ICDRecolour, Daltonize, and three variations of Kuhn's recolourer.

- ICDRecolour was configured as described above. This tool uses an individualized model of colour differentiation, so the participant performed a calibration procedure (Chapter 5) to generate the model in each situation. To reduce the effect of calibration input errors, participants perormed three calibrations and the median values were used for calibration. Three calibrations took 6.84 minutes, on average.
- For the Daltonize tool, the original colours were submitted as an image to the Daltonize website with the following settings (guided by the site's FAQ): red-green stretch factor

was set to 1.3 for all situations, blue-yellow correction was set to 0.2 for all situations, and luminance correction was set to 0.0 for the isoluminant situation-specific tasks, and to -1.3 for the other situation-specific tasks and the Excel task (described below). A red-green stretch factor greater than 1.0 causes the Daltonize recolouring tool to emphasize the perceptual differences between reds and greens. A blue-yellow correction greater than 0.0 causes Daltonize to map reds to blues and greens to yellows – a common recolouring strategy for existing recolouring tools. A luminance correction of 0.0 results in no luminance remapping (colours maintain their luminance); a negative luminance correction causes reds to be made darker and greens to be brighter. Coupling the blue-yellow and non-zero luminance factors used in the study, reds become bluer and darker and greens become yellower and brigher.

- For Kuhn recolouring, the original colours were submitted as an image to Kuhn's recolouring tool [90]. As Kuhn's tool is a dichromatic recolouring system, it requires the type of dichromatism (protanopia, deuteranopia, or tritanopia). I therefore created three versions of the Kuhn recolourer, one for each type (referred to below as Kuhn-P, Kuhn-D, and Kuhn-T).

Task and Apparatus

A custom-built Java application presented a matching task in which the user had to click on a single colour that matched a given cue colour (see Figure 7.9). Two tasks were developed that used two different colour sets: a situation-specific set of nine colours, and the first fifteen colours from the Excel 2011 chart colour set (Mac version 14.1.3). The monitor used was an ASUS VH242H 24-inch LCD at 1920x1080 resolution.

The study system presented the participant with a 3x3 grid (situation-specific task) or 3x5 grid (Excel task) of coloured squares. The arrangement of colours in this grid was randomized, and one of the grid colours was presented as the cue colour on the right of the screen. To reduce the contextualizing effects of colour perception (where adjacent colours influence the perception of a colour), all colours were presented on a neutral white background with sufficient gap between colour squares to remove contextualizing effects. The participant's task was to click on the grid square that matched the cue colour. Each colour in every nine-

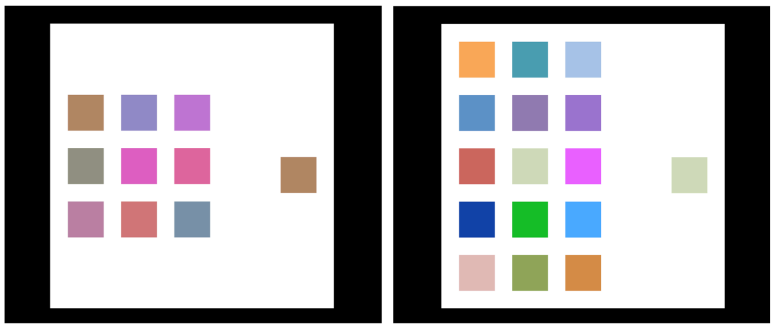


Figure 7.9: Study tasks: nine situation-specific colours (left) and fifteen Excel colours (right), after recolouring. Participants clicked on the square that most closely matched the cue colour at right.

colour and fifteen-colour set was chosen once to appear as a cue colour. Whether a participant clicked the correct square (match) and the time it took a participant to make a selection were recorded by the study system.

The study ran in a controlled environment with standard office lighting on a Windows 7 PC with a 24-inch 1920x1080 LCD monitor. The darkened monitor situation was carried out on a second (identical) computer and monitor, with the display pre-set to the appropriate values for the condition (see above).

Colour Sets for Situation-Specific and Excel Tasks

Each situation-specific task used a different colour set, chosen to highlight difficulties induced by the situation. The Excel task used the same colours for each situation, and was chosen to explore aspects of colour use in a commonly-used chart generation application.

Normal situation. Nine isoluminant colours were selected to focus on the ability of the recolouring tools to deal with hue and saturation differences (rather than luminance). Healy suggests that a maximum of seven isoluminant colours should be used in a visualization [53]; I added two more to increase the difficulty of the task for the recolouring tools. I chose the nine colours from the CIE L*u*v* isoluminant plane with luminance of 53.3 (obtained using default sRGB transforms (see Appendix A). Eight colours were taken from equally-spaced points around a circle of radius 100 on this plane, centred at the approximate center of the plane (Figure 7.10). These eight plus the central colour made up the set of nine.

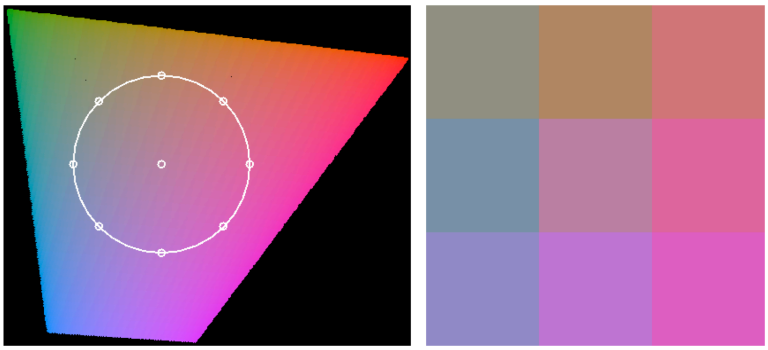


Figure 7.10: Colours used in 'normal' situation. The isoluminant circle used to generate colours (left); the resulting colours (right).

Tinted glasses. Here, the same procedure was used, but colours were chosen from the blue corner of the isoluminant plane. Blue colours were chosen because yellow glasses filter out the blue end of the spectrum, so colours that vary in their 'blueness' will be difficult to tell apart (Figure 7.11).

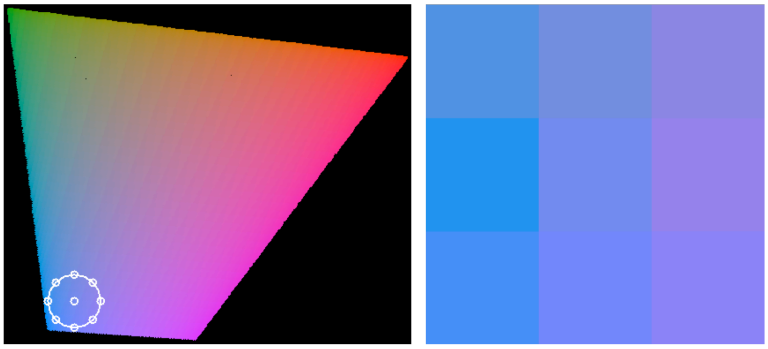


Figure 7.11: Colours used in 'tinted glasses' situation. The isoluminant circle used to generate colours (left); resulting colours (right).

Broken monitor – no red. Nine colours that varied only in red channel value were chosen by fixing the green and blue channels at 127 (out of 255), and the red value was set to one of 25, 51, 76, 102, 127, 153, 178, 204, or 229 (the colours before removing red are shown in Figure 7.12, left). This resulted (after red removal) in nine identical colours, simulating an extreme colour discrimination situation.

Darkened monitor. I chose nine achromatic RGB colours (i.e., r = g = b) equally spaced in the RGB colour space. Achromatic colours were chosen because the settings used in the darkened monitor situation affect the luminance range of the monitor without affecting hue and saturation, so colours that vary only in brightness were used (Figure 7.12, right).

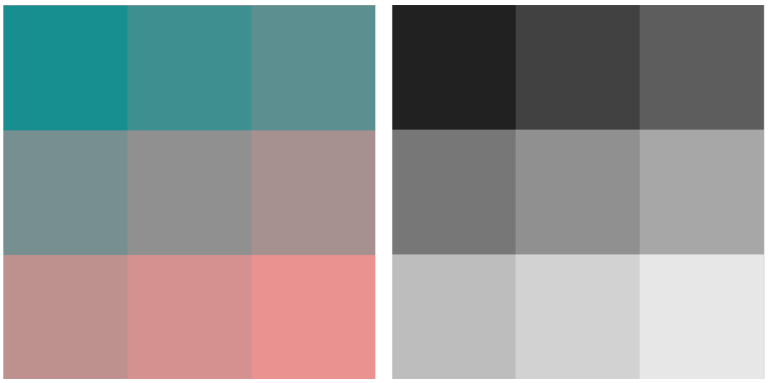


Figure 7.12: Red-channel-varying colours used in the broken monitor condition (left). Brightness-varying achromatic colours used in the darkened monitor condition (right).

Excel task. The second task used the first fifteen chart colours from Excel 2011 for Mac (version 14.1.3). Although fifteen categorical colours in a visualization is an extreme case, I chose these to introduce enough colour differentiation problems such that the tools would actually perform some recolouring. These colours were the same for all situations (see Figure 7.13).

Figure 7.13: Fifteen Excel colours used in all situations for evaluating ICDRecolour.

Study Design and Procedure

The study used a 4x6x2x2 mixed factorial design with three within-participants factors (Situation, Recolouring tool, and Task), and one between-participants factor (presence of inherited CVD). The levels of these factors were:

- *Situation*: Normal, Tinted glasses, Dark monitor, Broken monitor (no red).
- *Recolouring tool*: No recolouring, ICDRecolour, Daltonize, Kuhn-P, Kuhn-D, Kuhn-T.
- *Task*: Situation-Specific, Excel.
- *Inherited CVD*: CVD, non-CVD.

Dependent variables were accuracy (the proportion of trials with a correct colour match) and the time taken to respond. Participants carried out tasks in each situation type (Latin square counterbalanced). In each situation, the participant performed an ICD calibration to capture their colour differentiation abilities for that specific situation. Participants then carried out colour-matching tasks as described above. With three within-participants factors, and either nine or fifteen trials per task, there were 4x6x(9+15) = 576 trials in the study. At the end of the session, participants filled out a demographics questionnaire (see Appendix C). The entire study (including HRR test) took less than 90 minutes.

The instructions presented to participants for the calibration are shown in Figure E.3. Based on observations from the evaluation of ICD-2 (Section 5.3.1), the time limit for the display for each gapped circle stimulus was limited to two seconds. During the study for ICD-2, participants were allowed as much time as they wanted to report the location of the gap. This resulted in some participants taking an extraordinarily long time to report responses for some stimuli, doing some atypical things to help see the circle (e.g., holding their eyes closed and quickly opening and shutting them again, moving their head around to change the viewing angle of the screen). As these behaviours should not have to be employed in typical colour use situations in digital environments (e.g., for discerning the categories of a visualization apart), allowing this behaviour would result in models that suggest the colour differentiation abilities for a user are better than they actually are. To prevent these behaviours, each gapped circle was displayed for two seconds. Balancing speed and accuracy, participants were also encouraged to be careful reporting their observations – the instructions contained a warning stating that there was no 'undo' function, so the participant should be careful.

The instructions presented to participants for the colour-matching task are shown in Figure E.4. Once again, a warning about no 'undo' functionality was used to help participants

carefully report their responses, but with this task no explicit time limit was given. This was because the task had a varying amount of difficulty; some tasks would be very easy (e.g., the reference colour was clearly distinct within the grid of colours), and other tasks would be quite time consuming (e.g., they would require a careful linear scan of every colour in the grid to identify the reference colour). As a result, an explicit time limit was not imposed. To help participants to report their responses as quickly as possible, the number of responses they would have to provide during the study (576) was given verbally at the beginning the experiment, and a visual progress bar was shown between trials.

7.3.2 Results

Analysis of Match Accuracy

To look for effects of Situation, Recolouring tool, Task, and CVD on match accuracy, an omnibus 4x6x2x2 RM-ANOVA was conducted. The first three factors showed main effects (Situation: $F_{3,57}=71.1$, $p<0.001$; Recolouring tool: $F_{5,95}=81.7$, $p<0.001$; Task: $F_{1,19}=230.1$, $p<0.001$). No difference was found between CVD/non-CVD groups: $F_{1,19}=0.16$, $p=0.69$).

The significant effect of Situation was expected because some of the situations presented substantially easier colour matching tasks than others (e.g., the 'Normal' situation presented far fewer difficulties than the 'Dark Monitor' situation). Likewise, the significant effect of Task was also expected because there were large differences in colour-matching difficulty between different tasks (e.g., the situation-specific tasks were chosen to emphasize the colour-differentiation problems caused be different situations). No difference found between CVD and non-CVD participant performance is of interest, and suggests that recolouring works in a similar fashion for both of these groups, at least for the colours chosen for this experiment.

My main interest is in the differences between recolouring tools. Overall results for the different recolouring tools are shown in Figure 7.14; as can be seen in the figure, ICDRecolour allowed participants to find at least 20% more correct matches than the other tools.

To explore the factor interactions, post-hoc Bonferroni-corrected (15 comparisons; adjusted alpha=0.003) paired *t*-tests between the individual recolouring tools were carried out. These comparisons showed that accuracy with ICDRecolour was significantly higher than

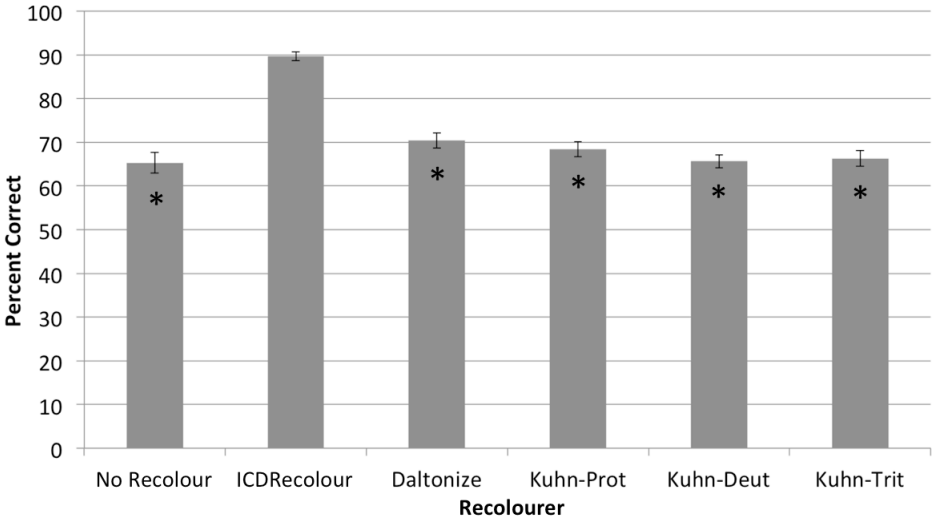


Figure 7.14: Overall mean accuracy of participants, by recolourer. Error bars indicate standard error. Asterisks indicate pairwise significant differences (15 comparisons; Bonferroni-corrected alpha=0.003) versus ICDRecolour.

any of the other tools (all $p<0.001$). None of the other comparisons showed any differences – in particular, none of the remaining recolouring tools were significantly better than no recolouring at all.

There were, however, several interactions among the factors, and so these overall results must be interpreted in light of these additional analyses. First, no four-way interaction (Situation x Task x Recolouring tool x CVD/non-CVD) was found ($F_{15,285}=1.422$, $p=0.136$), but two significant three-way interactions were identified:

- Situation x Recolouring tool x Task ($F_{15,285}=12.3$, $p<0.001$). This interaction highlights the differences between recolouring tool performance in different situations with different tasks. Although the Excel task was the same for every situation, the nine-colour situation-specific task was customized to each situation with some situations presenting more differentiable colours than others (e.g., in the 'broken monitor' situation with 'no recolouring', the colours were completely identical). Figure 7.14 illustrates the consistent performance of ICDRecolour, but other recolouring tools performed well in some situations (e.g., the Kuhn-Protan recolouring tool is well-suited to the broken monitor situation), as shown below in Figures 7.15–7.18. As a result, different situations presented different levels of difficulty, which was compounded by the situation-specific

colour tasks; most recolouring tools performed variably across different situations, but ICDRecolour consistently performed well.

- Situation x Task x CVD ($F_{15,285}$=5.512, p<0.01). Continuing with the connection between situations and the nine-colour situation-specific task, participant colour-matching success was also influenced by whether the participant had CVD. Some of the situation-task combinations resulted in situationally-induced CVD that was not aligned with the red-green inherited CVD of some participants. As a result, these situations presented additional challenges for participants with inherited CVD, as there was reduction in colour discrimination abilities along both the blue-yellow colour axis or the luminance axis (because of the situationally-induced CVD) and the red-green colour axis (because of the inherited CVD). Participants with typical colour vision were not as severely restricted, resulting in higher colour matching accuracy.

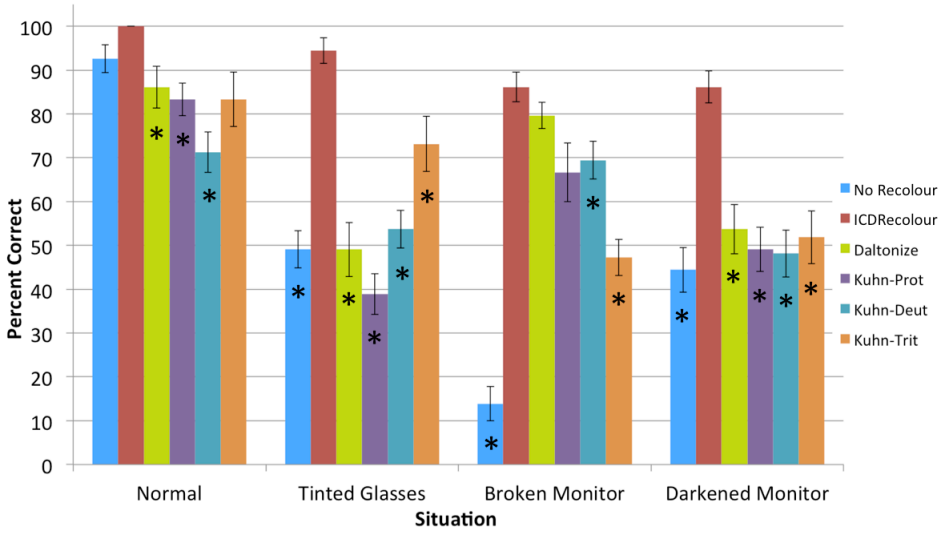


Figure 7.15: Nine-colour situation-specific task accuracy for participants without inherited CVD (in Figures 7.15–7.18 error bars show standard error; asterisks indicate pairwise significant differences (five comparisons; Bonferroni-corrected alpha=0.01) versus ICDRecolour within a situation.

In addition to the three-way interactions just described, there were also four significant two-way interactions, detailed below and illustrated in Figures 7.15–7.18:

- Situation x Task ($F_{3,57}$=40.4, p<0.001). The different situations had markedly different effects on the two tasks, as seen by comparing Figures 7.15 and 7.16 (or 7.17 and 7.18

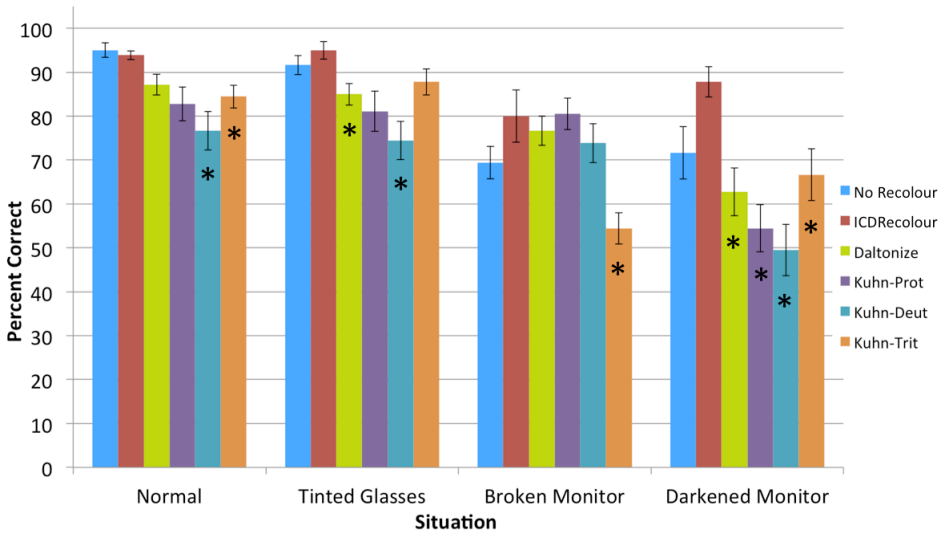


Figure 7.16: Fifteen-colour Excel task accuracy ± s.e. for participants without inherited CVD.

for CVD participants). For example, tinted glasses had a much stronger effect on the situation-specific task than the Excel task. This is again because the situation-specific task was designed to highlight problems induced by the situation. The Excel colours contain enough colour variation to reduce the effect of the situation.

- Situation x Recolourer ($F_{15,285}$=28.0, p<0.001). Different recolouring tools performed very differently in the different situations, because the CVD induced by some situations aligned well with some forms of inherited CVD. For example, a display with no red produces colours very similar to what an individual with protanopia experiences, so the Kuhn-Protan recolourer does well.

- Recolourer x Task ($F_{5,95}$=24.9, p<0.001). Similarly, different recolouring tools performed differently on different tasks, e.g., Kuhn-Deutan performed poorly in the tinted glasses situation (Figures 7.17 and 7.18). This recolourer assumes deutaneropic CVD, so it recolours by transferring red-green variations to blue-yellow variations. However, the yellow tinted glasses reduce perception of blue-yellow axis variations, so this recolourer performed poorly for the nine-colour task. In the fifteen-colour task, this transition from red-green to blue-yellow still occurred, but the colours contain variations in luminance which provide redundant cues that allow better colour differentiation.

- Recolourer x CVD ($F_{5,95}$=2.83, $p<0.05$). There was a small difference in performance between CVD and non-CVD participants with the different recolouring tools. CVD participants performed better overall than non-CVD people when using the Kuhn-Protan and Kuhn-Deutan tools. This is because these tools are optimized for these participants' inherited CVD, performing a 'worst case' recolouring which significantly aided these participants, but overly restricted the replacement colour set, limiting participants with typical colour vision.

Figures 7.15–7.18 show results for different recolouring tools in terms of the different Situations. I also carried out separate follow-up *t*-tests in each combination of factors shown in the charts, to compare ICDRecolour with the other tools; in the figures, asterisks (*) are placed on individual bars where there was a significant difference between that tool and ICDRecolour within the given situation.

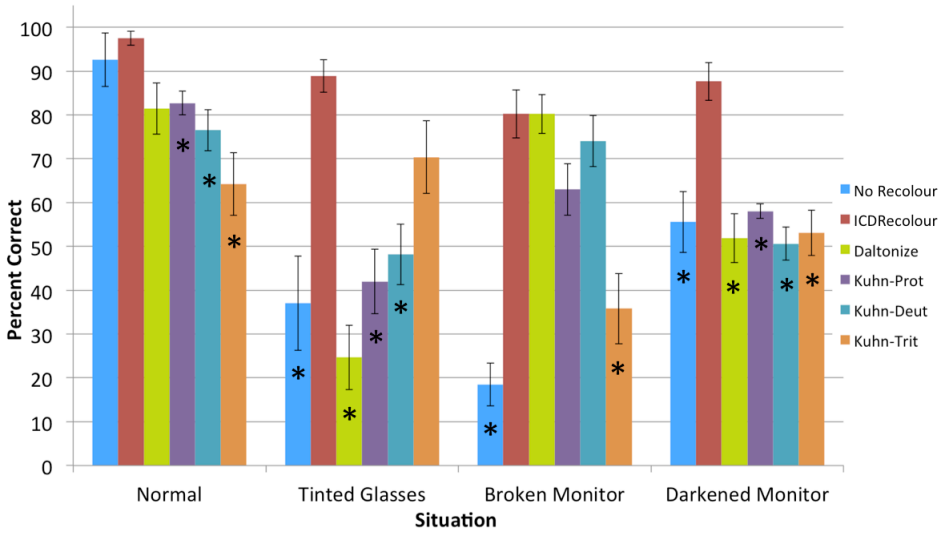


Figure 7.17: Nine-colour situation-specific task accuracy ± s.e., for participants with CVD for each recolouring tool and situation.

Analysis of Response Time

To explore the effects of the main factors on response time, a second 4x6x2x2 RM-ANOVA was carried out. Main effects were found for factors Situation ($F_{3,57}$=6.88, $p<0.001$) and Recolouring tool ($F_{5,95}$=17.8, $p<0.001$), but not for Task ($F_{1,19}$=0.12, $p=0.74$) or CVD

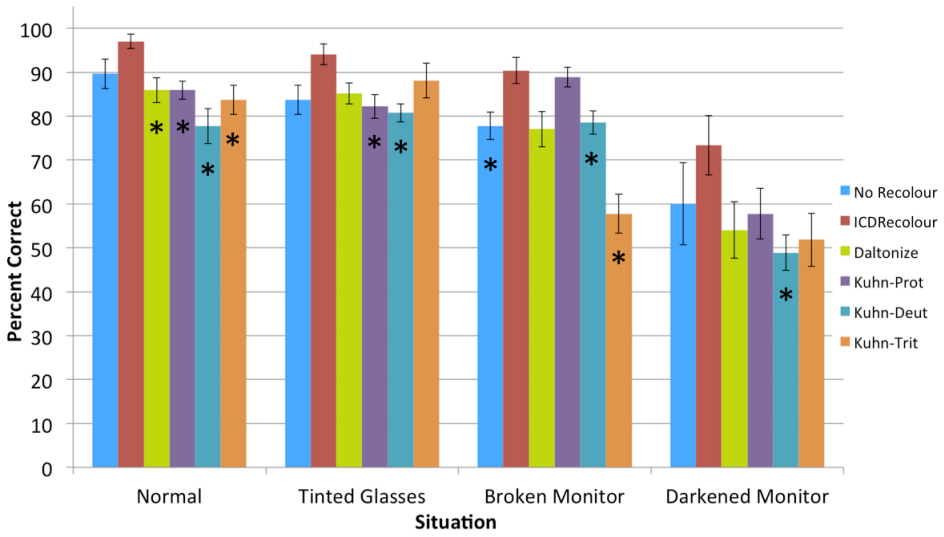


Figure 7.18: Fifteen-colour Excel task accuracy ± s.e., for participants with CVD for each recolouring tool and situation.

($F_{1,19}$=1.52, p=0.23). There were also significant interactions between Situation and Recolouring tool ($F_{15,285}$=4.71, p<0.001), and between Situation and Task ($F_{3,57}$=14.1, p<0.001).

Figure 7.19 shows the mean overall response times for the different recolouring tools. Post-hoc Bonferroni-corrected (15 comparisons; adjusted alpha=0.003) paired *t*-tests between individual recolouring tools showed that ICDRecolour allowed significantly faster response time than all of the other tools (all p<0.001), and that there were no significant differences between any of the other tools (or the 'no recolouring' case).

Figures 7.20–7.21 show response times (seconds) by recolourer and situation to show differences indicated by the interactions.

7.4 Discussion

This evaluation provided three main results:

1. Accuracy of ICDRecolour is 20% higher than existing recolouring tools.

2. Selection time for ICDRecolour is almost two seconds faster than existing recolouring tools.

3. Increased accuracy and reduced selection time of ICDRecolour is consistent across a

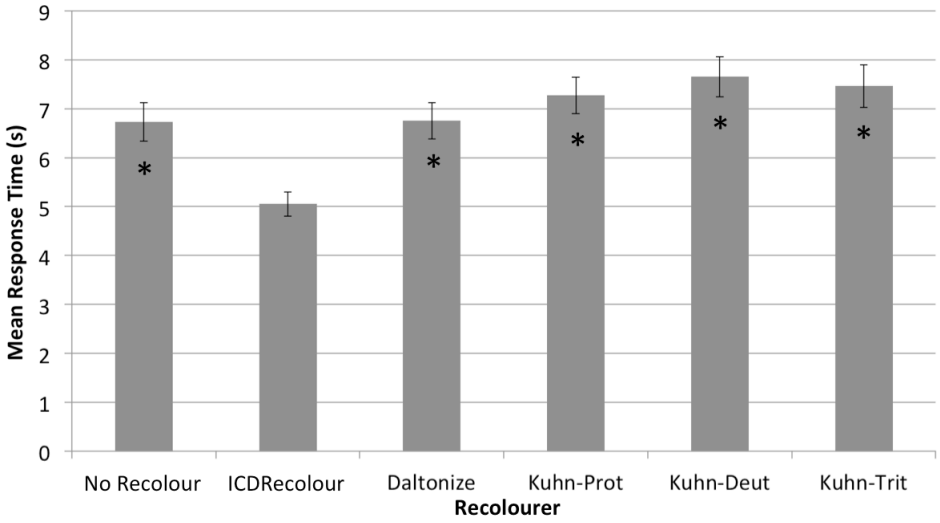


Figure 7.19: Mean response times (seconds) by recolouring tool. Error bars indicate standard error. Asterisks indicate pairwise significant differences (15 comparisons; Bonferroni-corrected alpha=0.003) versus ICDRecolour.

variety of situations, colour sets and users.

7.4.1 Explanation of Main Results

ICDRecolour achieves consistently higher accuracy and reduced selection time because the individualized model of colour differentiation more accurately represents the colour differentiation abilities of the user. By capturing the abilities of the user with a performance-based in-situ calibration, any factors that influence colour abilities are automatically encoded into the model. This allows accurate predictions to be made regarding the user's ability to differentiate between two colours. These predictions are the heart of ICDRecolour's recolouring algorithm, allowing both the accurate identification of problem colours as well as the selection of sufficiently differentiable replacement colours.

7.4.2 Generalizing the Results of the Evaluation

Although our study examined specific types of situation-induced CVD, our findings should extend to real-world situations that are similar to our experimental conditions.

- *Hue variations.* Any condition or situation that causes the perception of hue to change (e.g., cataracts, retinopathy, coloured lighting, tinted sunglasses, uncalibrated hard-

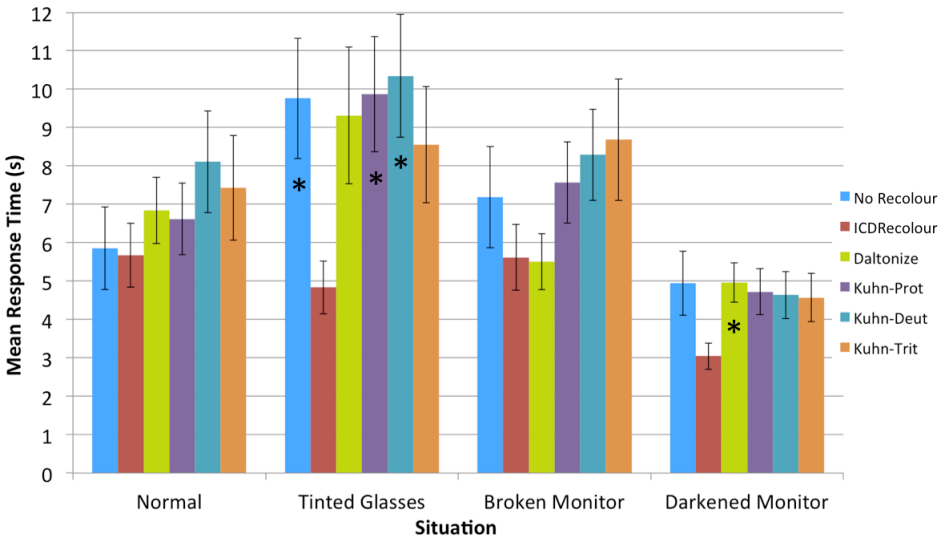


Figure 7.20: Response time (9-task), by recolourer and situation. Error bars represent standard error. Asterisks indicate pairwise significant differences (five comparisons; Bonferroni-corrected alpha=0.01) versus ICDRecolour within a situation.

ware) can benefit from ICD-based recolouring. Our evaluation shows that ICD-based recolouring is applicable to a wide range of situations (including extreme situations), and should significantly improve the differentiability of colours on digital displays in these situations.

- *Luminance challenges.* Any situation that results in reduced luminance range and contrast (e.g., using a mobile device in bright sunlight, using power-saving reduced-brightness displays) can also benefit from ICD-based recolouring. The study results from the darkened monitor situation show great promise for assisting users in these low-contrast situations.

7.4.3 Extending ICDRecolour

Even though the SSM approach achieves higher accuracy and reduced selection time, there are a number of changes that can be made to improve it further. These include preserving the naturalness of the original image, and extending it to other uses of colour.

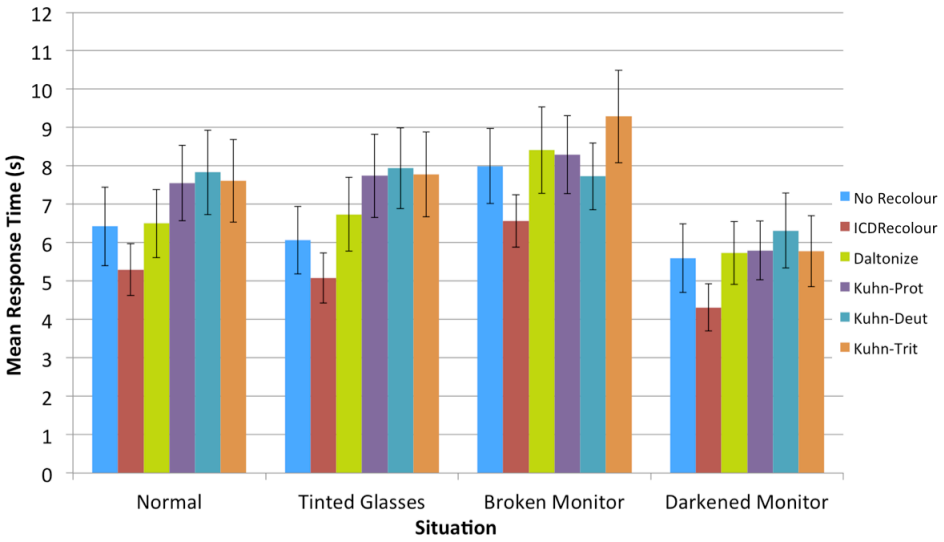


Figure 7.21: Response time (15-task) by recolourer and situation. Error bars represent standard error.

Preserving Naturalness

Kuhn's recolouring tool [90] attempts to maximize the similarity between the original image and the recoloured image. This is achieved by restricting each replacement colour to the luminance of its respective key colour, and by attempting to maintain the visually-perceived difference between key colours and between replacement colours. The existing recolouring algorithm for ICDRecolour is quite naive by comparison, but can be extended to incorporate the luminance and visually-perceived difference consistency as in Kuhn' s method. To do this, the replacement colour set can be divided into a replacement colour set for each key colour, such that each set is isoluminant with its respective key colour. Multiple recolour mappings can then be generated, selecting the recolouring that maximally preserves the visually-perceived differences in the original image.

Other Uses of Colour

Current situation-specific models only encode basic colour differentiation abilities. Although differentiation is the central aspect of many colour uses, there are additional uses of colour that go beyond differentiation, such as popout (using colour for pre-attentive processing) and highlighting (using colour to draw attention) in information visualizations [146]. The current

version of ICDRecolour makes no effort to preserve these perceptual properties through a recolouring, but the SSM approach is extensible to other colour uses in visualization, allowing SSM-based recolouring to preserve perceptual properties in addition to differentiability.

Adjustable Recolouring

SSMs for colour differentiation have a tuneable parameter (called the limit offset) that can be used to adjust how conservative the model is in its predictions of colour differentiability. Increasing the limit offset causes colours that are actually differentiable to be predicted as not differentiable. This parameter can be used in ICDRecolour to adjust the number of original colours that are modified, and to adjust the differentiability of the replacement set of colours. By increasing the limit offset, more of the original colours will be flagged as not differentiable, resulting in their subsequent recolouring. An increased limit offset will also cause colours that are more differentiable to be used as replacements. When the limit offset is small, a minimal number of colours will be modified; this can help maintain the naturalness of the original image. With a large limit offset, many colours will be modified, and the resulting replacement colours will be more differentiable.

In the next chapter, the results from the development of the SSM-based ICD (Chapters 4–6 and the ICD-based recolouring tool (this chapter) will be discussed.

Chapter 8

Discussion

In this chapter, I will discuss implications, limitations, and extensions of my dissertation research work. This chapter addresses the thesis problem presented in Chapter 1 and explores how the solution presented in this dissertation can be improved and extended.

8.1 Summary of Findings

The evaluations above show the main benefits of the individualized model of colour differentiation model (in this section, I will use 'ICD' to describe the combination of the four different versions):

- ICD provides accurate differentiability predictions for a wide range of CVD types that are not handled by the existing dichromatic models, and also work well for inherited CVD;
- ICD is responsive to environmental factors, but not inordinately sensitive to small changes in the physical setting;
- ICD's current calibration time (about two minutes) means that the technique is feasible, and that it can be used to generate multiple models for a user if necessary;
- ICD's current model prediction time (0.005 ms per prediction) is low enough to make real-world constraint-optimization recolouring possible without undue delays for the user;
- ICD's revised prediction mechanism provides degree-of-differentiation information that increases the range of visualizations that can be recoloured.

- Incorporated into a recolouring tool, ICD consistently improves colour match accuracy and selection time over existing recolouring tools for a broad range of situations, users, and colour sets.

8.2 Explanation of Findings

In Chapter 1, I identified four types of CVD: inherited, acquired, situationally-induced, and simultaneous. I will now explore the main findings of this dissertation in terms of each of these.

8.2.1 Inherited CVD

ICD and ICDRecolour were evaluated using data collected from participants with inherited CVD. In each case, the group of CVD participants contained some individuals with protan (long-wavelength) and some with deutan (medium-wavelength) inherited CVD, as determined by the Ishihara Plates in the ICD-1 study, and the HRR Polychromatic Plates for the ICD-2 and ICDRecolour studies.

In cases of inherited CVD, the angular orientation of the major axis of the discrimination ellipsoid aligns with the confusion lines for the type of inherited CVD the user has, regardless of whether the inherited CVD is anomalous trichromacy or dichromacy [96]. In ICD-2, discrimination ellipsoids were used to represent the user's differentiation abilities, and I found that the major axis for the ICD-2 discrimination ellipsoid for each participant aligned with his/her HRR-diagnosed type of CVD.

In addition to discrimination ellipsoid orientation, the length of the major axis also correlates with the severity of inherited CVD [11]. The HRR plates also roughly diagnose severity of inherited CVD, and in most cases the size of the ICD-2 ellipsoid aligned with the HRR-plate severity. Figure 5.7 illustrates the alignment and variation in sizes of discrimination ellipses identified by ICD-2.

These results from the ICD studies indicate that the ICD calibration and modelling components are able to detect and represent the colour differentiation abilities of individuals with inherited CVD. The results from the evaluation of ICDRecolour support this as well;

Figures 7.17 and 7.18 show an improvement in colour matching accuracy for participants with CVD in the 'Normal' condition. These figures further highlight the success of ICD-based recolouring with nearly 100% matching accuracy achieved by participants with CVD, and the substantial (often significant) improvement in matching accuracy compared to existing recolouring tools.

These findings strongly suggest that ICDRecolour is able to perform recolouring that is aligned with the specific type and severity of inherited CVD of a particular individual, regardless of the type and severity of CVD he/she has.

8.2.2 Acquired CVD

Despite acquired and inherited blue-yellow CVD having different causes, the assessment of each is performed using standard plate tests, such as the HRR Polychromatic Plates. This test does not differentiate between these two types of CVD and simply reports the presence and severity of any blue-yellow CVD in a participant. As a result of this, it is unknown whether any of the study participants had acquired CVD, however two ICD-2 participants and one ICDRecolour participant were classified with mild blue-yellow CVD. Considering that inherited blue-yellow CVD is very rare (less than 0.1% of all cases of inherited CVD [10, 20]), and acquired CVD appears to be quite common (e.g., 64% of British participants over 65 exhibited some degree of acquired CVD [25]), it is likely that these participants had acquired blue-yellow CVD. All three of these participants were over the age of 60, further supporting this hypothesis.

The discrimination ellipsoids for the two ICD-2 participants with acquired CVD both exhibited a lengthened blue-yellow axis compared to younger participants (with CVD and without), suggesting that the ICD model is able to detect and model the effects of acquired CVD.

In terms of the ICDRecolour evaluation, the 'Tinted Glasses' condition was included as an approximation of acquired CVD because the yellow tinting of these glasses induces a reduced ability to distinguish colours along the blue-yellow axis. Considering Figures 7.15 and 7.16, ICDRecolour achieved approximately 94% matching accuracy in the 'Tinted Glasses' condition, which was significantly higher than all other recolouring tools in the nine-colour task,

and higher (some significant) than all other recolouring tools in the 15-colour task.

These results suggest that the ICD is able to capture and represent acquired CVD. This ability allows ICDRecolour to provide recolouring results that provided very high matching accuracy in a condition that mimics acquired CVD.

8.2.3 Situationally-Induced CVD

When evaluating ICD-1, I included a number of environmental variations that are common causes of situationally-induced CVD (e.g., changes in ambient light levels, non-optimal monitor settings). The results from this study suggest that the ICD is able to capture and represent the influence of these variations.

This sensitivity to situationally-induced CVD is further supported in the results from the ICDRecolour evaluation. Two conditions were included in the experimental design ('Broken Monitor' and 'Darkened Monitor') to determine how well ICDRecolour improves colour matching accuracy in cases of situationally-induced CVD. The results were presented in Figures 7.15 and 7.16, and show that ICDRecolour achieved significantly better results than all other recolouring tools for the 'Darkened Monitor' condition and as good or better results than existing recolouring tools for the 'Broken Monitor' condition. ICDRecolour achieved 80-86% matching accuracy for these conditions.

These results show that the ICD is able to capture and model a broad range of situational effects. This ability helps an ICD-based recolouring tool to perform recolouring that accommodates the large population of people who experience situationally-induced CVD (e.g., when using mobile phones in bright sunlight). The 'Darkened Monitor' condition of the ICDRecolour study was designed to mimic using a display in bright sunlight, and its significant improvement over existing recolouring tools indicates that ICD shows promise for this use scenario.

8.2.4 Simultaneous CVDs

As described in Chapter 1, acquired and situationally-induced CVDs are common. An individual with inherited CVD can develop acquired CVD as he/she ages and may also expe-

rience situationally-induced CVD, leading to simultaneous CVDs. The evaluation of ICD-1 included simultaneous CVDs in the study design by gathering data from participants with CVD in situations that can induce CVD (e.g., lamp producing glare on the screen). The results from this evaluation showed that the ICD is able to capture and model the influence of simultaneous CVDs.

In evaluating ICDRecolour, the ability of the ICD to help people experiencing simultaneous CVDs was tested. Participants with CVD performed the matching tasks in the 'Tinited Glasses', 'Broken Monitor', and 'Darkened Monitor' conditions, all of which result in simultaneous CVDs (the user's inherited CVD coupled with the CVD induced by the experimental condition). In all cases, ICDRecolour performed as well or better than existing recolouring tools, and always helped the participants improve their colour matching performance over using no recolouring tool at all.

With the in-situ calibration technique used by the ICD, any factor that influences a user's colour perception can be captured. Traditional models of CVD fail when more than one factor influences colour perception, but these results show that ICD is sensitive to differences between users and to environmental variations. ICDRecolour is able to take advantage of this ability, resulting in higher colour matching accuracy than existing recolouring tools.

8.3 Is the Dissertation Problem Solved?

In Section 1.1, the problem to be addressed in this dissertation was presented: *Existing recolouring tools do not help most people with CVD.* In this dissertation, I demonstrated that the SSM-based ICD is able to capture and represent the colour differentiation abilities of individual users, and that applying the ICD to recolouring significantly improved colour matching performance for users with a wide variety of abilities in a diverse set of environments. However, three outstanding issues remain: real world applicability, ICDRecolour matching accuracy, and cases where the ICD approach fails.

8.3.1 Real World Applicability

Although ICDRecolour was able to show consistently-improved recolouring performance (20% increase in colour matching accuracy overall), the evaluation was conducted in a controlled lab environment, and not in real world situations. The difficulties that people with inherited, acquired, situationally-induced, and simultaneous CVDs encounter are not typically in a controlled lab environment, but in real world situations like trying to buy fruit at the supermarket or trying to find a subway route. Recolouring tools are just beginning to move from lab-based prototypes to real world recolouring via mobile phones [113] and mobile projector systems [1, 154], potentially expanding their applicability. As the evaluation of ICDRecolour used a wide range of users and situations, I look forward to replicating the success of ICD and ICDRecolour in more general situations.

8.3.2 Is ICDRecolour Good Enough?

ICDRecolour did not achieve 100% accuracy for every condition in the study described in Section 7.3. One possible reason for this is that the situations used in the study were difficult from a recolouring standpoint (i.e., for some users these situations did not have a recolouring that would achieve 100% accuracy). If this is the case, then ICDRecolour should do better in less extreme situations, and indeed that is what is seen (e.g., ICDRecolour achieved near 100% performance for every 'Normal' situation). Even in these less extreme cases, however, perfect accuracy is unlikely – for example, even in the control case with non-CVD participants, people still made a few errors due to the difficulty of the task or selection mistakes.

8.3.3 Cases Where the ICD Approach Fails

As highlighted above, there are cases where the ICD approach did not result in 100% matching accuracy. This suggests there are cases where the ICD approach will not succeed in expanding the application of recolouring tools to most users with CVD.

Recolouring tools have an inherent limitation that prevents recolouring from being a complete solution. When the number of differentiable (or salient) colours perceivable by the user is less than the total number of colours that need to be distinct, recolouring will not

be able to find a suitably-differentiable set of replacement colours. Two extreme cases help illustrate this. First, if a visualization calls for a large number of differentiable colours (e.g., 100 milliion), this can easily overwhelm the colour perception capabilities of even typical users who can see approximately ten million distinct colours [153]. A second case involves users who have no vision at all. A hypothetical calibration involving a user who is blind would result in him/her responding with 'not differentiable' for every single stimuli, resulting in a very large discrimination ellipsoid. A very large discrimination ellipsoid will prevent ICDRecolour (or any other recolouring tool) from finding a single pair of differentiable colours.

Another case where the ICD (and the SSM approach in general) will not succeed is in environments that change the ambient lighting very frequently in an unpredictable manner (e.g., a dance club). In these situations, the current two-minute calibration for ICD is not fast enough, as the lighting can change potentially multiple times a second. This situation, although extreme, highlights several areas for improving ICD, which I will cover next.

8.4 Improvements to ICD

Some of the improvements to ICD presented in this dissertation (reduced calibration time, improved prediction speed, and improved precision) can be extended to improve ICD further.

8.4.1 Reduce Calibration Time

Calibrating ICD currently requires measuring eight differentiation limits. Six of these lie on the three confusion lines that intersect the background colour, and two lie along the luminance axis, and can be considered as 'pairs' of differentiation limits – one pair for the protan confusion line, one for the deutan, etc. As CIE L*u*v* is a perceptually-uniform colour space, each differentiation limit in a pair actually measures the same value (differentiation limit for the background colour along a confusion line or luminance axis). As a result, both measurements in a pair are not needed; one of the limits along a line or axis can be measured using the calibration and simply duplicated by placing it on the opposite side of the background colour in CIE L*u*v* colour space to generate the other value in a pair. This allows the calibration to be shortened to finding four differentiation limits instead of eight.

As a result, the calibration time for ICD can be halved to about one minute, but does this influence the accuracy of the model? Using the study data from the evaluation for ICD-2, I performed a brief analysis comparing the prediction accuracy for ICD-2 to this revised ICD (called ICD-5 below), and found that accuracy for non-CVD participants increased slightly, and the accuracy for participants with CVD decreased slightly (see Figure 8.1). Pairwise *t*-tests between ICD-2 and ICD-5 accuracy results did not show any significant differences (non-CVD p=0.14, CVD: p=0.19). These results warrant follow-up, which I will pursue in the future.

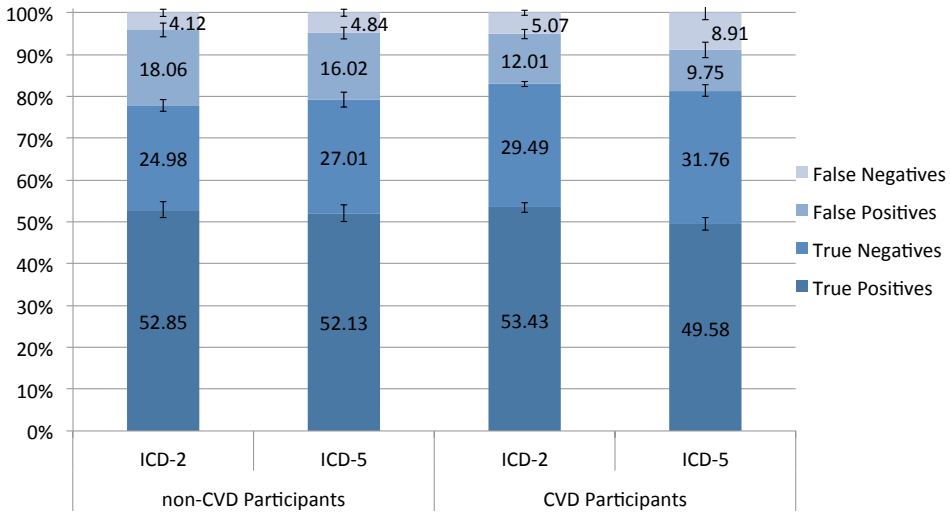


Figure 8.1: Model accuracy ±s.e., by CVD status and model (true positive rate + true negative rate = overall accuracy).

8.4.2 Eliminate Calibration Completely

No matter how little time a calibration takes, a user still has to task switch to the calibration in order to perform it (at least for active calibrations). In this section, three approaches for reducing or eliminating calibration are explored.

Sensing the Environment: Most mobile phones are equipped with sensors for detecting location and ambient light. If calibrations are stored locally along with readings from relevant sensors (e.g., location, ambient light levels), the calibration can be loaded automatically the next time the user enters the same location or an environment with similar light levels. This

would allow a single calibration to be used repeatedly over time and in different locations, reducing the frequency of calibration for the user.

In addition to using a pre-existing calibration for a situation that the user has been in before, it should be possible to adjust an existing calibration for a new situation. In particular, if two calibrations are available that are for situations similar, but not an exact match, to the current situation (e.g., calibrations for slightly brighter and slightly darker ambient light levels), then a calibration for the current situation can be generated by interpolating between the existing models. How to interpolate or blend existing calibrations (or models) is an open research question for situation-specific modelling, but the study results for ICD-1 (Section 4.3.2) suggest that the ICD model is sensitive, yet robust, to small environmental variations, thereby allowing models to be generalized beyond the exact situation and user for which they were created.

Even in the situation where an exact model does not exist and similar models are not available, environmental sensing could still be used as a way to tell the user that a new calibration would result in better performance.

Community of Models: Another method for increasing the number of available models is to share models with a community, and find ways of selecting others' models that fit well with the current user and the current environment. For example, if there is no ICD model that fits the user's current situation, a model could be retrieved from the community that closely matches the user and his/her environment, enabling more specific predictions without performing a calibration. Similar to that discussed above in *Sensing the Environment*, it may be possible to interpolate between existing models to try and fit a particular environmental situation. It is not yet known whether there is too much inter-user variation to prevent sharing another person's ICD model, but this approach has worked well in other individualization settings (e.g., movie recommendations) where there is a large database of users.

To achieve this, additional work is required to determine the set of properties that should be used to organize such a community of models (e.g., ambient light level, type and severity of inherited CVD). In addition, it would be very beneficial to group human members within this community into collections of people with similar inherited and acquired CVDs, as these are much less subject to variation over time in comparison to situationally-induced CVDs. This

can be extended to a full community with concepts such as trust, ratings, and reputation, where 'friends' who have similar colour differentiation problems can be found

Passive Calibration: In Section 3.4.1, a distinction was made between active calibrations and passive calibrations. Active calibrations require the user to purposefully engage with the calibration task, whereas passive calibrations collect data while the user attends to some other task. One approach for moving ICD to utilizing passive calibration is to incorporate the calibration into a game [35]. Instead of the user executing a calibration, the user gets to play a game which should be more enjoyable and rewarding. I have developed one colour differentiation calibration game [35] (Figure 8.2) based on the ICD-1 calibration, but intend to explore it more.

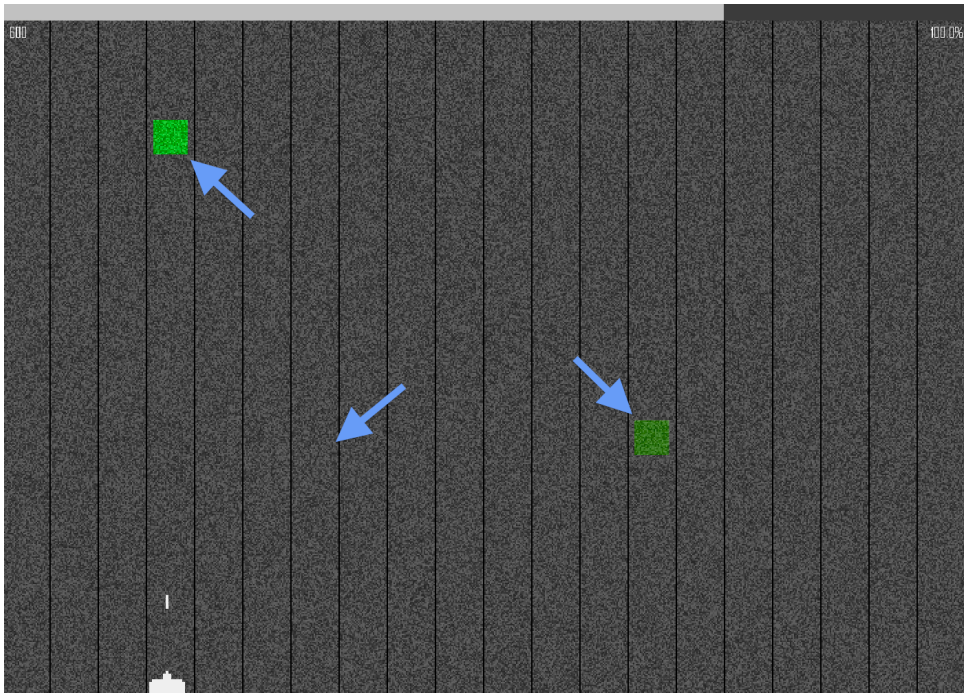


Figure 8.2: An ICD calibration game in which the user shoots falling targets and is penalized for missing targets. Blue arrows are not part of the game, but highlight a highly visible target, a less-visible target, and an invisible target. If a user shoots a target, the game assumes that the user can differentiate the target from the background. When a user misses a target, the game assumes that the user cannot differentiate the target from the background.

One additional benefit of calibration games is that the user may choose to engage in the calibration for longer than the ICD requires. This enables the collection of more calibration data, allowing the modelling performance to be improved. These improvements are discussed

below (Section 8.4.4).

Note that the three approaches for eliminating calibrations just described are not mutually exclusive; a community of models expands the set of calibrations or models that can be drawn from to find an appropriate model, sensing the environment helps detect which model in the community of models is the most appropriate, and passive calibration should increase the number of models in the community.

8.4.3 Reduce Prediction Time

Although Chapter 6 significantly reduced the time to perform an ICD differentiation prediction, it is clear that recolouring tools can benefit from even more reductions in prediction time. This section discusses three ways achieve this.

Parallelization: The final version of the ICD performs predictions by translating and rotating the basic calibration ellipsoid to fit it to one of the input colours, and then checks the proximity of the second colour parameter to the adjusted ellipsoid. This is purely a read-only set of actions (no state changes occur), so there is nothing preventing full parallelization of the prediction functionality. Suppose a recolouring tool is searching for a replacement colour set for n representative colours. This may require n^2 predictions per iteration of the constraint optimization function. As the predictions do not influence the internal state of the model, and do not influence each other, they can be executed in parallel using n^2 different instantiations of the ICD model. This has the benefit of reducing the time to perform recolouring without actually modifying the current ICD.

Caching: Constraint optimization in recolouring tools often requires a large number of differentiation predictions to be performed where at least one of the parameters does not change (e.g., it needs to find how different every other colour is from a particular colour). In addition, candidate replacement colours may take many iterations before they are changed. When a small set of colours make up a majority of the parameters to a prediction, the translated and rotated ellipsoids for these colours can be stored for later user. When a differentiation prediction is made with one of these colours, its translation and rotation steps can be avoided, thereby saving time.

Pre-generating Ellipsoids: A third approach to reducing prediction time is to generate the discrimination ellipsoid for every possible colour (e.g., every one of the 2^{24} RGB colours). As the discrimination ellipsoid can be represented by seven floating point values (major axis length, minor axis length, height, 3D location in L*u*v* space, and angular rotation), this could easily be stored in memory (e.g., if 8-byte double-precision floating point values are used, only 128 MB of memory is needed to store every ellipsoid). When a prediction is made, the ellipsoid for one parameter is read from memory and the other parameter's location is compared to the loaded ellipsoid to generate a prediction.

8.4.4 Increase Prediction Precision

Chapter 6 presented a technique for incorporating a real-valued `howDifferentiable` prediction function into the ICD, which provides more power to recolouring tools. This extension utilizes a generic sigmoid function that is based on psychometric functions, but is not fit to each user's particular psychometric response. Additionally, the binary search approach to finding differentiation limits is very fast, but is very sensitive to errors. These two issues give rise to two additional ways to improve the precision of a prediction.

Fit Sigmoid to User: As mentioned above, utilizing a calibration game to gather differentiation limits may allow more data points to be gathered during calibration. More data points would allow the sigmoid function of the ICD to be fitted to the psychometric responses of the user by repeatedly querying the user about the differentiability of colours that lie along the psychometric function as it transitions from 0 to 1. This would allow the ICD to more precisely represent the colour differentiation abilities of the user by more clearly defining the ellipsoidal surface as well as the regions just inside and just outside the ellipsoid surface.

Reduce Sensitivity to Errors: In addition to more precisely measuring the user's psychometric function, increasing the number of measurements taken during calibration would help offset the influence of errors during calibration. For example, each differentiation limit could be measured three times, and the median of the three could be used as the final differentiation limit. Additionally, the binary search approach could be made less strict, e.g., requiring the user to 'not see' a colour difference two times before increasing the differentiability of the test colours.

8.5 Extensions to ICD

The ICD technique, and the SSM approach in general, can both be generalized to beyond extending recolouring tools to a wide range of user abilities and situations. Some extensions to the solution presented in this dissertation are discussed below.

8.5.1 Extending Colour Perception Modelling

In addition to problems with colour differentiation, the ICD also shows promise for other issues related to colour perception. Four of these explored below are colour matching, communicating with colour, simultaneous colour contrast, and interface and visualization colour selection.

Colour Matching: Colour matching is the task of finding a colour that has been identified in another part of a display (e.g., using a legend in a bar chart to look for a particular data category). Matching is directly based on differentiability, but with an added constraint – that colours must also be salient and memorable between glances at one part of the image (e.g., the chart) and another (e.g., the legend). As the evaluation for ICDRecolour used a colour matching task, it appears that the ICD approach incorporates colour matching as well, but in practice colour matching may require increasing the co-differentiability of the entire colour set, since colours that are farther apart are likely to be better remembered when there is a long gap between looking at image.

Colour Communication: Objects are often referred to using their colour when communicating with another person. For example, it is common to say "the red one" to indicate a specific object. Recolouring an image makes this kind of communication difficult; although some recolouring tools attempt to minimize the number of colours that change from the original (by preserving naturalness), the problem cannot be completely solved in this way. As an alternative, the ICD can be used to generate a simulation that allows non-CVD users to see an image in exactly the way that a collaborator with CVD sees it. Although simulation tools based on the traditional models of CVD exist, an ICD-based tool could provide a more specific model, and could also indicate how images have been recoloured. This is discussed in more detail below (Section 8.5.2).

Simultaneous Colour Effects: Perception and differentiability issues arise when colours are affected by the presence of surrounding colours (e.g., colours on a dark background look lighter than they do on a light background) [45, 130]. To accurately account for simultaneous colour effects in differentiability predictions, the ICD would need additional calibration to determine the user's response to relationships between colours (for example, foreground/background relationships, relative size of the different areas of colour, and different layouts of the colour areas). This is potentially a substantial extension, since there are many potential ways in which one colour can influence another, but recent work on colour appearance models (CAMs) [108] may help.

Interface Colour Selection: The personalized modelling capabilities of ICD can also be applied to the way in which colours are chosen for information presentations and interfaces. One possibility based on the approach of SmartColor [141] is enabling designers specify a set of requirements for a colour, rather than the colour itself. These requirements would be specified in a description language (e.g., the way that XML-based languages such as XUL [32] are used to specify interfaces), and would be related to the function of the colour rather than its visual properties. For example, a colour could have the requirement that it provide a certain degree of differentiability compared to another colour used in the presentation. If a model such as ICD is used, the system could automatically choose colours that will satisfy the requirements, taking into consideration the user and the local environment. It is even possible to model other perceptual interactions with colour and include these in the specification. For example, colour-based 'popout', which works when two colours are sufficiently different from one another, could potentially be defined as a particular amount of separation using the modelling architecture described in this dissertation. A requirement for popout could then be constructed – and when the system encountered this requirement, it could calculate (based on the user's individual model) what colour difference would provide popout, and choose colours accordingly.

8.5.2 Personalized Simulation of CVD

Many people live with CVD, but those with typical colour vision – even friends and family members – have a difficulty time understanding the experience of having CVD. Simulation

tools have been proposed as a solution (see Section 2.3), but existing simulations suffer the same challenges dealt with in this dissertation, namely they do not handle acquired, situationally-induced, and simultaneous CVDs, and they require input that users with CVD usually do not have available. As a result, existing simulations do not provide a precise representation of the experience of CVD to those with typical colour vision.

To address this problem, the ICD presented in this dissertation can be extended to provide personalized simulations of CVD, in which the exact type(s) and severity of CVD an individual has can be captured (via the ICD calibration component) and used to steer the simulated appearance of an image. I have developed and evaluated an early personalized simulation of CVD [34], and am currently working to move this simulation technique to mobile platforms to enable real time simulation of real world situations using the mobile device's camera.

8.6 Extensions to SSM

In addition to the extensions of ICD to other colour uses, the SSM approach proposed in Chapter 3 can be extended to other cases not involving colour.

8.6.1 Super Abilities

Human-Computer Interaction (HCI) seeks to improve every user's interaction with computers by developing new computer input and output technologies. Despite many advances, this discipline has been criticized for focusing only on users with 'typical' abilities, resulting in the exclusion of users with motor, cognitive or perceptual disabilities (e.g., CVD). To help address interaction challenges for users with disabilities, the sub-discipline of accessibility was established and has produced design guidelines, adaptive input and output technologies, and assistive devices.

However, by focusing only on so-called typical users and users with disabilities, the full range of abilities possessed by computer users has been ignored – little work has been done to improve the interaction experience for people with what I call super-abilities – those with perceptual, motor, or cognitive function greater than that found in the typical popu-

lation. Super-abilities are common: they can arise from inherited causes (e.g., higher visual acuity, greater spatial memory capacity), acquired through training (e.g., expert software users, the refined colour 'sense' of graphical designers), or induced through special situations or surroundings (e.g., high-resolution monitors, ultra-precise input devices). Current HCI techniques do not allow people with super-abilities to reach their full potential productivity, diminishing their satisfaction and reducing the value of making the effort to acquire a super-ability and the technological developments that allow super-abilities.

Situation-specific modelling is perfectly situated to help expand the applicability of HCI research to users who experience inherited, acquired, or situationally-induced super-abilities. The calibration component of SSM can be designed in such a way as to gather the full range of abilities of users, and the modelling component can be used to adapt the input or output mechanisms to take full advantage of each user's unique abilities.

Chapter 9

Conclusion and Future Work

Many people have difficulty using colour in digital environments because they have some form of colour vision deficiency. Recolouring tools have been developed to help people with CVD, but existing recolouring tools do not help most people with CVD because they rely on models that cannot represent the colour differentiation abilities of most people with CVD.

In this dissertation, I presented Situation-Specific Modelling (SSM) and applied it to colour differentiation to develop the Individualized model of Colour Differentiation (ICD). SSM is composed of a calibration component that measures a user's abilities, and a modelling component that generalizes the calibration measurements into a full model of the user's abilities. SSM calibration is performed by a user in situ to automatically incorporate almost every factor that influences the user's abilities.

To apply the SSM approach to colour differentiation, I first designed and developed a feasibility implementation using RGB colour space (ICD-1). Through a user study, I found that ICD-1 is sensitive to environmental changes and differences between users, but robust to small changes so that very frequent calibration is not needed. This study also identified that a 32-minute calibration provided the right balance between sampling data and calibration length.

Identifying that the 32-minute calibration of ICD-1 was too long to be practical, I revised the ICD to use a perceptually-uniform colour space (ICD-2). This allowed a significant reducing in the number of calibration measurements. I also revised the internal representation of ICD to more closely approximate actual human colour differentiation than the internal representation used in ICD-1. Through a user study, I found that the new calibration was about 24 times faster (down to about 2 minutes), and the accuracy of the model was slightly but significantly improved.

Recolouring tools utilize constraint optimization to find an optimal or near-optimal replacement colour set. This requires hundreds of thousands or millions of colour differentiability predictions to be made during recolouring. The modifications that resulted in the reduced calibration time for ICD-2 also resulted in the prediction time being much longer than in ICD-1. Prediction time was also not constant, varying proportionally with the severity of the user's CVD. To reduce prediction time, I developed a new prediction mechanism that replaced the time-consuming operations of ICD-2's prediction algorithm with constant time affine transformations (ICD-3). Comparing the prediction time using the data from ICD-2's user study, the new prediction function was found to be 56.4 times faster, on average. This improvement in speed did not come at a cost in accuracy – the accuracy of ICD-2 and ICD-3 were found to be not significantly different.

The `areDifferentiable` function introduced in ICD-1 is not powerful enough to recolour certain visualizations properly (e.g., gradients). To address this, I added a real-valued `howDifferentiable` function, resulting in ICD-4. This function allows the degree of differentiability between two colours to be predicted so that the perceived difference between two colours can be maintained during recolouring.

Having reduced the calibration time, improved the accuracy, reduced the prediction time, and increased the prediction power of ICD, I then applied the ICD to a recolouring tool to build ICDRecolour. ICDRecolour serves as the final evaluation of the ICD, showing that ICD enables significantly improved colour matching accuracy and reduced match selection time across a wide variety of user abilities and situations.

Through modelling the colour differentiation abilities of a particular user in a particular environment with the SSM-based ICD, recolouring tools can be extended to help most people with inherited, acquired, situationally-induced, and simultaneous CVDs.

9.1 Contributions

The central contribution of this dissertation is the identification of the situation-specific modelling (SSM) approach and the demonstration of how it can be applied to colour differentiation to produce the individualized model of colour differentiation (ICD). The ICD

uses in-situ, empirically-based performance tasks to measure a user's ability to differentiate colour, and then utilizes these measurements to construct a prediction model that is specific to the user and his/her situation.

This dissertation also presented a number of secondary contributions:

1. A summary of the wide variety of factors that influence users' abilities to differentiate colours resulting in inherited, acquired, or situationally-induced colour vision deficiency (Section refCVD1).

2. An overview of the different applications of colour differentiation in digital systems (Section 2.4).

3. An evaluation of the feasibility of using situation-specific modelling to develop individualized models of colour differentiation (Section 4.3).

4. A performance-based ICD calibration technique that rapidly collects SSM calibration data (Section 5.2).

5. Improved modelling techniques that reduce the time to perform a prediction and increase the power of a prediction for ICDs (Sections 6.1 and 6.3).

6. An implementation of an ICD-based recolouring tool (Section 7.2).

7. An evaluation of the ICD-based recolouring tool showing ICD extends the applicability of recolouring tools to many types and severities of colour vision deficiency (Section 7.3).

9.2 Future Work

In the future, I will pursue a number of the extensions to the SSM-based ICD discussed in Chapter 8.

Improving ICD: Although there are many possible improvements to the prediction time and internal representation of colour differentiation in the ICD, the most valuable improvement is reducing the frequency of calibration. To achieve this, a community of models needs to be established first. This community will provide a large pool of models for adaptation

systems to draw from. The models will need to be categorized according to the user and environmental characteristics of their calibration, and environmental sensing will help automate this. Finally, I will explore methods for building new models by interpolating between existing models. This will help extend the applicability of each model, further reducing the frequency and length of calibration.

Passive ICD Calibration: In addition to improving ICD by reducing calibration frequency, I also plan to explore calibrating the ICD through calibration games. As identified in Section 8.3.1, recolouring on mobile devices allows the benefits of recolouring to be extended to real world situations. To extend ICDRecolour to mobile recolouring, calibrations can be provided by short and motivating games. This allows calibrations to be gathered while the user is 'killing time' waiting in line, for example. If these game-based calibrations are shared with the community of models, the number and diversity of models will be increased.

Extending ICD: Interface and visualization designers often have difficulty choosing colours that fit their design criteria and are accessible to a wide audience. To facilitate this, I plan to extend existing interface markup languages (e.g., XUL [32]) to include colour-based constraints. These constraints will include colour relationship properties identifying interface items as differentiable from each other (or the background) or popping out from the rest of the interface items. Using the ICD, the colours can be chosen for the interface to meet the designer's goals by incorporating the colour differentiation abilities of the user.

Personalized CVD Simulation: As the ICD calibration component measures the exact colour differentiation abilities of a particular user, it can be used to show people with typical colour vision exactly how the world appears to someone with CVD. However, the personalized simulation of CVD I developed previously [34] only works for still images because it takes too long to compute the simulation. This results in personalized CVD simulation being unavailable for many dynamic, real world situations (e.g., fireworks displays, grocery shopping). By improving the speed to perform a simulation, this technique can be extended to operate on the video stream from mobile device cameras in real time. This will expand the number of situations where someone without CVD can improve his/her understanding of CVD by facilitating inquisitive exploration of the world as it appears to people with CVD.

Extending SSM: Finally, the SSM approach presented in this dissertation can be extended to many other situations. Chapter 3 presented the idea of using SSM to measure and model a particular individual's hearing ability. Hearing ability is another great application area for SSM because people experience inherited, acquired, and situationally-induced hearing loss, so the dynamic modelling nature of the SSM can be extended to modelling a user's auditory capabilities.

Another extension of the SSM was highlighted at the end of Chapter 8, that of modelling user's 'super abilities'. By exactly measuring the full range of user abilities, interfaces can be adapted to not only accommodate disabilities, but also super abilities. Super abilities can be inherited, acquired through training, or situationally-induced through specialized hardware, so they are a natural extension to SSM.

References

[1] Toshiyuki Amano and Hirokazu Kato. Appearance control by projector camera feedback for visually impaired. In *CVPRW 2010: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 57–63. IEEE, 2010.

[2] Joseph J. Atick and A. Norman Redlich. Towards a theory of early visual processing. *Neural Computation*, 2(3):308–320, 1990.

[3] Jibin Bao, Yuanyuan Wang, Yu Ma, and Xiadong Gu. Color image transformation based on an adaptive mapping algorithm. *Chinese Journal of Scientific Instrument*, 28:115–118, 2007.

[4] Jibin Bao, Yuanyuan Wang, Yu Ma, and Xiadong Gu. Re-coloring images for dichromats based on an improved adaptive mapping algorithm. In *ICALIP 2008: IEEE International Conference on Audio, Language and Image Processing*, pages 152–156. IEEE, 2008.

[5] James Bednar and Risto Miikkulainen. Joint maps for orientation, eye, and direction preference in a self-organizing model of v1. *Neurocomputing*, 69:1272–1276, 2006.

[6] Mike Bennett. *Designing For An Individuals Eyes: Human-Computer Interaction, Vision And Individual Differences*. PhD thesis, University College Dublin, Ireland, 2009.

[7] Mike Bennett and Aaron Quigley. A method for the automatic analysis of colour category pixel shifts during dichromatic vision. *Advances in Visual Computing*, pages 457–466, 2006.

[8] Mike Bennett and Aaron Quigley. Creating personalized digital human models of perception for visual analytics. *UMAP 2011: User Modeling, Adaption and Personalization*, pages 25–37, 2011.

[9] Tos T. J. M. Berendschot, Wendy M. R. Broekmans, Ineke A. A. Klöpping-Ketelaars, Alwine F. M. Kardinaal, Geert van Poppel, and Dirk van Norren. Lens aging in relation to nutritional determinants and possible risk factors for age-related cataract. *Archives of Ophthalmology*, 120(12):1732, 2002.

[10] Jennifer Birch. *Diagnosis of Defective Colour Vision*. Butterworth Heinemann, Linacre House, Jordan Hill, Oxford, second edition, 2001.

[11] Jennifer Birch, John L. Barbur, and Alister J. Harlow. New method based on random luminance masking for measuring isochromatic zones using high resolution colour displays. *Ophthalmic & Physiological Optics: the Journal of the British College of Ophthalmic Opticians (Optometrists)*, 12(2):133–136, April 1992.

[12] Cosimo Birtolo, Paolo Pagano, and Luigi Troiano. Evolving colors in user interfaces by interactive genetic algorithm. In *NaBIC 2009: World Congress on Nature & Biologically Inspired Computing*, pages 349–355. IEEE, 2009.

[13] Steven Blake, David L. Black, Mark A. Carlson, Elwyn Davies, Zheng Wang, and Walter Weiss. An architecture for differentiated services. Memorandum rfc 2475, The Internet Society, December 1998.

[14] David A. Booth and Richard P. J. Freeman. Discriminative feature integration by individuals. *Acta Psychologica*, 84(1):1–16, 1993.

[15] Seth E. Bouvier and Stephen A. Engel. Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex*, 16(2):183, 2006.

[16] Hans Brettel, Françoise Viénot, and John D. Mollon. Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America A*, 14(10):2647–2655, October 1997.

[17] Pascual Capilla, María Amparo Díez-Ajenjo, María José Luque, and Jesús Malo. Corresponding-pair procedure: A new approach to simulation of dichromatic color perception. *Journal of the Optical Society of America A*, 21:176–186, February 2004.

[18] Pascual Capilla, María José Luque, and María Amparo Díez-Ajenjo. Looking for the dichromatic version of a colour vision model. *Journal of Optics A: Pure and Applied Optics*, 6:906, 2004.

[19] Weifeng Chen, Wei Chen, and Hujun Bao. An efficient direct volume rendering approach for dichromats. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2144–2152, 2011.

[20] Barry L. Cole. The handicap of abnormal colour vision. *Clinical and Experimental Optometry*, 87(4-5):258–275, July 2004.

[21] Barry L. Cole and Ka Yee Lian. Search for coloured objects in natural surroundings by people with abnormal colour vision. *Clinical and Experimental Optometry*, 89(3):144–149, May 2006.

[22] Barry L. Cole, Jennifer D. Maddocks, and Ken Sharpe. Visual search and the conspicuity of coloured targets for colour vision normal and colour vision deficient observers. *Clinical and Experimental Optometry*, 87(4–5):294–304, July 2004.

[23] Stephen J. Dain. Clinical colour vision tests. *Clinical and Experimental Optometry*, 87(4-5):276–293, July 2004.

[24] H.J.A. Dartnall, James K. Bowmaker, and John D. Mollon. Human visual pigments: microspectrophotometric results from the eyes of seven persons. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1218):115–130, 1983.

[25] Ian R. L. Davies, Glynis Laws, Greville G. Corbett, and David J. Jerrett. Cross-cultural differences in colour vision: Acquired 'colour-blindness' in africa. *Personality and Individual Differences*, 25(6):1153–1162, 1998.

[26] Lawrence Davis. Bit-climbing, representational bias, and test suite design. In *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 18–23. Morgan Kaufmann Publishers, 1991.

[27] Hy Day, Jeffrey Jutai, William Woolrich, and Graham Strong. The stability of impact of assistive devices. *Disability & Rehabilitation*, 23(9):400–404, 2001.

[28] Yinhui Deng, Yuanyuan Wang, Yu Ma, Jibin Bao, and Xiadong Gu. A fixed transformation of color images for dichromats based on similarity matrices. *Advanced Intelligent Computing Theories and Applications with Aspects of Theoretical and Methodological Issues*, pages 1018–1028, 2007.

[29] James Dillon, Lei Zheng, John C. Merriam, and Elizabeth R. Gaillard. Transmission of light to the aging human retina: possible implications for age related macular degeneration. *Experimental Eye Research*, 79(6):753–759, 2004.

[30] Bob Dougherty and Alex Wade. Vischeck daltonize: Color blind image correction. http://www.vischeck.com/daltonize/, March 2005.

[31] Dean Farnsworth. Farnsworth-munsell 100-hue and dichotomous tests for color vision. *Journal of the Optical Society of America (1917-1983)*, 33, October 1943.

[32] Kenneth C. Feldt. *Programming Firefox: Building rich internet applications with XUL*. O'Reilly Media, Incorporated, 2007.

[33] Andrew W. Fitzgibbon, Maurizio Pilu, and Robert B. Fisher. Direct least square fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.

[34] David R. Flatla and Carl Gutwin. "So that's what you see!" building understanding with personalized simulations of colour vision deficiency. In *ASSETS '12: Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 167–174, 2012.

[35] David R. Flatla, Carl Gutwin, Lennart E. Nacke, Scott Bateman, and Regan L. Mandryk. Calibration games: making calibration tasks enjoyable by adding motivating game elements. In *UIST '11: Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 403–412, 2011.

[36] Brand Fortner and Theodore E. Meyer. *Number by Colors: A Guide to Using Color to Understand Technical Data*. Springer-Verlag New York, Inc., New York, NY, USA, 1996.

[37] Brendan J. Frey and Delbert Dueck. Clustering by passing messages between data points. *Science*, 315:972–976, February 2007.

[38] Elizabeth R. Gaillard, Lei Zheng, John C. Merriam, and James Dillon. Age-related changes in the absorption characteristics of the primate lens. *Investigative Ophthalmology & Visual Science*, 41(6):1454, 2000.

[39] Krzysztof Z. Gajos, Jing Jing Long, and Daniel S. Weld. Automatically generating custom user interfaces for users with physical disabilities. In *ASSETS '06: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 243–244. ACM, 2006.

[40] Krzysztof Z. Gajos and Daniel S. Weld. SUPPLE: automatically generating user interfaces. In *IUI '04: Proceedings of the 9th international conference on Intelligent user interfaces*, pages 93–100. ACM, 2004.

[41] Krzysztof Z. Gajos and Daniel S. Weld. Preference elicitation for interface optimization. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 173–182. ACM, 2005.

[42] Krzysztof Z. Gajos, Daniel S. Weld, and Jacob O. Wobbrock. Automatically generating personalized user interfaces with supple. *Artificial Intelligence*, 174(12):910–950, 2010.

[43] Krzysztof Z. Gajos, Jacob O. Wobbrock, and Daniel S. Weld. Automatically generating user interfaces adapted to users' motor and vision capabilities. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 231–240. ACM, 2007.

[44] Krzysztof Z. Gajos, Jacob O. Wobbrock, and Daniel S. Weld. Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces. In *CHI '08: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 1257–1266. ACM, 2008.

[45] E. Bruce Goldstein. *Sensation and Perception*. Wadsworth-Thomas Learning, sixth edition, 2002.

[46] Sabine Graf, Kinshuk, and Tzu-Chien Liu. Identifying learning styles in learning management systems by using indications from students' behaviour. In *ICALT'08. Eighth IEEE International Conference on Advanced Learning Technologies.*, pages 482–486. IEEE, 2008.

[47] Clarence H. Graham and Yun Hsia. Color defect and color theory. *Science*, 127(3300):675, 1958.

[48] Web Accessibility Initiative Group. Web content accessibility guidelines. Website, June 2011.

[49] Radim Halır and Jan Flusser. Numerically stable direct least squares fitting of ellipses. In *The Sixth International Conference in Central Europe on Computer Graphics and Visualization*, pages 59–108, 1998.

[50] Legrand H. Hardy, Gertrude Rand, and M. Catherine Rittler. HRR polychromatic plates. *Journal of the Optical Society of America*, 44(7):509–521, 1954.

[51] Meredith G. Harris and C.R. Cabrera. Effect of tinted contact lenses on color vision. *American Journal of Optometry and Physiological Optics*, 53(3):145, 1976.

[52] John A. Hartigan. Representation of similarity matrices by trees. *Journal of the American Statistical Association*, pages 1140–1158, 1967.

[53] Chris G. Healey. Choosing effective colours for data visualization. In *Proceedings of Visualization '96*, pages 263–270. IEEE, 1996.

[54] Donald Hearn and M. Pauline Baker. *Computer Graphics with OpenGL*. Pearson Prentice Hall, Upper Saddle River, NJ, USA, third edition, 2004.

[55] Paul S. Heckbert. *Color Image Quantization for Frame Buffer Display*, volume 16. ACM, 1980.

[56] M Heim and J Morgner. Disturbed color vision in endogenous psychoses. *Psychiatrische Praxis*, 28(6):284–286, 2001.

[57] Ewald Hering. *Outlines of a Theory of the Light Sense*. Harvard University Press, Cambridge, Massachusetts, USA, 1964. Translated by L. Hurvich and D. Jameson.

[58] Jeffery K. Hovis. Long wavelength pass filters designed for the management of color vision deficiencies. *Optometry & Vision Science*, 74(4):222–230, 1997.

[59] Yinghua Hu. Visual simulating dichromatic vision in cie space. Technical report, University of Central Florida, 2004.

[60] Chun-Rong Huang, Kuo-Chuan Chiu, and Chu-Song Chen. Key color priority based image recoloring for dichromats. In *PCM 2010: 11th Pacific Rim Conference on Multimedia: Part II: Advances in Multimedia Information Processing*, pages 637–647. Springer-Verlag, 2010.

[61] Chun-Rong Huang, Kuo-Chuan Chiu, and Chu-Song Chen. Temporal color consistency based video reproduction for dichromats. *IEEE Transactions on Multimedia*, 13(5):950–960, October 2011.

[62] Jia-Bin Huang, Chu-Song Chen, Tzu-Chen Jen, and Sheng-Jyh Wang. Image recolorization for the colorblind. In *ICASSP 2009: IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1161–1164. IEEE, 2009.

[63] Jia-Bin Huang, Yu-Cheng Tseng, Se-In Wu, and Sheng-Jyh Wang. Information preserving color transformation for protanopia and deuteranopia. *IEEE Signal Processing Letters*, 14(10):711–714, October 2007.

[64] Jia-Bin Huang, Sih-Ying Wu, and Chu-Song Chen. Enhancing color representation for the color vision impaired. Technical Report 1, INRIA, September 2008.

[65] Amy Hurst, Scott. E. Hudson, and Jennifer Mankoff. Dynamic detection of novice vs. skilled use without a task model. In *CHI '07: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 271–280. ACM, 2007.

[66] Amy Hurst, Scott. E. Hudson, Jennifer Mankoff, and Shari Trewin. Automatically detecting pointing performance. In *IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 11–19. ACM, 2008.

[67] Gennaro Iaccarino, Delfina Malandrino, Marco Del Percio, and Vittorio Scarano. Efficient edge-services for colorblind users. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 919–920, New York, NY, USA, May 2006. ACM Press.

[68] Gennaro Iaccarino, Delfina Malandrino, and Vittorio Scarano. Personalizable edge services for web accessibility. In *W4A: Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A)*, pages 23–32, New York, NY, USA, 2006. ACM Press.

[69] Manabu Ichikawa, Kiyoshi Tanaka, Shoji Kondo, Koji Hiroshima, Kazuo Ichikawa, Shoko Tanabe, and Kiichiro Fukami. Web-page color modification for barrier-free color vision with genetic algorithm. In E. Cantú-Paz, J. A. Foster, K. Deb, L. D. Davis, R. Roy, U. M. O'Reilly, H. G. Beyer, R. Standish, G. Kendall, S. Wilson, J. Wegener, D. Dasgupta, M. A. Potter, and A. C. Schultz, editors, *GECCO 2003: Proceedings of the Genetic and Evolutionary Computation Conference*, Lecture Notes in Computer Science 2724, pages 2134–2146, Berlin Heidelberg, 2003. Springer-Verlag.

[70] Manabu Ichikawa, Kiyoshi Tanaka, Shoji Kondo, Koji Hiroshima, Kazuo Ichikawa, Shoko Tanabe, and Kiichiro Fukami. Preliminary study on color modification for still images to realize barrier-free color vision. *2004 IEEE International Conference on Systems, Man and Cybernetics*, 1:36–41, October 2004.

[71] Shinobu Ishihara. Tests for colour-blindness. Pseudoisochromatic Plates, 1917.

[72] Z. Jakab and Klára Wenzel. Detecting tetrachromacy in human subjects. *Perception*, 33:S64, 2004.

[73] Luke Jefferson. *Computer Accessibility for Colour-Blind People*. PhD thesis, University of East Anglia, Norwich, UK, November 2008.

[74] Luke Jefferson and Richard Harvey. Accommodating color blind computer users. In *ASSETS '06: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 40–47, New York, NY, USA, 2006. ACM.

[75] Luke Jefferson and Richard Harvey. An interface to support color blind computer users. In *CHI '07: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 1535–1538, New York, NY, USA, April 2007. ACM.

[76] Jennifer Birch. Extreme Anomalous Trichromatism. In J.D. Mollon and J. Pokorny and K. Knoblauch, editor, *Normal and Defective Colour Vision*, chapter 38, pages 364–369. Oxford University Press, New York, New York, 2003.

[77] Deane B. Judd. Color perceptions of deuteranopic and protanopic observers. *Journal of the Optical Society of America*, 39(3):252–256, 1949.

[78] Shaun K. Kane, Jacob O. Wobbrock, and Ian E. Smith. Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 109–118. ACM, 2008.

[79] Tapas Kanungo, David M. Mount Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):881–892, 2002.

[80] Scott Kirkpatrick, Charles D. Gelatt Jr., and Mario P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671, 1983.

[81] Heidi H. Koester, Edmund F. LoPresti, Glen Ashlock, William W. McMillan, Pam Moore, and Richard C. Simpson. Compass: Software for computer skills assessment. In *CSUN 2003: International Conference on Technology and Persons with Disabilities*, 2003.

[82] Heidi H. Koester, Edmund F. LoPresti, and Richard C. Simpson. Measurement validity for compass assessment software. In *RESNA'06: Proceedings of the RESNA 29th Annual Conference*, 2006.

[83] Heidi H. Koester, Richard C. Simpson, Don Spaeth, and Edmund F. LoPresti. Reliability and validity of compass software for access assessment. In *RESNA'07: Proceedings of the RESNA 30th Annual Conference*, 2007.

[84] Shoji Kondo. A computer simulation of anomalous color vision. *Color Vision Deficiencies*, pages 145–159, 1990.

[85] Gábor Kovács, Itala Kucsera, György Ábrahám, and Klára Wenzel. Enhancing color representation for anomalous trichromats on CRT monitors. *Color Research and Application*, 26(S1):S273–S276, 2001.

[86] Vassili Kovalev. Towards image retrieval for eight percent of color-blind men. In *ICPR 2004: Proceedings of the 17th International Conference on Pattern Recognition*, volume 2, pages 943–946. IEEE, 2004.

[87] Vassili Kovalev and Maria Petrou. Optimising the choice of colours of an image database for dichromats. *Machine Learning and Data Mining in Pattern Recognition*, pages 456–465, 2005.

[88] Vassili Kovalev and Maria Petrou. Mining dichromatic colours from video. *Advances in Data Mining: Applications in Medicine, Web Mining, Marketing, Image and Signal Mining*, pages 431–443, 2006.

[89] Giovane R. Kuhn. Image recoloring for color-vision deficients. Master's thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, April 2008.

[90] Giovane R. Kuhn, Manuel M. Oliveira, and Leandro A. F. Fernandes. An efficient naturalness-preserving image-recoloring method for dichromats. *IEEE Transactions on Visualization and Computer Graphics*, 14:1747–1754, November 2008.

[91] Alessandro Lai, Alessandro Soro, and Riccardo Scateni. Interactive calibration of a multi-projector system in a video-wall multi-touch environment. In *UIST '10: Adjunct proceedings of the 23rd annual ACM symposium on User interface software and technology*, pages 437–438. ACM, 2010.

[92] Bruce Lindbloom. Useful color equations, March 2007. `http://www.brucelindbloom.com` Last Accessed: February 5, 2013

[93] Bo Liu, Meng Wang, Linjun Yang, Xiuqing Wu, and Xian-Sheng Hua. Efficient image and video re-coloring for colorblindness. In *ICME 2009: IEEE International Conference on Multimedia and Expo*, pages 906–909. IEEE, 2009.

[94] Richard B. Lomax, Peter Ridgway, and Maureen Meldrum. Does occupational exposure to organic solvents affect colour discrimination? *Toxicological Reviews*, 23(2):91–121, 2004.

[95] Yu Ma, Xiadong Gu, and Yuanyuan Wang. Color discrimination enhancement for dichromats using self-organizing color transformation. *Information Sciences*, 179(6):830–843, 2009.

[96] David MacAdam. Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32(5):247–274, 1942.

[97] Gustavo M. Machado. A model for simulation of color vision deficiency and a color contrast enhancement technique for dichromats. Master's thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, September 2010.

[98] Gustavo M. Machado and Manuel M. Oliveira. Real-time temporal-coherent color contrast enhancement for dichromats. *Computer Graphics Forum*, 29(3):933–942, 2010.

[99] Gustavo M. Machado, Manuel M. Oliveira, and Leandro A. F. Fernandes. A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1291–1298, 2009.

[100] Curtis E. Martin, John G. Keller, Steven K. Rogers, and Matthew Kabrinsky. Color blindness and a color human visual system model. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 30(4):494–500, 2000.

[101] Geoffrey J. McLachlan and David Peel. *Finite Mixture Models*. Wiley-Interscience, 2000.

[102] Alina Mereuţă, Sébastien Aupetit, and Mohamed Slimane. Improving web accessibility for dichromat users through contrast preservation. In K. Miesenberger, A. Karshmer, P. Penaz, and W. Zagler, editors, *ICCHP 2012: 13th International Conference on Computers Helping People with Special Needs*, volume LNCS 7382, pages 363–370. Springer-Verlag, July 2012.

[103] Gary W. Meyer and Donald P. Greenberg. Color-defective vision and computer graphics displays. *IEEE Computer Graphics and Applications*, 8:28–40, September 1988.

[104] Rika Mochizuki, Tatsuya Nakamura, Jinhui Chao, and Reiner Lenz. Color-weak correction by discrimination threshold matching. In *CGIV 2008: 4th European Conference on Colour in Graphics, Imaging, and Vision*, pages 9–13, 2008.

[105] Rika Mochizuki, Satoshi Oshima, and Jinhui Chao. Fast color-weakness compensation with discrimination threshold matching. *Computational Color Imaging*, pages 176–187, 2011.

[106] Rika Mochizuki, Satoshi Oshima, Reiner Lenz, and Jinhui Chao. Exact compensation of color-weakness with discrimination threshold matching. *Universal Access in Human-Computer Interaction: Applications and Services*, pages 155–164, 2011.

[107] F.M. De Monasterio. Center and surround mechanisms of opponent-color x and y ganglion cells of retina of macaques. *Journal of Neurophysiology*, 41(6):1418–1434, 1978.

[108] 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 2002: 10th Color Imaging Conference*, pages 23–27. Society for Imaging Science and Technology, 2002.

[109] Shigeki Nakauchi and Tatsuya Onouchi. Detection and modification of confusing color combinations for red-green dichromats to achieve a color universal design. *Color Research & Application*, 33(3):203–211, June 2008.

[110] Jeho Nam, Yong Man Ro, Youngsik Huh, and Munchurl Kim. Visual content adaptation according to user perception characteristics. *IEEE Transactions on Multimedia*, 7(3):435–445, June 2005.

[111] Maureen Neitz and Jay Neitz. Numbers and ratios of visual pigment genes for normal red-green color vision. *Science*, 267(5200):1013–1016, February 1995.

[112] Noboru Ohta and Alan Robertson. *Colorimetry: Fundamentals And Applications*. John Wiley & Sons, Wiley-IS&T Series in Imaging Science and Technology, 2005.

[113] Jinsan Park, Jongho Choi, and Dongil Han. Applying enhanced confusion line color transform using color segmentation for mobile applications. In *First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering (CNSI)*, pages 40–44. IEEE, 2011.

[114] Allen B. Poirson and Brian A. Wandell. The ellipsoidal representation of spectral sensitivity. *Vision Research*, 30(4):647–652, 1990.

[115] Thomas Porter and Tom Duff. Compositing digital images. *ACM SIGGRAPH Computer Graphics*, 18(3):253–259, 1984.

[116] Karl Rasche, Robert Geist, and James Westall. Detail preserving reproduction of color images for monochromats and dichromats. *IEEE Computer Graphics and Applications*, 25(3):22–30, 2005.

[117] Karl Rasche, Robert Geist, and James Westall. Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum*, 24(3):423–432, September 2005.

[118] B.C. Regan, J.P. Reffin, and John D. Mollon. Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vision Research*, 34(10):1279–1299, May 1994.

[119] Yong Man Ro and Seungji Yang. Color adaptation for anomalous trichromats. *International Journal of Imaging Systems and Technology*, 14(1):16–20, 2004.

[120] Carlos Eduardo Rodriguez-Pardo and Gaurav Sharma. Dichromatic color perception in a two stage model: Testing for cone replacement and cone loss models. In *IEEE 10th IVMSP Workshop*, pages 12–17. IEEE, 2011.

[121] Jacek Ruminski, Jerzy Wtorek, Joanna Ruminska, Mariusz Kaczmarek, Adam Bujnowski, Tomasz Kocejko, and Artur Polinski. Color transformation methods for dichromats. In *HSI 2010: 3rd Conference on Human System Interactions*, pages 634–641. IEEE, 2010.

[122] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. When computers fade...pervasive computing and situationally-induced impairments and disabilities. In *International Conference on Human Computer Interaction*, volume 2, pages 1298–1302, 2003.

[123] Donald Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*, pages 517–524. ACM, 1968.

[124] Steven K. Shevell. *The Science of Color*. Elsevier Science Ltd, 2003.

[125] Jonathan R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain, 1994.

[126] Jae Chul Shin, Naoki Matsuki, Hirohisa Yaguchi, and Satoshi Shioiri. A color appearance model applicable in mesopic vision. *Optical review*, 11(4):272–278, 2004.

[127] Louise L. Sloan and Lorraine Wollach. A case of unilateral deuteranopia. *Journal of the Optical Society of America*, 38(6):502–509, 1948.

[128] Jaeil Song, Seungji Yang, Cheonsug Kim, Jeho Nam, Jin Woo Hong, and Yong Man Ro. Digital item adaptation for color vision variations. In Bernice E. Rogowitz and Thrasyvoulos N. Pappas, editors, *Human Vision and Electronic Imaging VIII*, volume 5007, pages 96–103. SPIE-IS&T, 2003.

[129] Judy M. Steward and Barry L. Cole. What do color vision defectives day about everyday tasks? *Optometry and Vision Science*, 66(5):288–295, May 1989.

[130] Maureen C. Stone. *A Field Guide to Digital Color*. A. K. Peters, Natick, Massachusetts, USA, 2003.

[131] Shari Trewin and Helen Pain. Dynamic modelling of keyboard skills: Supporting users with motor disabilities. In A. Jameson, C. Paris, and C. Tasso, editors, *Proceedings of the 6th International Conference on User Modeling*, pages 135–146. Springer Verlag, 1997.

[132] Shari Trewin and Helen Pain. Keyboard and mouse errors due to motor disabilities. *International Journal of Human Computer Studies*, 50(2):109–144, February 1999.

[133] Luigi Troiano, Cosimo Birtolo, and Gennaro Cirillo. Interactive genetic algorithm for choosing suitable colors in user interface. Technical report, University of Sannio, 2010.

[134] Luigi Troiano, Cosimo Birtolo, and Maria Miranda. Adapting palettes to color vision deficiencies by genetic algorithm. In *GECCO '08: Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1065–1072, New York, NY, USA, July 2008. ACM.

[135] Edward R. Tufte. *Envisioning Information*. Graphics Press, Cheshire, Connecticut, USA, tenth edition, 1990.

[136] Russell L. De Valois and Gerald H. Jacobs. Primate color vision. *Science*, 162(3853):533–540, 1968.

[137] Françoise Viénot and Hans Brettel. Color display for dichromats. *Proc. SPIE: Color Imaging*, 4300:199–207, 2001.

[138] Françoise Viénot, Hans Brettel, and John D. Mollon. Digital video colourmaps for checking the legibility of displays by dichromats. *Color: Research and Applications*, 24(4):243–252, 1999.

[139] Françoise Viénot, Hans Brettel, L. Ott, A. Ben M'Barek, and John D. Mollon. What do colour-blind people see? *Nature*, 376:127–128, July 13 1995.

[140] Thomas Wachtler, Ulrike Dohrmann, and Rainer Hertel. Modeling color percepts of dichromats. *Vision Research*, 44(24):2843–2855, 2004.

[141] Ken Wakita and Kenta Shimamura. Smartcolor: Disambiguation framework for the colorblind. In *ASSETS '05: Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 158–165, New York, NY, USA, 2005. ACM Press.

[142] Jan Walraven and Johan W. Alferdinck. Color displays for the color blind. In *IST and SID 5th Color Imaging Conference*, pages 17–22, 1997.

[143] Brian A. Wandell. *Foundations of Vision*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA, 1995.

[144] Meng Wang, Bo Liu, and Xian-Sheng Hua. Accessible image search. In *MM 2009: Proceedings of the Seventeenth ACM International Conference on Multimedia*, pages 291–300. ACM, 2009.

[145] Meng Wang, Bo Liu, and Xian-Sheng Hua. Accessible image search for colorblindness. *ACM Transactions on Intelligent Systems and Technology*, 1(1):8, 2010.

[146] Colin Ware. *Information Visualization - Perception for Design*. Morgan Kaufmann Publishers, San Francisco, California, USA, 2000.

[147] Colin Ware and John C. Beatty. Using color dimensions to display data dimensions. *Human Factors*, 30(2):127–142, 1988.

[148] Klara Wenze, Karoly Ladunga, and Krisztian Samu. Measurement of color defective and normal color vision subjects' color and luminance contrast threshold functions on CRT. *Periodica Polytechnica Series Mechanical Engineering*, 45(1):103–108, 2001.

[149] Andrew D. Wilson. Playanywhere: a compact interactive tabletop projection-vision system. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 83–92. ACM, 2005.

[150] Jacob O. Wobbrock, James Fogarty, Shih-Yen Liu, Shunichi Kimuro, and Susumu Harada. The angle mouse: target-agnostic dynamic gain adjustment based on angular deviation. In *CHI '09: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 1401–1410. ACM, 2009.

[151] Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. Ability-based design: Concept, principles and examples. *TACCESS: ACM Transactions on Accessible Computing*, 3(3):9, 2011.

[152] Alexander Wong and William Bishop. Perceptually-adaptive color enhancement of still images for dichromacy deficiencies. In *CCECE 2008: Canadian Conference on Electrical and Computer Engineering*, pages 2027–2032. IEEE, 2008.

[153] Gunter Wyszecki and W.S. Stiles. *Color science: concepts and methods, quantitative data and formulae*. Wiley-VCH, 2000.

[154] Atsushi Yamashita, Rie Miyaki, and Toru Kaneko. Color information presentation for color vision defectives by using a projector camera system. In *Computer Vision–ACCV 2010 Workshops*, pages 92–101. Springer, 2010.

[155] Seungji Yang and Yong Man Ro. Visual contents adaptation for color vision deficiency. In *ICIP 2003: International Conference on Image Processing*, volume 1, pages 453–456, 2003.

[156] Seungji Yang, Yong Man Ro, Jeho Nam, Jinwoo Hong, Sang Yul Choi, and Jin-Hak Lee. Improving visual accessibility for color vision deficiency based on MPEG-21. *Electronics and Telecommunications Research Institute Journal*, 26(3):195–202, June 2004.

[157] Seungji Yang, Yong Man Ro, Edward K. Wong, and Jin-Hak Lee. Color compensation for anomalous trichromats based on error score of fm-100 hue test. In *IEEE-EMBS 2005: 27th Annual International Conference of the Engineering in Medicine and Biology Society*, pages 6571–6574. IEEE, 2005.

[158] Seungji Yang, Yong Man Ro, Edward K. Wong, and Jin-Hak Lee. Quantification and standardized description of color vision deficiency caused by anomalous trichromats–part I: Simulation and measurement. *EURASIP Journal of Image Video Processing*, 2008(487618):1–9, January 2008.

[159] Seungji Yang, Yong Man Ro, Edward K. Wong, and Jin-Hak Lee. Quantification and standardized description of color vision deficiency caused by anomalous trichromats–part II: Modeling and color compensation. *EURASIP Journal on Image and Video Processing*, 2008(246014):1–12, 2008.

[160] J. Terry Yates, Ioannis Diamantopoulos, and Franz-Josef Daumann. Acquired (transient and permanent) colour vision disorders. *Operational Colour Vision in the Modern Aviation Environment, NATO RTO Technical Report*, 16:43–47, 2001.

Appendix A

Colour Space Conversions

In this Appendix, the colour conversions between RGB and L*u*v* colour space are described. As this is not the main contribution of this dissertation, all of these conversions are taken from Bruce Lindbloom's colour conversion website [92] and Maureen Stone's text which describes fundamental colour conversion principles [130].

A.1 Converting RGB to CIE L*u*v*

In Java, RGB colours are typically expressed as 24-bit integer values with 8 bits for each of the red, green, and blue channels. The top 8 bits can optionally be used to encode alpha (transparency), but I will ignore the alpha channel here. The first step when converting from RGB to CIE L*u*v* is to extract the individual red, green, and blue channel values from the single int:

```
/* Unpackage the int rgb value into an array of int rgb values. */
public static int[] unpackageIntRGB(int rgb)
{
    int r1 = (rgb & 0x00FF0000) >> 16;
    int g1 = (rgb & 0x0000FF00) >> 8;
    int b1 = (rgb & 0x000000FF);
    int[] RGB = { r1, g1, b1 };
    return RGB;
}
```

Next, the individual channel values (which range in discrete values from 0 to 255) must be normalized (converted to double-precision real values between 0.0 and 1.0, inclusive):

```
/* Convert the supplied int rgb colour to 0.0-1.0
 * range double rgb colour. */
public static double[] intRGBToDoubleRGB(int r1, int g1, int b1)
{
    // safety: ensure RGB values are within (0-255), inclusive
    r1 = Math.max(0, r1);          // >= 0
    r1 = Math.min(255, r1);        // <= 255
    g1 = Math.max(0, g1);          // >= 0
    g1 = Math.min(255, g1);        // <= 255
    b1 = Math.max(0, b1);          // >= 0
    b1 = Math.min(255, b1);        // <= 255

    // normalize to 0.0-1.0 and return
    double R = (double) r1 / 255.0;
```

```
    double G = (double) g1 / 255.0;
    double B = (double) b1 / 255.0;
    double[] RGB = { R, G, B };
    return RGB;
}
```

Once the RGB colours are converted to the (0.0-1.0) range, they can now be converted to CIE XYZ colour coordinates. CIE XYZ colour coordinates commonly serve as an intermediary between colour spaces. This simplifies colour conversion libraries by only requiring conversions to and from CIE XYZ to allow conversion to and from the new colour space from any other colour space already supported by the library.

The conversion from normalized RGB colour to CIE XYZ is a two-step process. First, the RGB colours must be linearized using an inverse gamma function. Gamma functions are used to align the linear brightness capabilities of displays with the non-linear human perception of brightness (we detect changes in dark lighting much more easily than changes in bright lighting) [130]. As a result, RGB colours are typically modified to follow an exponential curve as the intensity of a channel increases from 0.0 to 1.0. CIE XYZ colour space does not incorporate this, so an inverse gamma function is used to transform exponentially-mapped RGB colours into linearly-mapped RGB colours. In my colour conversion library, I utilize the piecewise gamma function developed for sRGB [92].

The second step of this conversion is to translate the linearized RGB colours to CIE XYZ colour space. This is achieved through matrix multiplication, in which the RGB colours (represented as a 3 row $\times$ 1 column matrix) are multiplied with a 3×3 matrix that contains the measured CIE XYZ coordinates for each of the three RGB primaries for the display device being used:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_R & Y_R & Z_R \\ X_G & Y_G & Z_G \\ Z_B & Y_B & Z_B \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

In this colour conversion library, the XYZ coordinates for the sRGB primaries are used [92, 130]:

$$\begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix}$$

```
/* Convert the supplied sRGB colour into XYZ colour. */
public static double[] RGBtoXYZ(double[] RGB)
{
    // linearize the RGB values (sRGB gamma function)
    double R, G, B;
    if ( RGB[0] > 0.04045 ) {
        R = Math.pow(((RGB[0] + 0.055) / 1.055), 2.4);
    }
    else {
        R = RGB[0] / 12.92;
    }
```

```
    if ( RGB[1] > 0.04045 ) {
       G = Math.pow(((RGB[1] + 0.055) / 1.055), 2.4);
    }
    else {
       G = RGB[1] / 12.92;
    }
    if ( RGB[2] > 0.04045 ) {
       B = Math.pow(((RGB[2] + 0.055) / 1.055), 2.4);
    }
    else {
       B = RGB[2] / 12.92;
    }

    // determine XYZ values (from Bruce Lindbloom and Stone Book)
    double X = (R * 0.4124564) + (G * 0.3575761) + (B * 0.1804375);
    double Y = (R * 0.2126729) + (G * 0.7151522) + (B * 0.0721750);
    double Z = (R * 0.0193339) + (G * 0.1191920) + (B * 0.9503041);
    double[] XYZ = { X, Y, Z };
    return XYZ;
}
```

To convert the XYZ colour coordinates for a colour to CIE L*u*v* colour space, the following function is used [92]. Note how the relative luminance channel (L*) re-incorporates an element of the gamma function removed when the RGB colours were linearized to restore the non-linear nature of human perception of brightness:

```
/* Convert the supplied XYZ colour to Luv colour. */
public static double[] XYZtoLUV (double[] xyz)
{
    // retrieve the input values
    double X = xyz[0];
    double Y = xyz[1];
    double Z = xyz[2];

    // determine some constants
    double ySubR = (Y / Y_SUB_W);
    double denominator = X + (15.0 * Y) + (3.0 * Z);
    double uPrime;
    double vPrime;
    if (denominator == 0.0) {
       uPrime = U_PRIME_W;
       vPrime = V_PRIME_W;
    }
    else {
       uPrime = (4.0 * X) / denominator;
       vPrime = (9.0 * Y) / denominator;
```

```
    }

    // determine the 'L' component of the Luv colour
    double L;
    if (ySubR > EPSILON) {
       L = (116.0 * Math.cbrt(ySubR)) - 16.0;
    }
    else {
       L = KAPPA * ySubR;
    }

    // determine the 'u' and 'v' components of the Luv colour
    double u = 13.0 * L * (uPrime - U_PRIME_W);
    double v = 13.0 * L * (vPrime - V_PRIME_W);
    double[] LUV = { L, u, v };
    return LUV;
}
```

where

```
EPSILON = 216.0/24389.0
KAPPA = 24389.0/27.0
X_SUB_W = 0.9504
Y_SUB_W = 1.0000
Z_SUB_W = 1.0888
U_PRIME_W = (4.0*X_SUB_W)/(X_SUB_W + 15.0*Y_SUB_W + 3.0*Z_SUB_W)
V_PRIME_W = (9.0*Y_SUB_W)/(X_SUB_W + 15.0*Y_SUB_W + 3.0*Z_SUB_W)
```

A.2 Converting CIE L*u*v* to RGB

To convert a colour from CIE L*u*v* to RGB, the inverse of the procedure for converting from RGB to L*u*v* is applied. First, the L*u*v* colour is converted into CIE XYZ coordinates:

```
/* Convert the supplied Luv colour to XYZ colour. */
public static double[] LUVtoXYZ (double[] luv)
{
    // extract the L, u, and v components of the input
    double L = luv[0];
    double u = luv[1];
    double v = luv[2];

    // determine the 'Y' component of the XYZ colour
    double Y;
    if (L > (KAPPA*EPSILON)) {
       Y = Math.pow(((L + 16.0)/116.0), 3.0);
    }
```

```
    else {
       Y = L / KAPPA;
    }

    // temporary values
    double a = (1.0/3.0) *
               (((52.0 * L) / (u + (13.0 * L * U_PRIME_W))) - 1.0);
    double b = (-5.0 * Y);
    double c = (-1.0 / 3.0);
    double d = Y * (((39.0 * L)/(v + (13 * L * V_PRIME_W))) - 5.0);

    // determine the 'X' and 'Z' components of the XYZ colour
    double X = (d - b) / (a - c);
    if (Double.isNaN(X)) { X = 0.0; }
    double Z = (X * a) + b;
    if (Double.isNaN(Z)) { Z = 0.0; }
    double[] XYZ = { X, Y, Z };
    return XYZ;
}
```

The resulting CIE XYZ colour is then converted into normalized (0.0-1.0) RGB coordinates using the inverse of the sRGB primaries matrix [92]:

$$\begin{bmatrix} 3.3205 & -1.5371 & -0.4985 \\ -0.9693 & 1.8760 & 0.04156 \\ 0.0556 & -0.2042 & 1.0572 \end{bmatrix}$$

```
/* Convert the supplied XYZ colour into sRGB colour. */
public static double[] XYZtoRGB(double[] xyz)
{
    // extract the XYZ values
    double X = xyz[0];
    double Y = xyz[1];
    double Z = xyz[2];

    // determine RGB values (from Bruce Lindbloom and Stone Book)
    double R = (X*3.2404542) + (Y*-1.5371385) + (Z*-0.4985314);
    double G = (X*-0.9692660) + (Y*1.8760108) + (Z*0.0415560);
    double B = (X*0.0556434) + (Y*-0.2040259) + (Z*1.0572252);

    // de-linearlze the RGB values (sRGB gamma two-part function)
    if (R > 0.0031308) {
       R = (1.055 * Math.pow(R, 1.0/2.4)) - 0.055;
    }
    else {
       R = 12.92 * R;
```

```
    }
    if (G > 0.0031308) {
       G = (1.055 * Math.pow(G, 1.0/2.4)) - 0.055;
    }
    else {
       G = 12.92 * G;
    }
    if (B > 0.0031308) {
       B = (1.055 * Math.pow(B, 1.0/2.4)) - 0.055;
    }
    else {
       B = 12.92 * B;
    }
    double[] RGB = { R, G, B };
    return RGB;
}
```

Finally, the resulting RGB colours are normalized (0.0-1.0), so must be converted back to discrete 0-255 values and grouped together into a single integer value:

```
/* Convert the supplied double RGB colour
 * to 0-255 range int RGB colour. */
public static int[] doubleRGBToIntRGB(double[] rgb)
{
    // convert values back to integer
    int r1 = (int) Math.round(rgb[0] * 255.0);
    int g1 = (int) Math.round(rgb[1] * 255.0);
    int b1 = (int) Math.round(rgb[2] * 255.0);

    // return integer values
    int[] RGB = { r1, g1, b1 };
    return RGB;
}

/* Package the three rgb values into a single int value.
 * Forces the values to be bounded between 0 and 255. */
public static int packageIntRGB(int r1, int g1, int b1)
{
    r1 = Math.max(0, r1);          // >= 0
    r1 = Math.min(255, r1);        // <= 255
    g1 = Math.max(0, g1);          // >= 0
    g1 = Math.min(255, g1);        // <= 255
    b1 = Math.max(0, b1);          // >= 0
    b1 = Math.min(255, b1);        // <= 255
    return (r1 << 16) | (g1 << 8) | (b1);
}
```

Appendix B

Study Consent Forms

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: Colour Differentiation Study

Investigators: David Flatla, Department of Computer Science (966-2327)
Carl Gutwin, Department of Computer Science (966-4888)

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This study is concerned with observing how individuals differentiate between colours.

The goal of the research is to model individuals ability to differentiate colours.

The session will take 2 hours and 30 minutes, during which you will be asked look at a series of circles of colours presented on a computer screen and indicate (using keys or buttons) whether the colours are different or not.

At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

The data collected from this study will be used in articles for publication in journals, theses, and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (usually within two months). This summary will outline the research and discuss our findings and recommendations. This summary will be available on the HCI lab's website: http://www.hci.usask.ca/

All personal and identifying data will be kept confidential. No textual excerpts, photographs, or video recordings will be used in the dissemination of research results in scholarly journals or at scholarly conferences. Anonymity will be preserved by using pseudonyms in any presentation of textual data in journals or at conferences. The informed consent form and all research data will be kept in a secure location under confidentiality in accordance with University policy for 5 years post publication. If you have any questions about this aspect of the study, please feel free to ask the study supervisor.

You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca
- Office of Research Services, University of Saskatchewan, (306) 966-4053

Participant's signature:_____

Date:_____

Investigator's signature:_____

Date:_____

A copy of this consent form has been given to you to keep for your records and reference. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan.

Figure B.1: Participant consent form for ICD-1 user study.

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: Color Differentiation Study

Investigators: David Flatla, Department of Computer Science (966-2327)
Carl Gutwin, Department of Computer Science (966-4888)

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This study is concerned with observing how individuals differentiate between colors.

The goal of the research is to identify new modeling techniques for color differentiation.

The session will take approximately 60 minutes, during which you will complete two sets of tasks. For each task, you will be presented with a colored semicircle centered on a square background of another color. Your task is to identify, using the numeric keypad, the orientation of the semicircle. More details about each task will be given to you by the study supervisor.

At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

The data collected from this study will be used in articles for publication in journals, theses, and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (usually within two months). This summary will outline the research and discuss our findings and recommendations. This summary will be available on the HCI lab's website: http://www.hci.usask.ca/

All personal and identifying data will be kept confidential. No textual excerpts, photographs, or video recordings will be used in the dissemination of research results in scholarly journals or at scholarly conferences. Anonymity will be preserved by using pseudonyms in any presentation of textual data in journals or at conferences. The informed consent form and all research data will be kept in a secure location under confidentiality in accordance with University policy for 5 years post publication. If you have any questions about this aspect of the study, please feel free to ask the study supervisor.

You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca
- Office of Research Services, University of Saskatchewan, (306) 966-4053

Participant's signature:_____

Date:_____

Investigator's signature:_____

Date:_____

A copy of this consent form has been given to you to keep for your records and reference. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan.

Figure B.2: Participant consent form for ICD-2 user study.

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: Recoloring Tool Study (first portion – HRR Plates)

Investigators: David Flatla, Department of Computer Science (966-2327)
Carl Gutwin, Department of Computer Science (966-4888)

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This is the first part of a two-part study examining how different recoloring tools assist a variety of users in multiple situations.

The goal of the research is to identify superior recoloring techniques for helping people with any form of color vision deficiency.

This portion of the study will take less than 5 minutes, and assesses whether or not you have color vision deficiency (CVD – commonly called color blindness) using a common color vision test called the HRR Pseudoisochromatic Plates. If you are diagnosed with CVD, the type and severity of your CVD will also be assessed using the same test. More details about this test will be given to you by the study supervisor.

At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

The data collected from this study will be used in articles for publication in journals, theses, and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (usually within two months). This summary will outline the research and discuss our findings and recommendations. This summary will be available on the HCI lab's website: http://www.hci.usask.ca/

All personal and identifying data will be kept confidential. No textual excerpts, photographs, or video recordings will be used in the dissemination of research results in scholarly journals or at scholarly conferences. Anonymity will be preserved by using pseudonyms in any presentation of textual data in journals or at conferences. The informed consent form and all research data will be kept in a secure location under confidentiality in accordance with University policy for 5 years post publication. If you have any questions about this aspect of the study, please feel free to ask the study supervisor.

You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca
- Office of Research Services, University of Saskatchewan, (306) 966-4053

Participant's signature:_____

Date:_____

Investigator's signature:_____

Date:_____

A copy of this consent form has been given to you to keep for your records and reference. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan.

Figure B.3: Participant consent form for HRR-plate-test portion of ICDRecolour user study.

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: Recoloring Tool Study (second portion)

Investigators: David Flatla, Department of Computer Science (966-2327)
Carl Gutwin, Department of Computer Science (966-4888)

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This is the second part of a two-part study examining how different recoloring tools assist a variety of users in multiple situations.

The goal of the research is to identify superior recoloring techniques for helping people with any form of color vision deficiency.

This portion of the study will take less than 90 minutes, during which you will complete four repetitions of a two-stage process. In the first stage, your task will be to use the numeric keypad to identify the orientation of a colored semicircle on a shimmering background. In the second step, your task will be to use the mouse pointer to identify which square in a grid of colored squares most closely matches a reference colored square. Each repetition will be performed using a different system in this room. More details about each stage will be given to you by the study supervisor.

At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research.

The data collected from this study will be used in articles for publication in journals, theses, and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (usually within two months). This summary will outline the research and discuss our findings and recommendations. This summary will be available on the HCI lab's website: http://www.hci.usask.ca/

All personal and identifying data will be kept confidential. No textual excerpts, photographs, or video recordings will be used in the dissemination of research results in scholarly journals or at scholarly conferences. Anonymity will be preserved by using pseudonyms in any presentation of textual data in journals or at conferences. The informed consent form and all research data will be kept in a secure location under confidentiality in accordance with University policy for 5 years post publication. If you have any questions about this aspect of the study, please feel free to ask the study supervisor.

You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-4888, gutwin@cs.usask.ca
- Office of Research Services, University of Saskatchewan, (306) 966-4053

Participant's signature:_____

Date:_____

Investigator's signature:_____

Date:_____

A copy of this consent form has been given to you to keep for your records and reference. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan.

Figure B.4: Participant consent form for colour-matching portion of ICDRecolour user study.

Appendix C

Participant Demographics Questionnaires

Demographics survey - Participant #_____ - Colour Differentiation Study

1. Age: _____

2. Sex: **M** **F** (circle one)

3. Handedness: **Left** **Right** **Ambidextrous** (circle one)

3. Occupation: _____

4. If you are a student, list your level, university, major, and year:

_____

5. How many hours a week, on average, do you spend working with computers?

_____

6. List 5 applications that you use frequently (e.g., Firefox, Internet Explorer, Outlook):

i. _____ ii. _____ iii. _____ iv. _____ v. _____

7. How many hours a week, on average, do you spend playing electronic games (e.g., Wii, XBox, PlayStation, computer games, World of Warcraft, Enemy Territory)?

_____

8. Please list a few of the game(s) you play most frequently:

_____

9. Do you have any experience with information visualizations? **Y** **N** (circle one)

10. If yes, please list some visualizations and visualization tools you have used:

_____

11. Please rank your (corrected) visual acuity:

excellent **good** **fair** **poor**

12. Have you been previously diagnosed with colour blindness? **Y** **N** (circle one)

13. To the best of your knowledge, do any of the following relatives have colour blindness?

mother's father: **Y** **N** (circle one)

mother's brother(s): **Y** **N** (circle one)

your brother(s): **Y** **N** (circle one)

Figure C.1: Demographics questionnaire for ICD-1 user study.

Demographic Questionnaire

Please fill out this demographic survey.
* Required

ParticipantID? *
Please enter your participant ID (should be automatic).

Sex? *
Please select your gender from below.

- male
- female
- other

Age? *
Please enter your age.

Occupation? *
Please enter your occupation.

Which is your dominant hand? *
Please select your dominant hand from below.

- right
- left
- abidextrous

On average, roughly how many hours do you use a computer during a typical weekday? *

- I rarely or never use a computer.
- less than 1 hour
- 1-3 hours
- 3-6 hours
- more than 6 hours

Please rank your (corrected) visual acuity *
How well do you see when wearing glasses/contacts?

- excellent
- good
- fair
- poor

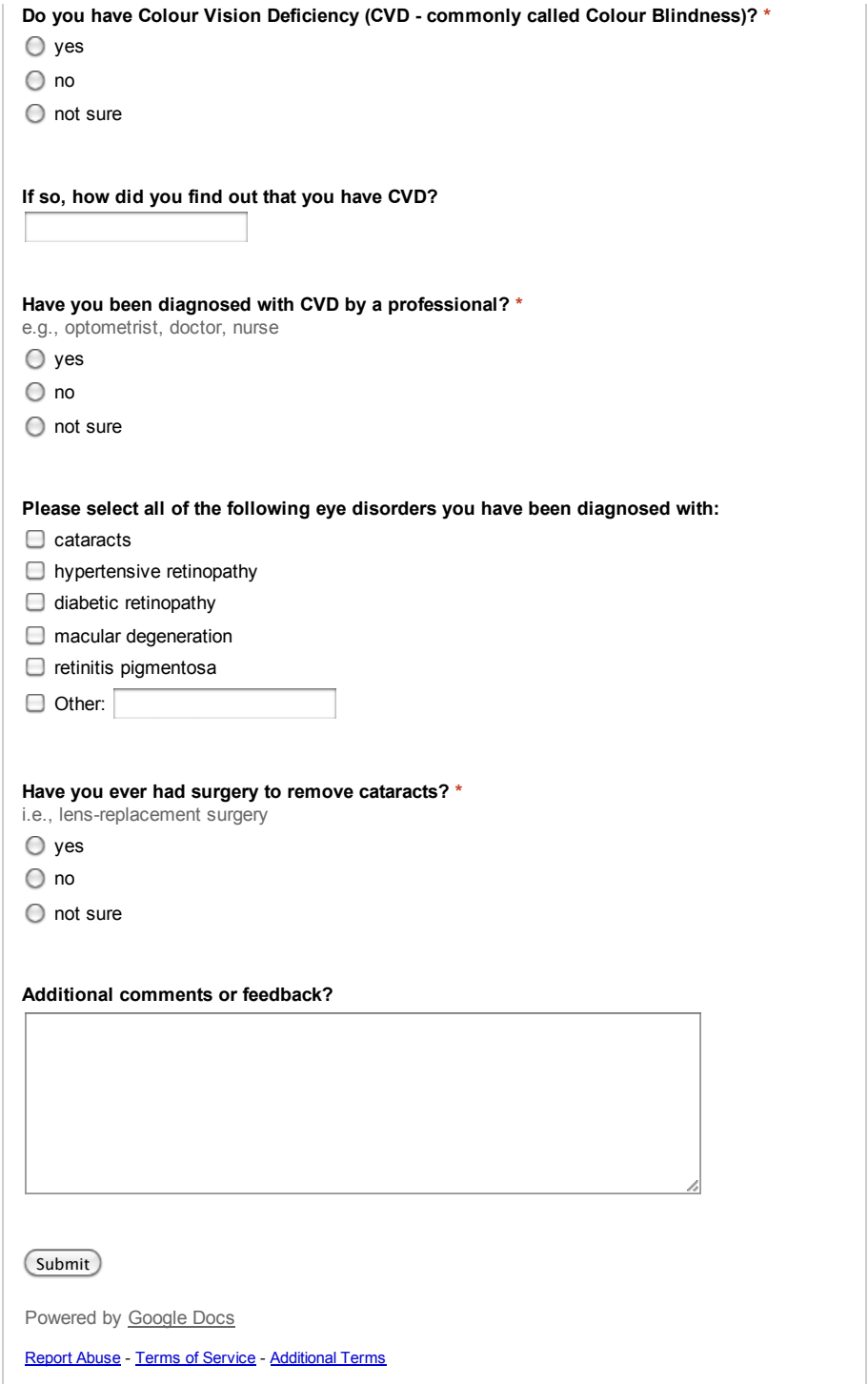
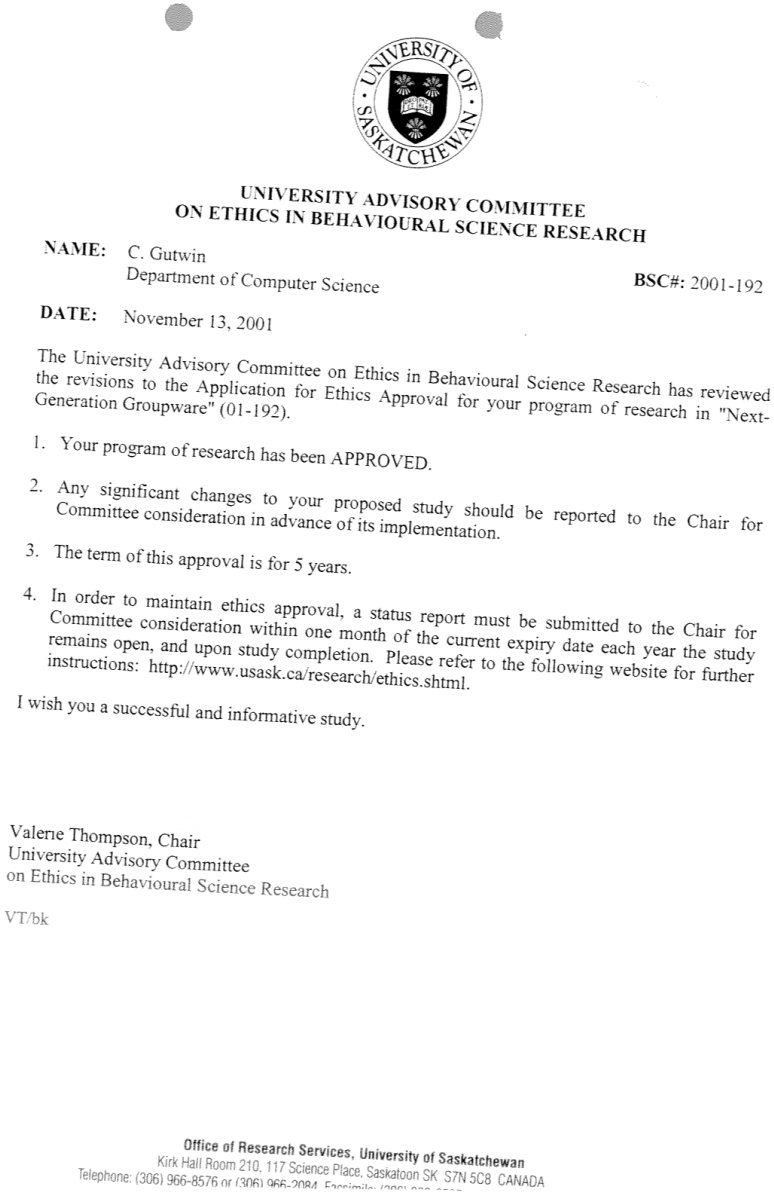


Figure C.2: Online demographics questionnaire for ICD-2 and ICDRecolour user studies.

Appendix D

Research Ethics Certificate



UNIVERSITY ADVISORY COMMITTEE
ON ETHICS IN BEHAVIOURAL SCIENCE RESEARCH

NAME: C. Gutwin
Department of Computer Science **BSC#:** 2001-192

DATE: November 13, 2001

The University Advisory Committee on Ethics in Behavioural Science Research has reviewed the revisions to the Application for Ethics Approval for your program of research in "Next-Generation Groupware" (01-192).

1. Your program of research has been APPROVED.
2. Any significant changes to your proposed study should be reported to the Chair for Committee consideration in advance of its implementation.
3. The term of this approval is for 5 years.
4. In order to maintain ethics approval, a status report must be submitted to the Chair for Committee consideration within one month of the current expiry date each year the study remains open, and upon study completion. Please refer to the following website for further instructions: http://www.usask.ca/research/ethics.shtml.

I wish you a successful and informative study.

Valerie Thompson, Chair
University Advisory Committee
on Ethics in Behavioural Science Research

VT/bk

Office of Research Services, University of Saskatchewan
Kirk Hall Room 210, 117 Science Place, Saskatoon SK S7N 5C8 CANADA
Telephone: (306) 966-8576 or (306) 966-2084 Facsimile: (306) 966-2069

Figure D.1: All studies in this dissertation were conducted under Dr. Carl Gutwin's ethics approval certificate.

Appendix E

Study Instructions for Participants

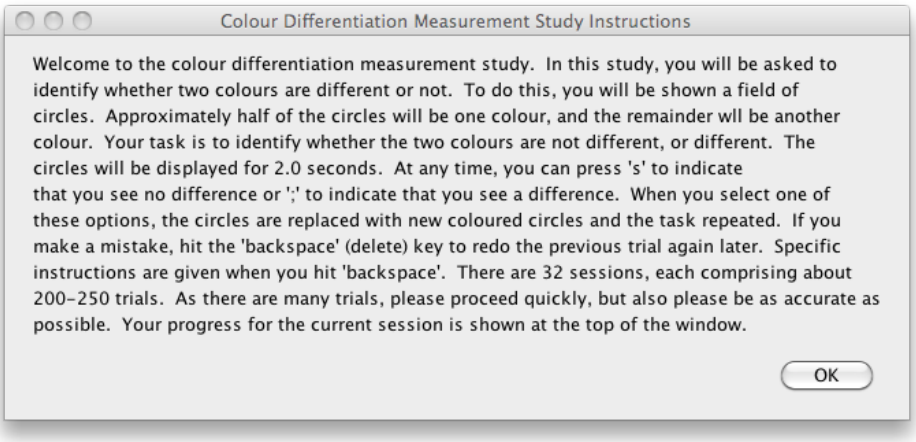


Figure E.1: Instructions given to participants for the study used to evaluate ICD-1.

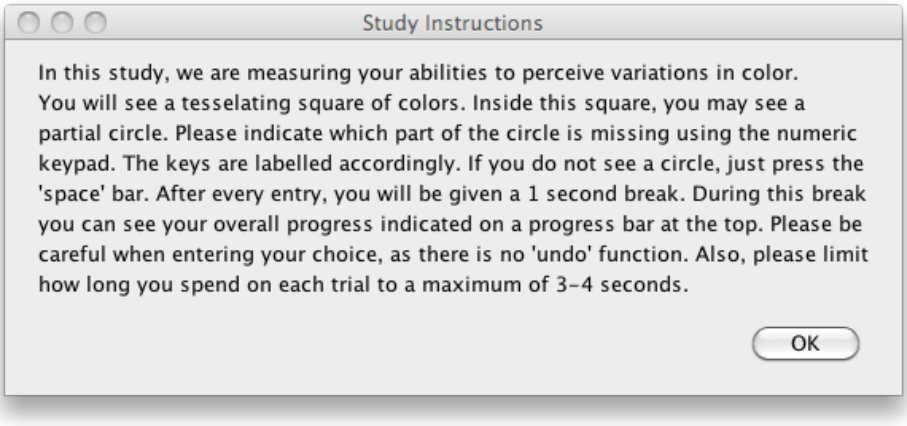


Figure E.2: Instructions given to participants for the study used to evaluate ICD-2.

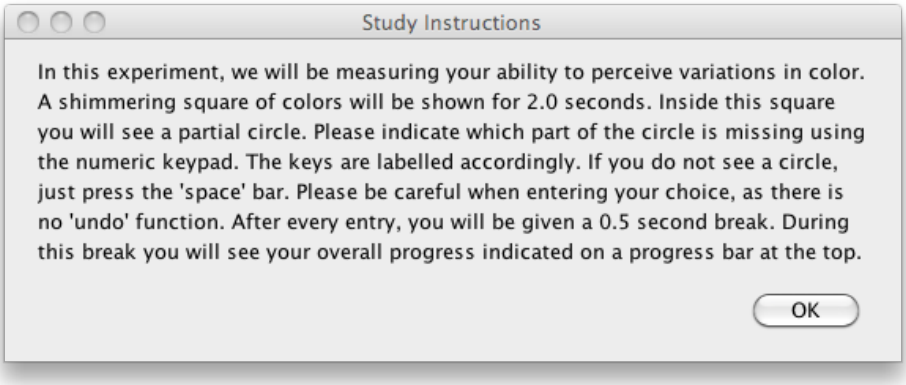


Figure E.3: Instructions given to participants for the calibration component of the study used to evaluate ICDRecolour.

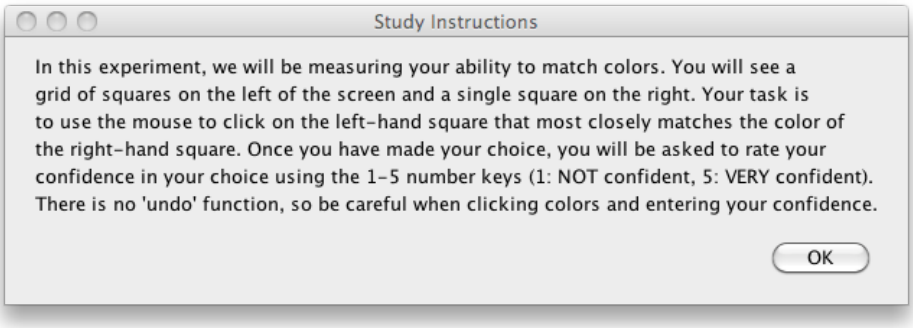


Figure E.4: Instructions given to participants for the matching component of the study used to evaluate ICDRecolour. Reported confidence values were not used in the evaluation of ICDRecolour.