UNIVERSITY of OTAGO
SAPERE · AUDE
Te Whare Wānanga o Otāgo
NEW ZEALAND

PhD Dissertation

Computational Glasses: Repurposing Augmented Reality Glasses for Vision Assistance

---

Jonathan Sutton
Dunedin, New Zealand, Feb 2022

Supervised By
Tobias Langlotz
Holger Regenbrecht
Yuta Itoh
Alexander Plopski

Abstract

A host of visual impairments affect the lives of those around us. These range from refractive errors such as short-sightedness (myopia), far-sightedness (hyperopia) and presbyopia, to colour perception variances such as *colour vision deficiency (CVD)* and cornea yellowing, to age-related issues that are expected to affect us all to some degree such as macular degeneration or glaucoma, and many more. In fact, visual impairments already affect the lives of some 2.2 billion people worldwide according to the *World Health Organization (WHO)* with at least 1 billion having issues that remain unaddressed or untreatable.

Traditional solutions to aid visual impairments involve optical lenses, generally in the form of glasses. Commonplace and widely accepted in society, glasses are used to treat low order refractive errors but are limited to those applications utilising static optical lenses with set refractive properties and limited within-optic variability.

Another means to modify perceived reality is *augmented reality (AR)* which has traditionally been used to introduce virtual objects to a user's reality, or provide virtual information such as guidance, instructions, or spatially located information.

Our research looks to integrate AR with the *human visual system (HVS)* to provide new aids for visual impairments that utilise AR techniques to controllably modulate the environment, aiding the varied and precise needs of the visually impaired. In particular, *optical see-through head-mounted displays (OSTHMDs)* provide the potential to do so in a form factor similar to that of socially accepted traditional glasses. We term this concept of utilising computer-controlled optics and displays to provide visual assistance in a manner akin to glasses: *Computational Glasses*.

This concept has application beyond aiding impairments to providing general assistance for the HVS. Similar to sunglasses being used to assist with high luminance and glare, being able to modulate the perception of the world enables a multitude of visual aids to be produced. The utilisation of Computational Glasses would allow for precise and targeted assistance that can be adjusted as needed.

In this thesis, we present our work demonstrating the feasibility of Computational Glasses created using OSTHMD, exploring their use, and developing some of the foundational steps to their creation. After covering the current state of OSTHMD and the developments still needed to realise them as complete Computational Glasses, we created a series of prototypes to demonstrate their application to aid colour discrimination for the colourblind through several user studies. We then demonstrate augmentation of the unimpaired HVS to place focus on areas of interest with a further series of prototypes and user studies.

Acknowledgements

Many people have had an influence on my research and the production of this thesis, either directly assisting with the research, providing feedback, or providing support. I would like to take this opportunity to thank them.

Firstly, I need to thank my various supervisors, Tobias Langlotz, Holger Regenbrecht, Yuta Itoh, and Alexander Plopski. Be it providing guidance and direction, having discussions, assistance in designing and running studies, writing of papers, or keeping me in scope and on track. It has all been appreciated. So has the patience.

I want to make particular mention of my primary supervisor Tobias Langlotz. I want to thank you for introducing me to research alongside Holger and the time taken, and patience shown, in assisting me in writing this thesis. I also want to acknowledge the many years of mentoring and guidance you have provided, the opportunities you made available to attend conferences, visit other labs, and tutor, and always having time for meetings and to offer advice. Thank you.

Next, I would like to thank the many members of the HCI lab, past and present, who have put up with me over the last years. Be it the hours of distraction, the catch ups over beers, the chats, or trips for coffee, you have made life in the lab enjoyable.

I would also like to thank Gail, Heather, and Karen who kept me out of strife, and organised even when I went out of my way to be the contrary. The many trips for the fresh coffee were a crucial part of my research.

Finally, I want to thank my family. You are always there for me providing support, a sympathetic ear, and offering help. I want to thank my father for the many hours of proof reading and providing comments on my writing.

To my parents, thank you, I could always find rest at home.

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be at the front (after the thesis abstract) of the thesis submitted for examination and library deposit.

Expectations of theses including publications:

The thesis must be an integrated and coherent body of work. The body of the thesis may include entire publications verbatim, but these need to be reformatted to a consistent chapter style, and commentary may need to be added to link the publications together. If the thesis contains multiple publications, candidates need to ensure that they have: included an introductory chapter; demonstrated their ability to critically engage with the literature; carefully considered their research design and justified choices of methodology and methods where necessary; and synthesised and discussed the findings from the publications. It is the candidate's responsibility to ensure that any published work (or parts thereof) included in the thesis comply with the copyright provisions of the publisher and that any guidelines with regard to self-citation are followed.

Details of publications included in and/or appended to this thesis (please add rows as needed)**.**

Chapter/ Append.	**Paper title**	**Authors**	**Contribution of candidate and co-authors – please detail the nature and extent (%)**	**Journal**	**Status (e.g. under review, forthcoming, published)**
2,3	Towards Indistinguishable Augmented Reality: A Survey on Optical See-Through Head-Mounted Displays	Yuta Itoh, Tobias Langlotz, **Jonathan Sutton,** Alexander Plopski	The candidate was involved in the writing of the paper, contributing two sections **(20%).**	*CSUR*	Published
3	Computational Glasses: Vision augmentations using computational near-eye optics and displays	**Jonathan Sutton**, Tobias Langlotz, Yuta Itoh	The candidate was the main author on the paper and work the initial draft, this was then iterated on with co-authors **(75%).**	*ISMAR-Adjunct*	Published

3,4	Seeing Colours: Addressing Colour Vision Deficiency with Vision Augmentations using Computational Glasses	**Jonathan Sutton,** Tobias Langlotz, Alexander Plopski	The candidate was the main author on the paper and wrote the initial draft, this was then iterated on with co-authors **(60%).** The candidate developed the design of the system and studies and running of the studies **(80%)**	*ToCHI*	Published
4	ChromaGlasses: Computational Glasses for Compensating Colour Blindness	Tobias Langlotz, **Jonathan Sutton,** Stefanie Zollmann, Yuta Itoh, Holger Regenbrecht	The candidate was not involved in the writing of this paper **(10%).** The candidate developed the design of the system and studies and running of the studies **(70%).**	*CHI*	Published
5	Look over there! Investigating Saliency Modulation for Visual Guidance with Augmented Reality Glasses	**Jonathan Sutton,** Tobias Langlotz, Alexander Plopski, Stefanie Zollmann, Yuta Itoh, Holger Regenbrecht	The candidate was the main author on the paper and wrote the initial draft, this was then iterated on with co-authors **(60%).** The candidate developed the design of the system and studies and running of the studies **(70%).**	-	Awaiting Submission
5	A Visual Guidance Design Space for On-Screen AR	**Jonathan Sutton,** Tobias Langlotz, Alexander Plopski	The candidate has written the first draft of the paper to date under advice from co-authors **(90%)**. The candidate developed the design of the system and studies and running of the studies **(75%).**	-	Being written

Certification by Primary Supervisor:

The undersigned certifies that the above table correctly reflects the nature and extent of the candidate's contribution to this co-authored work

Name:	Signature:	Date:

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Glasses have become commonplace in modern society. Traditionally they have been used to assist the *human visual system (HVS)* aiding impairments and, more recently, have also been used to enhance unimpaired vision (or augmenting it). Originally dating back to the 13th century, optical lenses have been used to correct for low order refractive errors and since then have become socially accepted, even being considered fashionable in some parts of modern society. Traditionally made with thick glass lenses, the composition, size, and precision of glasses have developed over time, leading to slim lenses with various dioptres. Whilst glasses have been developed to include different dioptres in different parts of the one lens, allowing for bifocals, transitional lenses and similar multi-application glasses, they are still static in nature and limited to low order refractive errors. This prohibits them from assisting other issues that are more complex in nature and requiring them to constantly modulate a set portion of the user's *Field of View (FoV)*. This prohibits adjusting the view only as, how, and where required. More recently glasses have been extended to include filters. To aid with visual impairments they have been used to adjust light saturation for *colour vision deficiency (CVD)*, although to limited effect (Gómez-Robledo et al., 2018), and to assist with photophobia (Clark et al., 2017).

Filtered lenses have also been used to assist the unimpaired HVS, augmenting it, with sunglasses being one of the most common examples of this. Sunglasses allow for greater reduction in light entering the eye than that provided by dilating irises. They enhance the ability to see in bright conditions and under glare, and find application not only to general everyday use but also to provide advantages in situations such as sports and driving. Another form of filter lenses that have seen increased popularity in recent years are blue light lenses used to reduce eye strain and the impact of bright *liquid-crystal display (LCD)* screens on sleep patterns and attention (Green et al., 2017). Further lenses enable protection from harm, either physical (safety glasses), or by radiation such as *ultra-violet (UV)* which can be filtered by protective coatings applied to standard glasses. Glasses even have application in entertainment, allowing the HVS to perceive 2D images as though 3D via filter lenses or active shutters. However, these applications retain the issues inherent in traditional glasses in being generally static in nature, having a set use case and are applied liberally without concern for need.

Figure 1.1: Examples of our solution for extended glasses, Computational Glasses. Left: example of them being used to augment the saliency of a scene. Center: an example prototype of our solution. Right: example of our solution being used to aid colour vision deficiency.

In this thesis we present first steps in overcoming these issues for glasses when assisting the HVS. We look to extend glasses beyond their current applications in aiding limited visual impairments and provide a new, more diverse means of augmenting the HVS. Figure 1.1 demonstrates the output from this thesis.

When considering the HVS, over the past centuries humans have become increasingly aware of its intricacies, strengths, weaknesses, and the myriad of conditions and impairments that can affect it. The traditional use of glasses is to provide aids and when looking at the conditions that affect the HVS, conditions such as low order refractive errors are some of the most well documented with short-sightedness (myopia), far-sightedness (hyperopia) and presbyopia being commonly understood. These are also the issues traditionally aided by glasses. However, less well understood issues by the general public, such as CVD, are also prevalent in our society too and remained unaided. According to the *World Health Organization (WHO)*, globally over two billion people are affected by vision impairments (Cieza et al., 2019). With ageing populations and increasing rates of diabetes (Roglic, 2016), the number of people affected by issues associated with age and diabetes, such as glaucoma or diabetic retinopathy, is expected to rise, adding to the 1 billion people already estimated to have unaddressed vision issues. These visual impairments have various impacts on the lives of those who have them. For conditions like CVD these can often be mundane inconveniences with communication and understanding, although they can also have more significant impacts such as reduced job prospects (Steward and Cole, 1989a). Whereas age-related issues can cause a number of problems such as difficulty with literacy, poor facial recognition, and reduced spatial perception with difficulties judging depth, low contrast sensitivity, and reduced peripheral or central FoV. These problems degrade quality of life and are correlated to increased risk of accidents and mortality. For example, poor spatial perception reduces the ability to identify and accurately judge variations in surfaces such as holes and steps, as well as impacting hazard identification (Wood et al., 2011; Zheng et al., 2012).

Given the impact of these unaddressed issues, there is a need to explore new forms of assistance that are capable of tackling a wider range of issues and overcoming the limitations of traditional optics.

Figure 1.2: Simulated examples of assisting vision. Left: loss of visual acuity can be corrected with glasses. Right: bright light can be reduced using sunglasses.

1.1 Motivation

Early glasses were limited in availability and application, however their use has grown to become commonplace and readily available. Today most people can find use for them, if not for vision aid then for general assistance. Whilst glasses were for some time stigmatised in society, they are now generally socially accepted. Figure 1.2 shows two common examples of their application. They are being worn by millions of people world-wide with over 16 million prescription and sunglasses being sold online in the US in 2019 alone[1]. They are even considered fashion accessories and are being worn for no other purpose than this (Pullin, 2009).

As an alternative to traditional glasses, a form of eyewear for modifying vision have been being developed since 1968 (Sutherland, 1968). *Head-mounted displays (HMDs)* have been used to alter visual perception, adjusting reality as *augmented reality (AR)* and replacing it with a virtual one as in *virtual reality (VR)*. Of particular interest for this thesis are AR glasses and *optical see-through head-mounted displays (OSTHMDs)*. AR research and development over the last five decades has enabled the introduction of virtual artefacts into the real environment. This has seen applications in providing guidance, information, training, and gaming (Mulloni et al., 2011; Tatzgern et al., 2016; Jeffri and Rambli, 2020; Poupyrev et al., 2000). AR is achieved in many ways with *video see-through head-mounted displays (VSTHMDs)* intercepting the view of the world with cameras and re-presenting the view from the cameras modified with virtual content, and OSTHMDs where virtual content is added to the view of the world by transparent displays. Although not eyewear, and often not even head worn, *spatial augmented reality (SAR)* also looks to modify vision by directly modifying the light perceived from the environment, being achieved in ways such as using projectors to illuminate the environment. Whilst all styles of AR have their advantages and disadvantages, OSTHMDs have seen much public interest and allow for transportable, personal AR. Early implementations of OSTHMDs were cumbersome and tethered, however from inception they have shared some optical properties with glasses in having stereo see-through lenses, and throughout their development have trended towards a form factor similar to that of glasses. Furthermore, OSTHMDs enable the user to perceive AR content whilst retaining their direct view of the real world, which has various benefits such as re-

[1] www.statista.com/statistics/256287/amount-of-eyewear-sold-in-the-united-states-online-by-type/

Figure 1.3: The development of glasses. Top: Glasses ranging from the 16th century to modern glasses, all demonstrating a similar style. Bottom: Examples of alternative forms of glasses now used to augment the HVS.

Glasses adapted from:

https://commons.wikimedia.org/wiki/File:Post-medieval_spectacles_Silver_spectacles_frame_(FindID_468635).jpg,

https://commons.wikimedia.org/wiki/File:Spectacles,_pair_(51371483677).jpg,

https://commons.wikimedia.org/wiki/File:Lunettes_de_Gandhi-National_Gandhi_Museum_(2).jpg,

https://commons.wikimedia.org/wiki/File:Rounded_glasses.jpg,

https://commons.wikimedia.org/wiki/File:Safety_Eyewear.jpg,

https://commons.wikimedia.org/wiki/File:Anachrome_Aviator%2B_3D_glasses.jpg

tention of vision, computational costs, more readily allowing mutual eye-contact, and not requiring power to allow the user to view the world. Figure 1.4 shows various modern OSTHMDs and their various form factors. In a similar manner to the fashion uses of traditional glasses, AR glasses are also being marketed as fashionable accessories with devices such as the Snap Spectacles[2] or NReal[3].

Although, in order to enable mobility of computing power and to overcome limitations on the current miniaturisation of the underlying technology, some recent devices have adopted a style more akin to safety glasses of face shields[4]. This breaks from the current trend of development in form factor to enable the practical development of applications to apply once the desired form factor can be reached.

Whilst traditional glasses and OSTHMDs show many similarities in their physical usage, their continued separation in practical application provides the motivation for the research outlined in this thesis.

1.2 Research Gap

This thesis presents our research into a new form of vision assistance that we term Computational Glasses. This concept draws from traditional glasses and AR, introducing new ideas to overcome the respective shortcomings of both and providing a modern means to more holistically assist the HVS. We investigated how this can be achieved using commercial OSTHMDs being the most mature form of technology that enables glasses-style modulation of the real world.

[2]https://www.spectacles.com/
[3]https://www.nreal.ai/
[4]https://www.microsoft.com/en-us/hololens

Figure 1.4: Examples of various modern OSTHMDs with different form factors and styles, including those similar to traditional glasses. From left the OSTHMDs are: Lumus DK-52, Moverio Bt-300, Daqri AR Glasses, Hololens 1, Vuzix Blade, Meta 2, and QD Retissa.

Traditionally glasses have been used to correct refractive errors helping millions of people, and whilst they have been extended to assist the HVS in other ways, such as reducing light with sunglasses, their use is still severely limited in capability. Ways to extend the capabilities of glasses have been looked at, such as adding prisms to adjust FoV, (Peli and Jung, 2017; Jung and Peli, 2018). However, these solutions still rely on static optics which limit their application and prevent them from being adjusted situationally.

AR provides us with an alternative means to adjust the HVS and provide it support. How we adjust it can be modified as desired, within the constraints of the AR device being used. Research has investigated this by using VSTHMD to provide zoom, contrast, and edge enhancement (Zhao et al., 2015), and to assist peripheral vision loss (Younis et al., 2017). However, the use of VSTHMD has drawbacks in applying further limitations to the user's view of the world based on the cameras and displays used. For example the front-running Varjo XR-3, whilst being able to match human resolution in the central vision, falls far short of entirely covering even the static field of view achieving 60% coverage, and <20% of the resolveable resolution based on the figures detailed in Chapter 2 (2.1.2)[5]. Entirely limiting the user's view to that of the cameras and display also requires them to operate constantly. Utilising OSTHMDs alleviates these drawbacks to a large degree and as such, have also been utilised in research with works looking to provide depth information to those with low vision (Min Htike et al., 2021) and to enhance field of view (Luo and Peli, 2006; Orlosky, 2014). However their applications suffer from poor alignment and occlusion. Overall, these techniques do not directly modify the user's vision by adjusting the world as they see it, or are not able to modify the environment in a personal and ubiquitous manner. An ideal vision aid would be one that is similar to traditional glasses in that they directly modify the wearer's view of the environment, can be worn ubiquitously, and are socially acceptable, however are also able to be adjusted as needed and provide aids beyond that of traditional glasses.

Whilst AR and OSTHMD provide a starting point, how this can be achieved with such devices remains to be properly addressed. This thesis provides some

[5]https://varjo.com/products/xr-3/

foundational research towards achieving this end. Namely, we present our research demonstrating the use of prototypical customised OSTHMD to modulate a wearers view of the real world with pixel precision, enabling new forms of assistance. We demonstrate examples of how this can be achieved to aid visual impairments and to augment users' capabilities.

In order to address this research gap, we proposed three hypotheses which we looked to test throughout the course of our research and cover in this thesis:

- H1: Utilising AR OSTHMDs we will be able to create effective Computational Glasses that can modulate the user's view with pixel precision, providing visual assistance.
- H2: Utilising these Computational Glasses we will be able to adjust the visual spectra seen to provide aid for visual impairments. In particular, provide effective compensations for CVD.
- H3: Computational Glasses will be able to not only assist impaired vision by providing visual aids, they will also be able to assist the unimpaired vision. In particular, they will be able to effectively guide user's attention by modulating the saliency of a scene.

1.3 Outline and Key Contributions

The following details the structure of the thesis and its key contributions. We also provide a brief introduction to the papers produced during the completion of this thesis.

1.3.1 Outline

To aid understanding of this thesis, we provide an introduction into the background of the HVS, the background of AR focusing on OSTHMD, and how wearable assistive technologies have been used to address visual impairments and to augment the HVS in Chapter 2. This chapter draws on the related work of all our papers, particularly drawing from (Sutton et al., 2019) and (Itoh et al., 2021).

In our research we built upon OSTHMDs to realise our concept of Computational Glasses. After having a concept of how we would provide visual assistance which we detailed in our workshop paper (Sutton et al., 2019) it became necessary to understand the limitations of the devices we were using, the potential for these issues to be solved in the form of the research being done to address them, and their lasting impacts on our research. We looked to cover the spatial requirements of OSTHMD in their applications for general AR, as well as the temporal and visual requirements. We covered this in a survey paper (Itoh et al., 2021).

We designed a prototypical system to realise our concept of creating Computational Glasses by using OSTHMDs and integrated world cameras that were able to capture the world as seen by the users. Our prototype is designed to allow for precise modulation of a user's view of the world, and whilst we detail this in most of our papers, Sutton et al. (2022a) provides the most complete explanation. Chapter 3 covers the details of this component of our research.

Having developed a system to demonstrate our concept, our first practical research step was to demonstrate that Computational Glasses using OSTHMDs can be used as aids for the HVS. To this end we looked to address a common visual impairment that is not currently aided by traditional glasses, CVD. We developed a basic understanding of the impairment, then completed a thorough investigation of the prior works for compensating for CVD on computers.

Using the works most closely related to ours we created a set of techniques to be deployed on Computational Glasses that could compensate for CVD. We then conducted a series of studies to verify our concept. We started with efficacy studies to prove that the system could work before comparing it against alternative methods to aid CVD. We then looked to open up further research in the area, investigating the reception of our concept and the potential directions for future techniques to be deployed on the glasses, based on those seen in the prior literature.

In doing this we were able to confirm the plausibility of our concept as a means to provide a new form of aid for visual impairments. We published this research in two papers (Langlotz et al., 2018; Sutton et al., 2022a) and Chapter 4 covers the specifics of our research.

Having demonstrated the plausibility of our concept as a means to provide new aids for visual impairments, we wanted to investigate how our concept could be extended to general visual assistance by providing visual augmentations to the unimpaired HVS. We decided to look at providing visual guidance and assistance in finding areas of interest. In particular we viewed providing visual guidance by subtle saliency modulation as a promising means to achieve this in Computational Glasses, based on the results of prior works on similar styles of guidance in AR. We investigated the related works around visual guidance, looking at the general field of on-screen visual guidance in AR. Based on our investigation we developed a design space for on-screen visual guidance in AR.

As the design space was lacking in a saliency technique applicable to our Computational Glasses we developed such a technique. We ran a short study to find parameters for our technique and then investigated its efficacy as a subtle means of providing visual guidance. We then once again looked to open up our research and investigate potential future research directions for further guidance techniques by exploring more of our design space. We cover this research in two papers (Sutton et al., 2022b, TBD) and detail it in Chapter 5.

In summary, our research has demonstrated the potential for Computational Glasses to provide a means of visual assistance both to aid and to augment the HVS. As a first step, our research has produced several findings, and opened up future research directions. There are also several limitations to the research we have conducted. All of this is discussed in Chapter 6.

1.3.2 Key Contributions

This thesis contributes to the fields of assistive aids, augmenting humans, OSTHMDs, and AR with the key contributions including the presentation of:

The concept of Computational Glasses as a means to provide a new generation of vision assistance to support the human visual system, both as a means of aiding impaired vision, and as a means of augmenting standard vision.

A practical implementation of Computational Glasses using OSTHMDs that can demonstrate their potential efficacy, and can be used for the development of visual aids and also be used to conduct user studies. We present several prototypes of our implementation which range from well controlled bench prototypes to stereoscopic statically mounted prototypes that allow users to directly look through the glasses, to mobile prototype users can wear and experience, to demonstrative miniaturised prototypes that exemplify the potential for future miniaturisation of our work.

An investigation into the use of Computational Glasses to assist those with CVD, reporting on a series of user-studies with people affected by CVD to investigate different aspects of efficacy, and in varying degrees of controlled environments. Our results show that by using Computational Glasses the effects of CVD can be compensated for, particularly in the most extreme cases. Our results also show that our concept performs on par with state-of-the-art approaches, even when they are provided with advantages, and that there is a general willingness from those afflicted with CVD to explore solutions such as our own. Finally, our results indicate future potential research directions for compensation techniques with colour shifts proving effective but potentially over applied, and need to explore more generalisable scenarios.

An investigation into augmenting human perception and the HVS by supporting visual guidance using Computational Glasses. This includes a design space for visual guidance methods in AR, the development and study of a saliency modulation algorithm tailored for OSTHMDs, and a study of the various styles of visual guidance present in the space. Our design space demonstrated a lack of exploration in various area, for example in OSTHMDs and in particular saliency modulations for them. We present our saliency modulation which we investigate as a subtle means of visual guidance. Our results demonstrate the desired efficacy of our approach in guiding attention whilst allowing continued viewing however, we were unable to attain the intended subtlety. Our subsequent exploration of the design space showed the potential for different approaches to providing visual guidance to be effective in different manners, with varying overtness so being applicable in different situations. Our results also show potentially interesting research directions for future research in our space, such as further combination of styles.

A discussion of the future outlook and research directions established by this thesis. We reflect on the addressing of our hypotheses and lessons learn in doing so such as the importance of utilising various prototypes that can account for differing confounding variables and help verify the results from each other. We briefly touch on some considerations for the concept, such as social acceptability and ethical application. We then discuss some of the limitations of our work, such as the constraints of current commercial OSTHMDs. Finally, we consider the future research directions our work reveals. This covers aspects such as directions for future modulation techniques and components therein, and extending study environments.

1.3.3 Publications

Throughout the course of this research, we have produced five papers which are either published or under review (four published, one under review) and have one paper that is a work in progress where we are running final studies as our initial time-frame was impact by Covid-19. The research done in these papers forms the basis of this thesis and many parts are included from them. Information may have been copied verbatim into this thesis in cases where I was the first author and completed the initial draft with remaining authors mainly revising the paper. I rewrote content for papers where I was involved but not first author unless I wrote the sections myself. The following provides more information on these papers and my role.

ChromaGlasses: Computational Glasses for Compensating Colour Blindness

Tobias Langlotz, ***Jonathan Sutton****, Stefanie Zollmann, Yuta Itoh, Holger Regenbrecht*
ACM CHI Conference on Human Factors in Computing Systems, *CHI '18*, Montreal, 2018

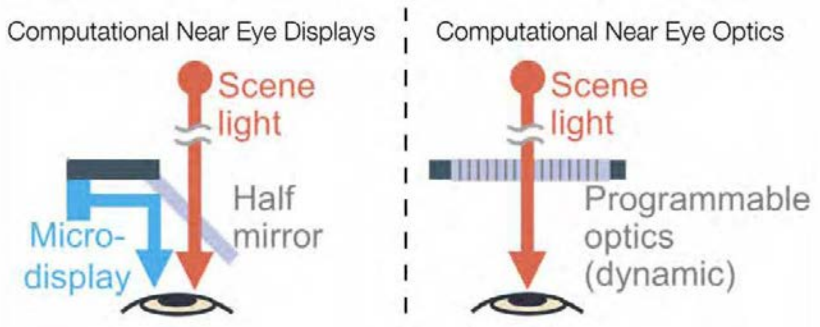
In this paper, we presented our first exploration of Computational Glasses. We looked to build on the prototype of (Langlotz et al., 2016) to modify OSTHMDs to create Computational Glasses that could compensate for CVD. For the purpose of this paper, we termed this implementation of Computational Glasses, ChromaGlasses. We created a set of hardware prototypes for demonstrating our concept, then implemented compensation techniques based on those in the related literature. We then ran three user studies with CVD affected participants to evaluate ChromaGlasses.

My main contributions to this paper were setting up the hardware prototypes, implementing the software compensations, and running the user studies, which I also helped design. As I did not have a large contribution to the writing of this paper, any information included from it has been rewritten.

The research from this paper is covered in Chapter 4 of the thesis.

(Langlotz et al., 2018)

Computational Glasses: Vision augmentations using computational near-eye optics and displays



Jonathan Sutton, *Tobias Langlotz, Yuta Itoh*
IEEE International Symposium on Mixed and Augmented Reality Adjunct,
ISMAR-Adjunct '19, Beijing, 2019

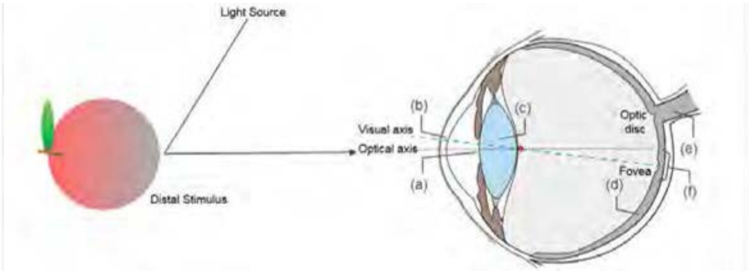
In this paper we formally covered the concept of Computational Glasses that was first detailed in *ChromaGlasses* (Langlotz et al., 2018). We covered the core idea and recent works that would be considered examples of Computational Glasses. We then detailed our own experiences working with the concept ((Itoh et al., 2019b; Langlotz et al., 2018)) and reflected on the lessons learned for future development.

As the first author my contribution to this paper was writing the first draft, which was then iterated on with help from my co-authors. As such, parts from this paper may be copied verbatim.

The research from this paper is covered in Chapter 3 of the thesis.

(Sutton et al., 2019)

Towards Indistinguishable Augmented Reality: A Survey on Optical See-Through Head-Mounted Displays



*Yuta Itoh, Tobias Langlotz, **Jonathan Sutton**, Alexander Plopski*
ACM Computing Surveys, *ACM CSUR*, 2021

This survey paper covers the state of OSTHMDs, looking at what is needed for them to achieve coherent AR when matched against the HVS. We summarised the current research looking to achieve indistinguishable AR in various key areas and reflect on future outlook and issues.

My primary contributions to this paper were writing one section on the physiology of the HVS and another on the state-of-the-art OSTHMDs designs to overcome resolution constraints. As such these sections may be copied verbatim in this thesis, all other information taken from this paper has been rewritten.

The research from this paper is covered in Chapter 3 of the thesis.

(Itoh et al., 2021)

Seeing Colours: Addressing Colour Vision Deficiency with Vision Augmentations using Computational Glasses

***Jonathan Sutton**, Tobias Langlotz, Alexander Plopski*
ACM Transactions on Human-Computer Interaction, *ACM ToCHI*, 2022

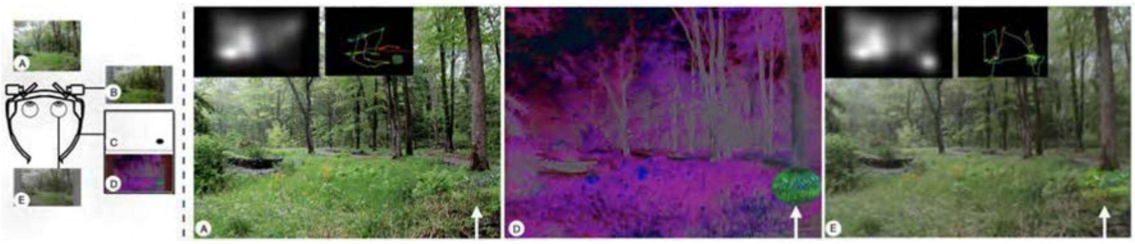
Extending our initial work on using Computational Glasses to assist CVD, we summarised the prior literature in general CVD compensations to inform further development of techniques to be deployed on Computational Glasses. We then evaluated a set of these via a user study, alongside a further user study to evaluate subjective feedback in less constrained conditions. We reported on these studies, alongside revisiting our prior studies to provide a cohesive overview and summary of our investigations into using Computational Glasses to assist CVD.

As the first author my contribution to this paper was writing the first draft, which was then iterated on with help from my co-authors. As such, content of this paper has been copied verbatim.

The research from this paper is covered in Chapter 3 and Chapter 4 of the thesis.

(Sutton et al., 2022a)

Look over there! Investigating Saliency Modulation for Visual Guidance with Augmented Reality Glasses



Jonathan Sutton, *Tobias Langlotz, Alexander Plopski, Stefanie Zollmann, Yuta Itoh, Holger Regenbrecht*
ACM Symposium on User Interface Software and Technology, *UIST '22*, 2022
(Conditional Acceptance)

In this paper we presented a method to adjust saliency in OSTHMD. We looked to demonstrate how subtle adjustment of the environment could be achieved using Computational Glasses using a custom algorithm. We present our algorithm and a user study investigating its efficacy. This paper is currently awaiting submission

As the first author my contribution to this paper was writing the first draft, which was then iterated on with help from my co-authors. As such, parts from this paper have been copied verbatim.

The research from this paper is covered in Chapter 5 of the thesis.

(Sutton et al., 2022b)

A Visual Guidance Design Space for On-Screen AR

***Jonathan Sutton**, Tobias Langlotz, Alexander Plopski*
TBD

In this paper we present a design space for visual guidance in AR. We present the space and works in it. We then look to investigate the space as it pertains to OSTHMD via a series of user studies. This paper is still work in progress with only the first study completed.

As the first author my contribution to this paper is writing the first draft, which will then iterated on with help from my co-authors. As such, parts from this paper have been copied verbatim.

The research from this paper is covered in Chapter 5 of the thesis.

(Sutton et al., TBD)

Chapter 2

Background

This chapter covers the relevant information to understanding the basis of this thesis and our research. It covers basic concepts on how the human visual system (HVS) works, light perception, and some of the many properties of it, focusing on those that impact our research. The conclusion of this includes information on the various ways the HVS can be affected, some currently able to be aided and others the output of this research hopes to aid in the future. It then briefly covers augmented reality (AR) techniques and the optical see-through head-mounted displays (OSTHMDs) that we built our prototypes on. Finally, it covers prior research into producing new wearable means to assist the HVS, both to aid and to augment it.
This chapter draws on the related work from all the papers, but primarily from Itoh et al. (2021); Sutton et al. (2019).

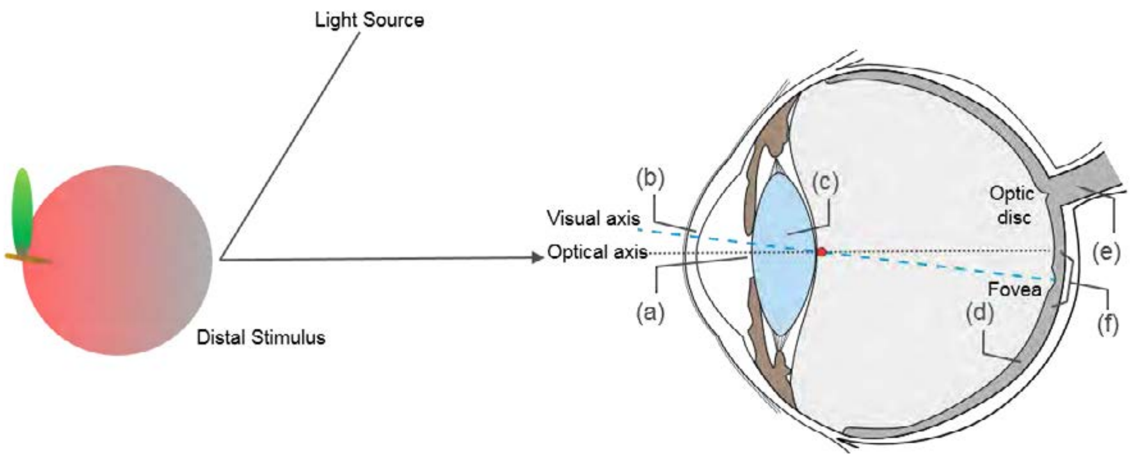


Figure 2.1: Schematic drawing of a cross-section of the human eye as viewed from the top of the head. The drawing is inspired by (Lindsay and Norman, 2013). (a) Pupil. (b) Cornea. (c) Lens. (d) Retina. (e) Optic nerve. (f) Macula.

In order to understand the research and contributions of this thesis, a brief understanding of several key areas of research is required, and these are covered here. Namely: knowledge of the physiology of the eye and the immense capabilities of the *human visual system (HVS)* is necessary to understand the system we are trying to interact with using our glasses; an understanding of how the HVS can be impaired and the impairments that have the potential to be assisted by Computational Glasses; *augmented reality (AR)*, the devices used and it's applications, as the foundations from where we developed our research; and finally, prior research in the field to assist the human visual system with wearable aids both to aid and to augment it.

2.1 Human Visual System

Before discussing the potential ways to assist the HVS, an understanding of some properties of the HVS is required. The following provides a quick overview of the human eye and the properties that are relevant to this thesis.

2.1.1 Light and the human eye

Visual light (with wavelengths from 390–750 nm) in the form of emitted photons is the basic requirement for visual perception. The HVS begins with the eyes (Lindsay and Norman, 2013) (Fig 2.1). Light is reflected from a distal stimulus (the stimulus providing information to the receiver) to the human eye. Once the light reaches the eye, it enters through the pupil (Fig 2.1a) and is focused onto the retina (Fig 2.1d) at the back of the eye by the cornea (Fig 2.1b) and lens (Fig 2.1c). The retina has receptors called rods and cones that contain light-sensitive chemicals known as visual pigments. Upon receiving light, an electrical signal is triggered by these pigments and sent to the brain by a network of neurons along the optical nerve (Fig 2.1e), where they are then processed. Colour vision is the result of different response functions from the cones to incoming light waves, in particular red, green, and blue. These cones are also termed *Long*, *Medium*, and *Short*, respectively, due to the wavelengths that produce the strongest response in each type. Compared to cones, rods are more sensitive to light (around 100 times more sensitive than cones) and thus contribute to seeing in dim environments. They contribute little to colour perception however, and are most sensitive to wavelengths in the green and blue spectrum of visible light. They also react more slowly to changing light conditions, whereas cones are responsible for colour vision and react faster, allowing us to perceive fine details.

The distribution of these rod and cone receptors on the retina is non-uniform, which affects our visual perception. The macula is a small area on the retina that is responsible for perceiving a sharp image, with a tiny area called the fovea being the area with the sharpest vision. The fovea has a high density of cone receptors. This ratio of cones to rods changes quickly as the density of cones decreases towards the periphery of the macula. Outside the macula, we find mainly rods, although with decreasing density when moving towards the periphery of the retina.

Whilst the subsequent processing of the electrical signals from the eye are complex and beyond the scope of this thesis (see Goldstein and Brockmole (2016) for a starting point), one area of relevance is the concept of saliency and its impact on attention. Early work in cognitive psychology has given evidence of a relationship between the properties of a scene and the attention applied to it. Neurons processing electrical signals from the eye will group them and respond to different features such as colour and contrast (Goldstein and Brockmole, 2016). Treisman and Gelade (1980) have shown how these various features are processed in parallel across the visual field, and that attention is placed based on these features to process them into complete objects. This feature-based process is commonly referred to as bottom-up saliency. It describes the influence of aspects of a visual scene upon where attention is placed, regardless of conscious influence. Features such as colours in contrast to their surroundings, and objects with different shapes or orientation will draw attention (Goldstein and Brockmole, 2016). Figure 2.2 demonstrates the effect saliency can have on attention. The other commonly given aspect of saliency is top-down saliency that describes the influence of conscious effort and goals on where attention is focused.

2.1.2 Properties of the human visual system

The HVS has several notable properties that are of relevance to our research into assisting it with modulations and *optical see-through head-mounted displays (OSTHMDs)*, such as perceivable wavelengths, dynamic range, visual field and resolution, latency and perceptible flicker, accommodation, and perceptual criteria such as the relevance of stereopsis or occlusions within the HVS and human perception.

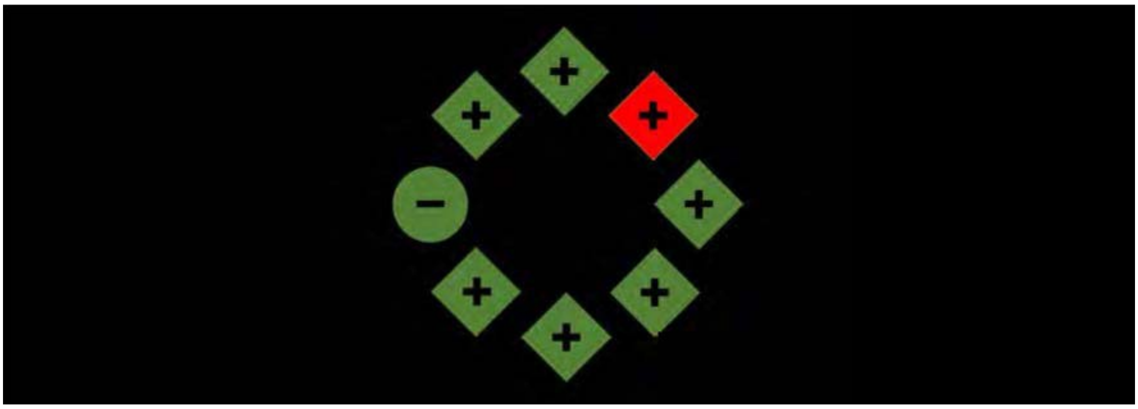


Figure 2.2: An example of the effect saliency can have on attention based on Goldstein and Brockmole (2016). When asked to find the circle then shown this image people will first look to the more salient orange diamond based on the bottom-up saliency of the image. This is despite top-down effort being placed to look for a circle.

Colour: The colour sensitivity of the human eye allows wavelengths between 390 nm and 750 nm to be perceived. Because of the distribution of cones, colour sensitivity is not uniform across the eye or across the spectrum of wavelengths. The fovea has a greater sensitivity to red and green chromatic stimuli and there is less across the periphery, while blue sensitivity is generally lower than that of red or green but more uniformly distributed between the fovea and the periphery (Noorlander et al., 1983).

Dynamic range: Research has found the static dynamic range of the human eye (assuming a constant aperture or pupil diameter) to cover 3.7 log units in cd/m^2 (Kunkel and Reinhard, 2010). However, the human eye has a focal length of 17 mm, and the pupil diameter can vary between 2 mm and 8 mm. This gives the human eye an equivalent aperture range of f/2.125–f/6.5. Changing the pupil and consequently the aperture allows us to drastically increase the perceivable overall dynamic range. In fact, the dynamic range can be further increased by mechanical processes and photo-chemical, neural-adaptive processes that transcend the scope of this paper (Goldstein and Brockmole, 2016). Research has shown that the cones in the eye are stimulated from 0.01 to 10^8 cd/m^2 when the cones are active and that rods are stimulated from 10^{-6} to 10 cd/m^2 (Kunkel and Reinhard, 2010), leading to a practical overall dynamic range of approximately 46.5 stops. Note that while the human dynamic range can go beyond 10,000 cd/m^2, this can be damaging (Robert, 2002).

Visual field: (Strasburger and Pöppel, 2002) found that the human eye can receive light information of up to 90° eccentricity from the vertical meridian towards the periphery, known as the temporal visual field, and 50° to 60° from the meridian to the nasal, known as the nasal visual field. The vertical visual field extends 50° to 60° in both directions from the horizontal meridian. Overlap in the visual field occurs in the nasal visual field, allowing true binocular depth perception and giving an overall horizontal visual field of up to 180°. Other sources have suggested that the *Field of View (FoV)* may be as wide as 190°, and some works have stated that when allowing the movement of the eye within a stationary head, the visual field increases up to 290° (Howard and Rogers, 1996). It should be noted that the FoV of the human eye is known to decrease with increasing age, so numbers have been taken with caution, as age is often not reported.

Acuity: Within this visual field, the acuity of the human vision is often measured in *cycles per degree (cpd)*, which is the number of cycles (a black circle with a white space around it) that can be distinguished within one degree of vision. Using a standard Snellen chart to assess visual acuity, a result of 20/20 (considered normal vision) is equivalent to 30 cpd (Yanoff and Duker, 2018). However, the visual acuity of the HVS is largely dependent on two factors: eccentricity and environment luminance. Under low lighting where scotopic vision is used, cpd can fall to as low as 2, while under high luminance and photopic vision, it can reach 50–60 cpd (Ferwerda et al., 1996; Wandell, 1995), and cpd falls off in an approximately linear fashion as eccentricity increases (Mckendrick and Johnson, 2003).

Resolvable resolution: The combination of the visual field and acuity indicates the required pixel resolution. To cover the visual field with static eyes in a singular display and have 20/20 vision (which is equivalent to 30 cpd or 1 arc minute (Yanoff and Duker, 2018)), a resolution of 11400×7200 is required, which equates to 82 megapixels, although we assume a rectangular FoV, which we know is false. If we expand this to the full 60 cpd approximated as the maximum under ideal conditions, these values increase to 22800×14400 and 328.3 megapixels. As the OSTHMD typically presents one image to each eye, the requirements increase further, each eye requiring 18000×14400 and 259.2 megapixels, giving 518.4 megapixels in total before considering motion: at 30 cpd, this total megapixel value falls to 129.6. Notably, these values are approximations that indicate what the equivalent pixel resolution to the human eye would be. To resolve this resolution, the human eye performs rapid movements to focus image areas in the fovea. These rapid movements are called saccades and help scan the environment. Saccades affect perceptions of scenes (e.g., saccadic masking), but a full exploration of this topic transcends the scope of this work.

Latency: VR studies have shown a limit to the latency which humans can recognise. Jerald and Whitton, for instance, suggested that immersive VR HMDs require a latency of 5 ms or less to be unnoticeable for the most sensitive subjects (Jerald and Whitton, 2009). These findings confirmed earlier work from psychophysical studies that found noticeable latency to be lower than 17 ms (Jota et al., 2013). Other work on interface interaction by Jota et al. (2013) showed that noticeable latency was 20 s to 100 ms, with 85% of participants unable to perceive latency below 40 ms when finger tapping (e.g., tapping-based input, such as tapping on touchscreens), but according to Ng et al. (2012), user performance was affected by a latency as low as 2.38 ms with a mean of 6.04 ms.

Critical flicker frequency (CFF): CFF is the fastest rate at which a flicker is perceived and not viewed as a stable image. This value varies with the adaptation level of the retina, size of the stimulus, luminance, and wavelength. It is not constant across the visual field, increasing up to 55° eccentricity before decreasing again toward the periphery, with a measured maximum value of approximately 90 Hz (Mckendrick and Johnson, 2003). Under extreme cases and using specific patterns, however, people have perceived flickers up to 500 Hz (Davis et al., 2015).

Depth perception: The ability of humans to perceive the position and size of objects is determined by several factors, including accommodation and occlusion. For the human eye to view objects at various distances in focus, the human eye has developed accommodation mechanisms. The shape of the lens can be adjusted by minute ciliary muscles in the eye, changing the curvature of the lens and subsequently moving the focal points. This allows distal stimuli from different distances to be focused on the retina. This is supported by vergence—the movement by both eyes to allow binocular vision—which is tightly coupled with eye accommodation. This works by triangulating the positions of objects in each eye with a fixed distance between the eyes, known as the *interpupillary distance (IPD)*, which varies between people and thus needs to be adjustable. Occlusions comprise more important information in the HVS eye, determining whether objects lie in front of one another and how large objects may be compared to familiar objects in view. As we later show, the ability to create virtual occlusions (virtual objects occluding real-world objects) is a major challenge for OSTHMDs but is important for visual perception and scene understanding.

Other: Other visual phenomena such as scattering occurs when light is randomly deflected in random directions, causing vision degradation due to image blur (Yanoff and Duker, 2018). Aberrations, of which there are many subcategories, occur when light incidents prevent image formation and otherwise limit human vision.

2.1.3 Visual Impairments

The number of impairments that can affect the HVS are myriad, and whilst some have been previously mentioned, many have overlapping impacts and here we instead focus on some of the common ways in which these impairments affect vision. Some of these have been simulated in Figure 2.3.

Inability to focus in certain depth ranges. Probably the most commonly known visual impairment is that of an inability to focus on certain depths. This can be an inability to focus on distant objects, caused by myopia (short-sightedness), an inability to focus on near objects caused by hyperopia (far-sightedness) or presbyopia, or a general blurring of the vision at all depths caused by astigmatisms. Other issues such as cataracts, diabetic retinopathy, and macular degeneration are also known to cause blurred vision. Less commonly, high order aberrations in the lens of the eye can also cause blurring.

Traditional treatments for these issues are the optical glasses we see in today's society. These static lenses are created with fixed dioptres to adjust light entering the eye such that it can be drawn into focus. More recent advancements have led to the bifocal glasses were one half (horizontally) has a different dioptre to the other. This can also be achieved with insets. More complex forms that introduce a

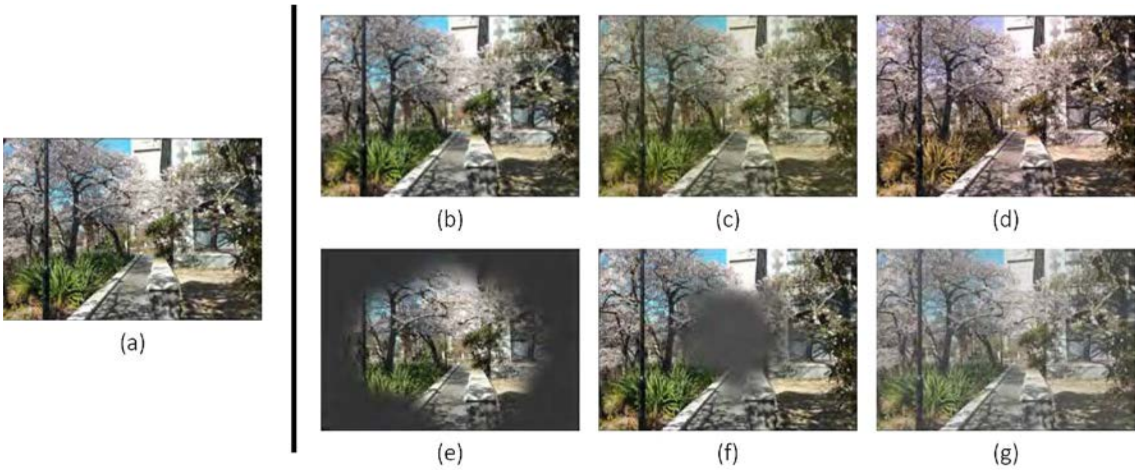


Figure 2.3: Simulated effects of visual impairments. (a) the original image, (b) blurred vision (astigmatism), (c) hazy vision (cataracts), (d) reduce colour vision (colour vision deficiency), (e) reduced peripheral vision (retinitis pigmentosa), (f) occluded central vision (macular degeneration), and (g) reduced contrast (diabetic retinopathy). Effect simulations created using "Impairment Simulator"[a] and "Coblis — Color Blindness Simulator"[b].

[a]inclusivedesigntoolkit.com/simsoftware/simsoftware.html
[b]www.color-blindness.com/coblis-color-blindness-simulator/

transition between the two halves are also possible and known as transitional lenses. Surgeries can also be performed to correct for these issues.

Cloudy vision. Whilst similar to a loss of focus, cloudy vision occurs when the transparency of the cornea is reduced, causing vision to take on a general hazy quality. This can make details harder to identify, such as edges and transitions of surfaces, text, or facial features.

This can be caused by several impairments with common ones being macular degeneration, cataracts, retinopathy, or glaucoma. Other than surgeries for cataracts, there is generally no accepted vision aid for cloudy vision.

Halos and glare. Accompanying hazy vision is an issue of halos appearing around light sources, often in darker lighting conditions. Similarly, light sources can produce a glare. These can affect vision, causing squinting and reducing contrast.

Various conditions and their solutions can cause halos and undue glare. For example, cataracts, and it can also be brought on by the corrective surgery. Myopia, hyperopia, and astigmatisms can also be a cause of halos, with glasses also introducing glare, particularly with thicker lenses.

Other causes for halos and glare include glaucoma and Fuch's dystrophy.

Reduced ability to differentiate between colours. Commonly caused by some form of *colour vision deficiency (CVD)* (covered in 4.1), a common issue with the HVS is a reduced ability to differentiate between colours which are normally considered readily differentiable. Whilst everyone's ability to differentiate between colours varies, it can be effected to such a degree that it is considered an impairment. This can be caused by genetics such as many forms of CVD, (Deeb, 2005). Another leading cause is age related with yellowing of the corneas which causes one's view to be tinted yellow, adjusting colours, and causing issues.

Currently CVD remains untreatable and commercial glasses to aid are of limited effect.

Reduced peripheral vision. Also known as tunnel vision, the peripheral vision of the HVS can deteriorate, leading to a reduced ability to detect objects in the periphery. This can continue to a complete loss of vision in the periphery causing the world to be viewed as though through a tunnel. It can develop slowly over time with age, going unnoticed if not checked up on.

The most common cause of peripheral vision loss is glaucoma causing damage to the optical nerve but can also be caused by conditions such as retinitis pigmentosa or strokes.

An alternative form of vision loss is when half of the FoV is lost, generally laterally, but can also be horizontally. This is the effect of hemianopsia and can be the result of strokes, brain trauma or tumours. This can affect either the outer side of the FoV in each eye, the left or right half of each eye, or the upper or lower half of each eye.

Currently there is no accepted form of assistance for tunnel vision, with treatment of the underlying condition and prevention of further deterioration being the only courses of action.

Blind spots in central vision. Whilst we all have a blind spot due to the optic nerve, we can naturally compensate for this. However, for many people their central vision is impaired with spots that are not naturally compensated for and cause an inability to detect and identify objects directly in view. This is known as a scotoma. These can be central (directly in the centre of the FoV), paracentral (slightly off centre) or scintillating (non-static).

This can be caused by various issues, such as migraines, macular degeneration, diabetic retinopathy, glaucoma, or stroke.

There are currently no accepted means of aiding scotoma.

Low contrast sensitivity. A loss of contrast sensitivity reduces the ability to tell contrast between objects, making separating the foreground from the background difficult. This can make this like judging depths and reading difficult and can impact locomotion. Low contrast sensitivity can often be caused by glaucoma, cataracts, amblyopia, and macular degeneration.

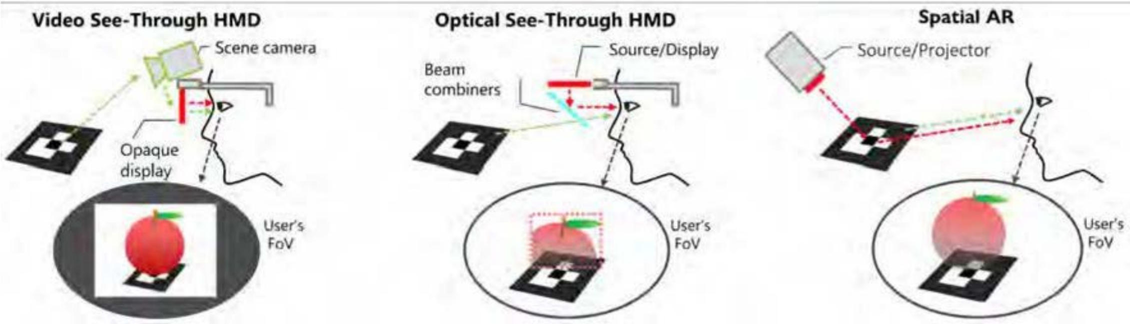


Figure 2.4: Common styles for creating AR.

Currently this remains an unaided effect of visual impairments and requires treatment of the underlying condition to be assisted.

2.2 Augmented Reality

Traditionally AR has been used to introduce virtual objects to the user's view with the classic example of this being a cube as shown by Ivan Sutherland when he first demonstrated an AR *head-mounted display (HMD)*. Today AR has seen many developments and is utilised to introduce various virtual objects to the real world, extending reality, for many different applications. Visualisation of information is a common example of this with applications for providing extra information about locations (Tatzgern et al., 2016) or objects (Tatzgern et al., 2014).

Further uses for AR have been found in providing support in various tasks such as navigation (Mulloni et al., 2011; Arntz et al., 2020), search (McNamara et al., 2008), and item picking (Schwerdtfeger and Klinker, 2008; Schwerdtfeger et al., 2009). In recent years, adding virtual content to assist with procedures in medicine (Eckert et al., 2019), and assembly and maintenance of machinery (Jeffri and Rambli, 2020) which has seen a particular interest in remote assistance (Oda et al., 2015), have become commonly noted applications.

Integration of virtual content has also seen use for creating AR games (Poupyrev et al., 2000), producing entertainment (Pucihar and Coulton, 2015) and providing education (Yuen et al., 2011).

2.2.1 AR Displays.

There are various ways in which virtual content can be integrated into the real world stimuli to create AR.

OSTHMDs were the first method demonstrated (Sutherland, 1968). This method works by mounting optics in front of the user's eye with a headset. The optics can be seen through, and whilst not obscuring the user's view of the world, can project light into their eyes or modify parts of the light from the real world entering their eyes, adding virtual content to it (Sutherland, 1968). This is shown in Figure 2.4 Center.

Because the view of the real world is not obstructed, the user is able to perceive the light from the real world and the virtual content simultaneously. The light is combined in the user's visual input augmenting reality in a perceptual space. Using tracking and computing per-eye images, virtual content can be displayed as though it is placed in the real world, with its location in the user's view being updated to remain consistent with the real world.

Spatial augmented reality (SAR) works by projecting light into the world, typically using projectors which are often static (Bimber and Raskar, 2005), although these can also be mobile (Raskar et al., 2003). By projecting light onto various surfaces, they can be augmented with textures, information, or when combined with shutter glasses, 3D objects. See Figure 2.4 Left. By projecting light into the world, the virtual content is combined with the real world in the real space and can be perceived by multiple viewers. Although, this can be problematic with perspective projections and integrating multiple sets of 3D shutter glasses for virtual content to be viewed.

Video see-through head-mounted displays (VSTHMDs) work in a manner similar to OSTHMD in that they project light into a user's eyes via optics mounted near to the user's eyes by a headset. However, instead of being able to see through these optics to the real-world, displays are placed behind the optics and the view of the real world is blocked off. The real world is instead captured by a camera(s) placed on the headset, facing the world. The feed from the camera(s) can then be modified to include virtual content and is displayed to the user in the headset. See Figure 2.4 Right. In this manner virtual content is combined with the real world in camera space and then perceived by the user. Doing so offers for greater control over the addition of virtual content than OSTHMDs or SAR as complete control over the camera image is possible. However, it disconnects the user's direct view of the real world and constrains their view of it to what the display can represent of the cameras view.

Mobile devices, such as phones, can provide an AR experience in a similar fashion to VSTHMD, by rendering virtual content on the camera feed from the phone camera and presenting this to the user. However, they do not provide the same 3D experience, or replace the user's view of the world.

2.2.2 Optical See-Through Head-Mounted Displays

As this research utilises OSTHMDs, rather than cover all the various devices used for AR in further detail, we instead present a quick overview on OSTHMD here.

Whilst the concept of HMDs has existed since the 19th century with stereograms (Wheatstone, 1838), it was in the mid-20th century that the first OSTHMD was demonstrated by Ivan Sutherland's 'Sword of Damocles' (Sutherland, 1968). Optical elements of half-silvered mirrors in the optical path to the user's eyes partially redirected light from a computer-generated image on a small *cathode-ray tube (CRT)* towards the eyes. This integrated the image with the user's view and augmented their view of the world with virtual content.

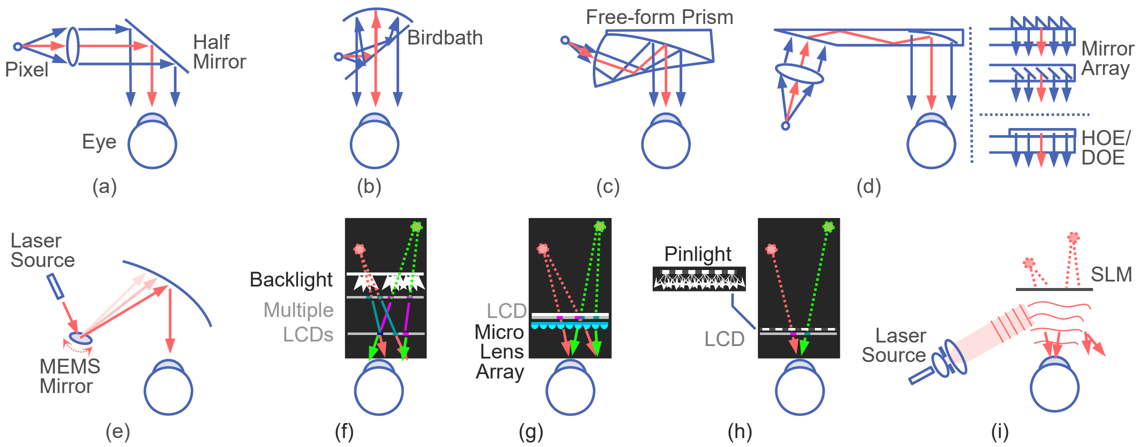


Figure 2.5: Schematic overview of different optical designs used in existing OSTHMDs. (a) Half mirror. (b) Birdbath combiner. (c) Free-form prism. (d) Waveguide (Kress and Chatterjee, 2021) with a coated curved mirror, a wedge-shaped mirror array, cascaded mirrors or holographic/diffractive optical elements. (e) Retinal scanning. (f) Light-field OST-HMD with multiple liquid crystal display (LCD) layers (e.g., (Maimone and Fuchs, 2013)). (g) Light-field OST-HMD with microlens arrays (e.g., (Yamaguchi and Takaki, 2016)). (h) Pinlight display (Maimone et al., 2014). (i) 3D holographic display.

Half-silvered mirrors provide the simplest of designs for creating an OSTHMD, making them readily manufactured and providing a good *eyebox* (Figure 2.5 (a)). However, their simple design limits their FoV relative to their size, and generally produces a bulky system as flat mirrors must be placed at angles to the eyes. The FoV can be expanded using a birdbath design (Figure 2.5 (b)), however at the cost of reduced light efficiency due to further half-silvered mirrors in the optical path.

To overcome the shortcomings of the traditional half-silvered mirrors designs, more complex optical elements have been introduced. Utilising waveguides where light is guided to the user's eyes via internal reflection of the element. The light is then projected into the user's eye via beam-combiners. Where the commercial Epson Moverio devices originally used simple half silvered mirror designs, modern examples use free-form optics as beam-combiners (Figure 2.5 (c))[1]. Other designs, such as the Lumus devices utilise arrays of mirror layers as optical combiners (Figure 2.5 (d)). Whilst free-form optics provide a reduction in relative size over their simpler counterparts, and arrays of mirrors further still, they remain relatively large. Research designs have also introduced further optical complexity using *holographic optical elements (HOEs)* (Figure 2.5 (d)). These elements are holographs that work by transmitting or reflecting light based on its frequency with the angle of reflection being dependent on frequency.

An alternative method to guiding light into the user's view is to directly project it into the user's eye via a laser light source. Using a laser projector or similar device specific to this purpose, this is known as a retinal-projection display or a

[1] moverio.epson.com/

retinal-scanning display (Figure 2.5 (e)) (Furness III and Kollin, 1995).

One common issue with all of the preceding styles, except retinal displays, is the *accommodation–vergence conflict*. Alternative designs for OSTHMDs that look to avoid this issue have also been developed. Stacking *liquid-crystal displays (LCDs)* in layers can be used to create multiple focal planes, akin to a light-field display but with very limited planes (Figure 2.5 (f)) (Akeley et al., 2004), or true light-field displays can be created using holographic displays (Figure 2.5 (g)) (Maimone et al., 2017). Designs that create an image that is always in focus have also been developed. Pinlight displays place an array of point light sources behind an LCD, creating a pinhole projection (Figure 2.5 (h)) (Do et al., 2019).

Whilst some research prototypes of OSTHMD demonstrate the ability to reduce the environment light when augmenting a user's view by using means such as LCDs (Wetzstein et al., 2010) or *phased spatial-light modulators (PSLMs)* (Itoh et al., 2019a,b, 2021), the vast majority of OSTHMD, particularly those commercially available, are constrained by their nature to only add light to a user's view. Whilst tinted shades can be used to provide a static reduction in environment light, the additive-only nature of OSTHMD remains a constant limitation. This limitation is reflected in our research as we rely on commercially available OSTHMD.

Note that Section 3.1 covers the limitations of current OSTHMDs, their impact on our research, and research looking to address them in greater detail.

2.3 Assisting the Human Visual System with Wearable Aids

Various wearable approaches have been demonstrated to assist with visual impairments. Whilst there are many approaches that look at augmenting the use of alternative senses like touch or sound (Ghafoor et al., 2019; Woźniak et al., 2015; Findlater et al., 2015), the focus of this research is on eyewear and aids that directly interact with and assist the HVS. As such these are focused on in the following.

Static Optics: As discussed, the traditional method for aiding visual impairments has been to use static optical lenses. Whilst these have been well covered in regard to their use to assist refraction errors, there are also research projects that have looked to modify them to produce new aids. Maybe the most prominent of these is the use of prisms integrated into the lenses in order to expand FoV for those with tunnel vision or hemianopsia (Peli, 2007; Peli and Jung, 2017; Jung and Peli, 2018; Houston et al., 2018). Further research on aiding hemianopsia has looked to integrate standard mirrors to re-present the lost information to another part of the FoV (Nooney Jr., 1972; Young, 1929). Other works creating new static optics have included contact lenses for CVD (Swarbrick et al., 2001).

Augmenting and assisting the HVS with wearable optical systems beyond aiding it is commonplace in today's society. Ubiquitous solutions such as sunglasses, are used as general aids by large numbers of people. Occupationally, society also regularly looks to augment the HVS to allow for the better accomplishment of tasks.

Examples include jobs such as welding, where masks are used to greatly reduce the amount of visible light, or loupes used for magnification in various occupations including dentistry and jewellers. Furthermore, there are many cases where we look to augment our visual system situationally. For example, when looking to see at extended distances binoculars or telescopes are employed, with magnifying glasses being employed in a similar fashion for small objects.

VSTHMD: The most controllable method for modulating vision is to use VSTHMDs. By capturing the environment via one or more cameras, the feed from the cameras can be processed and modified in real-time. This modified view is then presented to the users as a replacement of their own view of the world, effectively modulating their perception of the world. Examples of this style of aid include works such as CueSee (Zhao et al., 2016) and ForeSee (Zhao et al., 2015, 2019b) that look at providing guidance for *low vision* and various magnification and image enhancements for low vision respectively. The earliest works to demonstrate VSTHMDs as a potential method of vision assistance was the early work by Mann (1994) who described the use of cameras mounted on an opaque HMD with remote processing to provide zoom functionality and assistance with scotomas (Starner et al., 1997). This work was also realised in the form of a welding mask (Mann et al., 2012). Other research works include supporting impaired FoV (Younis et al., 2017; Bozzelli et al., 2020; Sayed et al., 2020), assisting those with night blindness to read (Fernandez et al., 2015) and systems to compensate for CVD (Melillo et al., 2017). Gonçalves et al. (2020) detailed a system to provide zoom on buttons to assist with macular degeneration. Commercially, devices such as the Samsung Relumino prototypes have even been demonstrated [2], which demonstrates the potential for miniaturised devices when custom built. Other commercial devices such as eSight [3], NuEyes [4] and IrisVision [5], provide magnified images with enhancements such as edge enhancements for low vision in a manner similar to ForeSee. An alternative form of VSTHMDs uses a depth sensor and an array of LEDs in a headset to provide visual assistance in the form of object shapes to partially sighted individuals (Hicks et al., 2013).

VSTHMDs have also been used to provide augmentations to the HVS. They have been used to expand the wearers FoV by replacing the world cameras with wide FoV lenses which are then displayed in the smaller FoV of the display (Yano et al., 2016; Orlosky, 2014). Reducing the effect of visual distractions in workplaces (Koshi et al., 2019), and integration of *infra-red (IR)* light with visible light in low-light conditions to improve visibility (Orlosky et al., 2017) have also been shown. An alternative approach to "augmenting" vision has been to enable those with normal vision to perceive the world as seen by those with visual impairments (Ates et al., 2015; Jones et al., 2020).

[2] samsungrelumino.com
[3] https://esighteyewear.com/
[4] https://www.nueyes.com/
[5] https://irisvision.com/

OSTHMD: An alternative way to produce wearable aids for visual assistance is to use OSTHMDs. Works by Peli et al. (2009) and Younis et al. (2017) looked at helping peripheral vision loss by reintroducing lost information into the central vision as an edge outline of the wider FoV (Peli et al., 2009) and providing object information in the remaining vision (Younis et al., 2017). Other works looked to use Google Glass to provide aid for CVD (Tanuwidjaja et al., 2014; Lausegger et al., 2017) by presenting modified cameras images to the user, provide zoom of phone screens (Pundlik et al., 2017) by presenting an enlarged version of the phone screen in the display, or edge enhancement (Hwang and Peli, 2014). Edge enhancement was also achieved with a HoloLens (Stearns et al., 2018). Other recent works that look to assist with visual impairments using the HoloLens focused on low vision, providing various forms of assistance such as coloured overlays (Kinateder et al., 2018; Angelopoulos et al., 2019), wall outlines and edge outlining (Min Htike et al., 2021), emotional cues (Lang et al., 2020), zoom (Stearns et al., 2017), assist targeting for interactions (Lang and Machulla, 2021), and reading signs (Huang et al., 2019). The HoloLens has also been used to provide eye tracking to aid amblyopia (Nowak et al., 2018) and FoV impairments (Zhao et al., 2019a). To assist with focus, the display of filters to adjust aberrations has been shown (Itoh and Klinker, 2015c). Also, to aid with night blindness by providing edge outlines (Bowers et al., 2004) and overlaying enhanced cameras image on the real world (Hu et al., 2015) have been proposed.

OSTHMDs have also been proposed to expand the wearers FoV (Orlosky, 2014). Further, they have been used to provide object tracking, (Itoh et al., 2016c), and occlusion capable OSTHMDs have been used to augment eye adaptation to varying light conditions (Hiroi et al., 2017).

Computational Glasses: Whilst this thesis presents leading work on Computational Glasses, since our initial work on presenting our concept (Langlotz et al., 2018), several other works have been published that could fall under our concept of Computational Glasses. Whilst we cover the concept of Computational Glasses further in the following chapter, these works are covered here.

A set of works that use OSTHMDs similar to our research, and completed subsequently, integrated an LCD and a user perspective camera to create a bench prototype to aid CVD (Tang et al., 2018, 2020).

Another work to use LCDs are the sunglasses developed by Hu et al. (2021) utilise two LCDs to provide variable dimming as an aid for the light sensitivity associated with autism spectrum disorder (Hu et al., 2021).

Various works have used near-eye optics to achieve vision assistance in a manner categorizable as Computational Glasses. Autofocals utilise focus-tuneable lens, eye-tracking, and depth sensors to adjust the dioptre of the glasses in real time to provide focus where the user is looking, extending traditional glasses into constantly variable ones (Padmanaban et al., 2019a). Another method to produce near-eye optics is to use PSLMs. By using one or more of these and a half-silvered mirror, prototypes can directly modulate the environment light being perceived by the user, modulating their vision by subtracting from the light they perceive (Itoh et al., 2019a,b; Kaminokado et al., 2020) with applications for aiding vision as variable bifocal glasses and colour aid for CVD, and augmenting vision with zoom or FoV shift.

One work that combines near-eye optics and near-eye displays to allow for accurate focus correction of both real and virtual content is FocusAR. These glasses work by using a varifocal display and tuneable focus lens (Chakravarthula et al., 2018). Alternatively, DehazeGlasses utilise a *digital micromirror device (DMD)* and an OSTHMD (Hiroi et al., 2020).

Singh et al. (2021) has also proposed the design of a contact lens for visual aid.

2.4 Summary

The HVS is a complex system that starts from the eye and allows for the perception and interpretation of light. The makeup of the HVS provides several properties and capabilities that we must consider when we look to assist it. It is not however invulnerable and can be affected by many conditions, either from its formation or developing over time. Whilst issues with focusing, such as for those with myopia or hyperopia, can be assisted with commonly used aids, many issues such as cloudy vision or inhibited colour vision cannot. Instead, these issues remain unassisted unless the underlying condition, such as cataracts, can be treated which is not possible in many cases.

In recent years, AR has arisen as a common way to modify the real world by adding virtual content to it by affecting the vision of perceivers. This can be achieved in the perceptual space of the eye using OSTHMDs, the real space using SAR, or camera space using VSTHMDs. In particular, OSTHMDs, working in a perceptual space to affect the vision of individuals whilst maintaining their view of the real-world show promise as a new means to produce vision aids. These devices project light into the user's eyes via a transparent optical path that the user can see through. The projected light combines with the light from the real world and allows for virtual content to be perceived by the user. Various methods with varying strengths and disadvantages have been developed for creating OSTHMDs.

Means to aid the HVS are commonly accepted and whilst some research has looked into creating new static optics to use as aids (Peli and Jung, 2017; Nooney Jr., 1972), others have looked to use head mounted AR devices and techniques to provide aid. Research has looked at using VSTHMDs in various ways, with approaches such as providing magnification and applying image enhancement techniques (Zhao et al., 2020), which have even seen commercial interest. However, such techniques apply constraints to the overall information perceived by the HVS, occluding the user's eyes from the real world. Alternatively, works have used OSTHMDs for assisting with depth perception (Min Htike et al., 2021) and assisting with peripheral vision loss (Peli et al., 2009). These works all rely on either presenting the user with an alternative perspective of the world overlaid on their own or suffer from misalignment issues.

Similar to aiding the HVS, some research has looked to augment it using VSTHMDs to expand FoV (Yano et al., 2016) amongst others, and OSTHMDs to augment the HVS's ability to adapt to variances in light (Hiroi et al., 2017). However, these works come with the same issues of those to aid the HVS.

There have been a limited number of works that have created new means to assist the HVS that fall under the same concept as covered in this thesis, such as those looking to create tuneable glasses (Padmanaban et al., 2019a; Chakravarthula et al., 2018), although they are limited to a single application. Other, adjacent work to our own has looked at directly adjusting the light from the real world that the user perceives with application to various forms of aid and augmentation (Itoh et al., 2019a,b).

However, as of yet, no research has demonstrated the means to create Computational Glasses utilising an OSTHMD that is capable of holistically assisting the HVS as both a means to aid and augment it. Whilst works have utilised OSTHMDs they are unable to modify the light perceived by a user on a per-pixel basis, modulating their view of the world as needed.

Chapter 3

Computational Glasses

This chapter covers the core concepts behind our research. We briefly introduce our concept of Computational Glasses before covering the state of optical see-through head-mounted displays (OSTHMDs) in research. We relate this to coherent integration with the human visual system (HVS), as we look to utilise OSTHMDs in modulating the HVS. We then cover in detail the concept of Computational Glasses and our implementation of them.

The contributions of this chapter are an introduction to the concept of Computational Glasses, a summary of the state of OSTHMD as it pertains to our research, and details on how we chose to implement Computational Glasses.

This chapter draws on Itoh et al. (2021), Sutton et al. (2019), and portions of Sutton et al. (2022a).

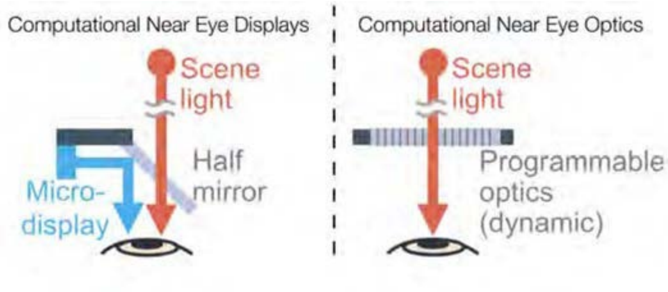


Figure 3.1: Illustration of two forms of Computational Glasses. (Left) Using computational near-eye display to change the appearance of the physical world. (Right) Using computational near-eye optics.

The vision of Computational Glasses is to precisely manipulate input from the real world to support the visual perception of the user. They look to provide visual assistance in aiding visual impairments above and beyond what is currently achieved in society, and in augmenting the *human visual system (HVS)*. Precise manipulation of input from the real world should be possible within a wearable device that, in future, will resemble traditional glasses and consequently be optical see-through. We see three ways that this can be achieved; with Computational Glasses using semi-transparent near-eye displays (Figure 3.1, Left), with Computational Glasses using computational near-eye optics (Figure 3.1, Right), and approaches that combine both, displays and programmable optics. Whilst near-eye displays could be classed as programmable optics, we choose here to differentiate them based on the long standing use of near-eye displays in the related literature as a *spatial-light modulator (SLM)* placed near to the users eye is used to introduce light to the user's view via a non-programmable optical element. For example an OLED display which is directed to the user's view via a free-form beam-combiner in some *optical see-through head-mounted displays (OSTHMDs)*[1] or a lens in many *video see-through head-mounted displays (VSTHMDs)*[2]. We differentiate this common understanding from other more recent approaches where optic elements are programmed not as traditional displays, but as means to directly operate on the existing scene light. For example focus tuneable lenses (Chen et al., 2021) which have been demonstrated as AutoFocals (Padmanaban et al., 2019a), and the programmable colour filter developed by Itoh et al. (2019a). In our research we utilise the most mature form in near-eye displays and OSTHMDs. However, these devices are not without their shortcomings when looking to create Computational Glasses.

[1]https://www.epson.co.nz/products/ProjectorAccessories/Moverio_BT-300.asp
[2]https://www.vive.com/nz/product/vive-pro-eye/specs/

Requirement	To Equal HVS	Hololens 2
Combined Static FoV	180–190° × 100–120°	43° × 29°
Per Eye Static FoV	140–150° × 100–120°	
˜Combined Resolution	11400 × 7200 @ 30 cpd	calculated as ˜2021 × 1363
	22800 × 14400 @ 60 cpd	
˜Per Eye Resolution	9000 × 7200 @ 30 cpd	2k [a]
	18000 × 14400 @ 60 cpd	
Static Dynamic Range	3.7 log units (cd/m^2)	500 cd/m^2 [b]
Overall Dynamic Range	10^{-6} cd/m^2 - 10^8 cd/m^2	
Latency	<2.38ms	Display: 16.66ms (60 FPS)*

Table 3.1: Requirements for an OST-HMD to match the HVS. We also present the equivalent values as they have been reported for the Hololens 2. This represents a recent, commonly known commercial application of OST-HMD and gives an idea of the current state of publicly available devices. The values used were reported by Microsoft[a] and KGuttag.com,[b] and we used reported diagonal FoV and pixels per degree to calculate the combined FoV and resolution.
**Note that the FPS for the Hololens 2 is only representative of the update rate and indicative of display latency. Overall latency will be higher.*

[a] docs.microsoft.com/en-us/windows/mixed-reality/develop/platform-capabilities-and-apis/rendering
[b] kguttag.com/2019/02/27/hololens-2-first-impressions-good-ergonomics-but-the-lbs-resolution-math-fails/

3.1 State of Optical See-Through Head-Mounted Displays

As summarised in Table 3.1, although not comprehensive (e.g., accommodation and sensitivity to geometric distortions), the capabilities of the HVS far exceed the capabilities of modern OSTHMDs and present a series of challenges that need to met to be able to achieve complete coherence between the real world and virtual content or modulations. Whilst these requirements do not always need to be strictly meet, commercially available devices are a way off the required capabilities to achieve this coherence (Table 3.1) and research has been investigating ways to better match the HVS. The issues required to achieve coherence and the research attempting to overcome them can be broken down into three categories: *spatial coherence*, looking at ensuring that virtual content is correctly located relative to the real world to ensure realism; *temporal coherence*, looking to achieve constant appearance of the content of time; and *visual coherence*, looking to match the visual properties of the virtual content with that of real counterparts.

3.1.1 Spatial Coherence

In order for virtual objects to be realistically rendered they must be correctly positioned spatially and maintain their spatial relationship with the real world. To

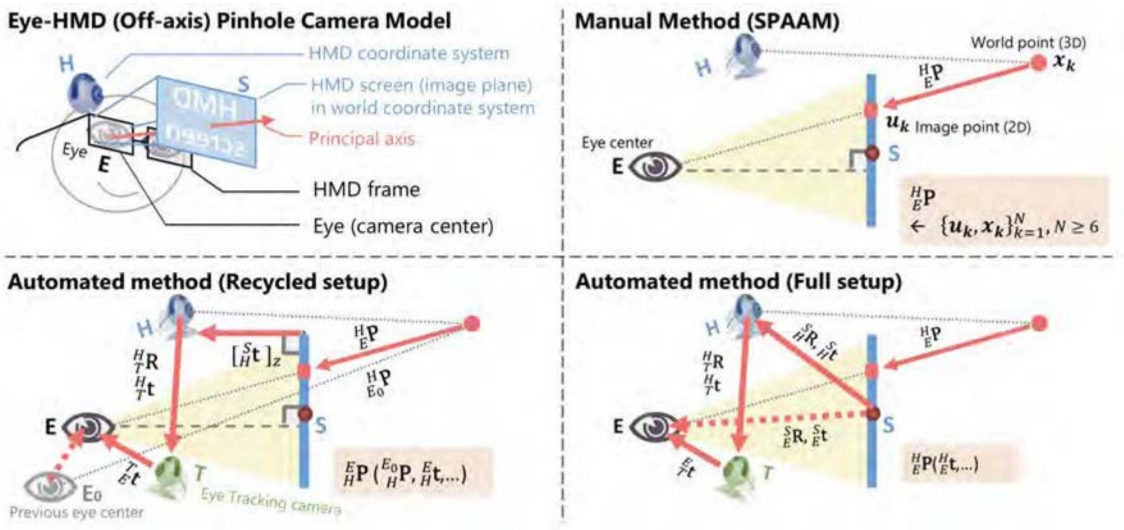


Figure 3.2: General overview of calibration techniques for calibrating the spatial relationship between the eye and display (Grubert et al., 2018).

achieve this correct registration of virtual content in the real world, to a user's eyes, the relationship between a user's eyes and the OSTHMD, as well as the relationship between the OSTHMD and the real world must be known.

The problem of finding the relationship between the OSTHMD and the real world is common to all *head-mounted display (HMD)* applications as one of motion or position tracking. For use in mobile situations, such as those provided by OSTHMD, the most popular solution is to use *inside-out* vision-based tracking, utilising environment-facing cameras. These cameras are used to either track synthetic markers, such as fiducial markers (Kato and Billinghurst, 1999; Wagner et al., 2008) or, more increasingly, use natural features as markers such as is demonstrated by *simultaneous localisation and mapping (SLAM)* (Klein and Murray, 2007; Reitmayr et al., 2010).

Whilst finding this relationship is an issue for general OSTHMDs, the nature of our Computational Glasses is not to add virtual content to the real world and therefore, with our solution, is not an issue in our work.

The second relationship (OSTHMD to eye or eye-display) is specific to OSTHMD and consequently of greater focus here. Grubert et al. (2018) detailed three categories of methods for computing eye-display relationships in their recent survey on calibration methods (Grubert et al., 2018): manual, semi-automated, and automated. An overview of these methods is covered in Figure 3.2. By assuming the eye-display relationship can be approximated using a pinhole camera placed at the position of the eye and viewing an image plane containing the virtual image in the OSTHMD the problem becomes one of estimating a 3D-to-2D projection. This assumption is common to all categories of methods, and solutions can be found by collecting corresponding data points between the 2D image and the 3D world. However, as directly fetching the user's view of the world is not viable, different methods are used to collect the data points.

Manual methods such as the commonly used method by (Tuceryan et al., 2002), *single-point active alignment method (SPAAM)*, rely on collecting data via user input. This is generally done by creating correspondences between a 2D crosshair and a known 3D point. These correspondences are then used to create a projection matrix by solving the system of linear equations.

Alternative to having using user input to collect all data points semi-automatic and automatic methods reduce the amount of online calibration completed by the user with semi-automatic methods such as Owen et al. (2004) calibrating the display portion of the eye-display relationship offline and only requiring the user to calibrate the eye component online. Fully automated methods completely remove the error prone user input with methods such as *interaction-free display calibration (INDICA)* (Itoh and Klinker, 2014a).

Finding the eye-display relationship is an important problem for running user studies in our research, and the general application of Computational Glasses. This problem is extended in our prototypes as we integrate world cameras and so we also need to solve for a camera-display relationship.

Distortion Correction. An often overlooked issue with OSTHMD is the distortion caused by imperfections in the displays and optics used. Accounting for these view-dependent distortions are important for achieving indistinguishable *augmented reality (AR)*, however is a challenge as the geometric distortions will vary with the location of the eye, requiring corrections at all possible viewing locations. This is coupled with the issue that most models used when displaying content consider the displayed image to be on a plane, however it is in fact projected onto a curved surface (Klemm et al., 2016; Owen et al., 2004). To solve this issue the use of user-perspective cameras to create parametric models and per-pixel mappings has been proposed (Klemm et al., 2016; Langlotz et al., 2016). Computed display parameters from sample points collected using a user-perspective camera can be used to inform further calibrations based on the position of a user's eyes (Owen et al., 2004). Itoh and Klinker (2015a); Itoh et al. (2016a) investigated the measurement and correction of these distortions by estimating a *point spread function (PSF)* from sample points on the display, highlighting the view dependence of the strong chromatic aberrations and blurring.

Distortion correction remains an issue for OSTHMD and therefore is an issue with the prototypes we develop. This is not addressed throughout our user studies where users are able to look directly through the prototypes. The eye position of the users varies and therefore the view dependent distortions will also vary. This distortion remains a limitation on the calibrations performed by participants.

3.1.2 Temporal Coherence

When moving an HMD around the virtual content displayed to the user must update to continue to correctly cohere to the real world. This introduces problems with updating the image displayed at high enough rates and with the latency introduced by the system itself and the display.

System Latency. Much like finding the relationship between OSTHMDs and the real world is common to all HMD, system latency is a common issue amongst all forms of HMD. Whilst OSTHMDs provide a more stringent set of latency requirements, initial steps in reducing latency can be found in *virtual reality (VR)* HMDs where it is often researched due to its contribution to motion sickness (Cobb et al., 1999). The overall latency of a system can be broken down into four components: the time between changes in physical pose and its measurement, known as the tracking delay; the time for the system to execute processing, or application delay; the time to generate the next image to be displayed, or image generation delay; and finally, the aforementioned time to display the image on the display, or display delay (Mine, 1993).

In AR latency leads to errors in the positioning of virtual content relative to the real world under motion. These can be seen by a user if the latency is too high, causing incoherence (Lincoln et al., 2016a; Livingston and Ai, 2008; Azuma, 1997). One method to reduce the latency of OSTHMD is to utilise low latency tracking methods developed for VR as demonstrated by Itoh et al. (2016b) who used an Oculus Rift's low latency tracking for an OSTHMD system.

Alternatively, to reduce the impact of system latency programmatic methods in both OSTHMDs (Buker et al., 2012) and general HMDs that can be applied to OSTHMD (Carmack, 2019) have been demonstrated by adjusting visual output based on the system latency and trying to account for errors.

For our Computational Glasses we do not have tracking delays, however we do have delays from integrated world cameras that introduce latency which needs to be considered. Furthermore, as the glasses need to be able to constantly modulate what the user is seeing as they are seeing it, minimising processing delay and production of images to display are important considerations.

Display Latency. Whilst a subset of the overall system latency, a notable part of the latency which has significant impact and seen dedicated research to solving is the display latency. To reduce this latency *digital micromirror devices (DMDs)* have been shown to allow an update rate of 50 μs when using binary states (Zheng et al., 2014). This prototype was however constrained to grey-scale and precomputed data whilst introducing a noticeable flicker to the user's view. This has subsequently been updated with reduced tracking latency using a mechanical tracker and no flicker (Lincoln et al., 2016b) and *high dynamic range (HDR)* (Lincoln et al., 2017).

The display latency for our prototypes is generally an issue that remains to be addressed. Most of the OSTHMDs we used for our research introduce significant display latency, particularly our use of the Epson Moverio Bt-300 which required displaying modulations via an Android device.

3.1.3 Visual Coherence

The final component needed for virtual objects to be coherently rendered in the real world is the need for the visual properties of the content to be coherent with the real world. Beyond the issues generic to all forms of AR, coherent lighting (Gao et al., 2013) and realistic rendering (Klein and Murray, 2008), there are issues with colour reproduction, dynamic range, creating occlusions, depth accommodation, FoV, and resolution that all require unique solutions for OSTHMDs.

Colour Reproduction. For a virtual object to be coherent with a physical counterpart the colour depth of both objects must be comparable to the point of indistinguishability. In VSTHMD this can be achieved in image space (Klein and Murray, 2008), whilst in *spatial augmented reality (SAR)* this introduces further challenges and requires radiometric calibration (Bimber et al., 2005a,b, 2008; Bimber and Raskar, 2005). The challenges in SAR are further compounded in OSTHMD as the displays are generally non-uniform which is common to all displays but amplified by the optics of the displays in OSTHMDs.

Display colour reproduction. The colour perceived by users of OSTHMDs is not equal to that displayed, due to inaccuracies in colour reproduction of displays. Even without considering the background that is being seen through the OSTHMD, virtual content will have incoherent colours to the real world due to light absorption from optical elements and lossy colour conversions (Itoh et al., 2015). Assuming a uniform background, research has shown the use of *look-up tables (LUTs)* (Sridharan et al., 2013) or using a semi-parametric approach to compute a more coherent display profile (Itoh et al., 2015).

Colour blending. Regardless of the optical design, all OSTHMD blend both the real world and virtual content together. In doing so the virtual object being displayed is impacted by the real world it is blending with. Depending on the real world being blended the resulting view will differ from that of the original display of the object (after colour reproduction considerations). This is compounded by the fact that all commonly implemented OSTHMD designs are unable to actively darken the environment, relying on tinted shades if doing anything at all. Global variance under environmental conditions is the most advanced application of this (Mori et al., 2018). The impact of this was first explored by (Gabbard et al., 2000) who analysed the problem aand suggested potential issues. Subsequently several works have looked to address this issue.

Under purely simulated conditions and building on prior works ((Fukiage et al., 2014; Sridharan et al., 2013)) Smart Colour presented a theoretical solution to this issue, however was still limited by the practicalities of physical implementations (David Hincapie-Ramos et al., 2014). First to present a practical approach, although requiring a static and calibrated system, (Weiland et al., 2009) computed an image to display that reduced the effect of colour blending by utilising a camera image of the scene. (Langlotz et al., 2016) provided further work in the direction of utilising cameras to compute mitigated images that neutralise the impact of the real world. They considered display and camera characteristics and their work can be applied to non-static environments.

Rather than compensate for present environment light a more desirable approach would be to filter it as needed. The use of either *liquid-crystal displays (LCDs)* (Wetzstein et al., 2010) or *phased spatial-light modulators (PSLMs)* (Itoh et al., 2019a; Kaminokado et al., 2020) have been demonstrated as ways to implement displays that can filter environment light in a subtractive manner. However, such approaches remain largely bench bound and would require the integration with more traditional OSTHMD or another means to add light to enable full control. To the best of our knowledge, this has yet to have been effectively demonstrated.

High dynamic range. To attempt to match the large dynamic range of the HVS methods such as multiple modular planes are popular (Seetzen et al., 2004), however are not directly translatable for OSTHMD due to exasperated issues with calibration and alignment inaccuracies (Xu and Hua, 2017; Zhao et al., 2020). Notably, the issue of dynamic range has been rarely approached in OSTHMDs. (Lincoln et al., 2017) utilised a DMD to extend dynamic range, however their system was relatively large and impractical as a worn OSTHMD. Alternatively, Itoh et al. (2018) utilised a high-contrast projector in conjunction with optical elements and filters to create a retinal projection display. However, whilst this hearkens to retinal projection, OSTHMDs their implementation was not *optical see-through (OST)*.

As our work involves direct modulation of the environment colour reproduction becomes an important consideration and has some unique considerations in its application. The display colour reproduction of displays needs to be considered as the actual colour displayed will have an impact on how any modulation is perceived and, subsequently, the effect it has on the world. Colour blending also needs considered, less in regard to how to produce the correct colour for virtual objects in spite of the real world, but in regard to what the final appearance of the real world will be. The dynamic range of the displays also has a different impact on Computational Glasses than in the standard use of OSTHMDs. The dynamic range of Computational Glasses has less affect on the ability for the glasses to be used and integrate with the real world, instead limiting the ways in which modulations can impact the world. The higher the dynamic range, the greater the range of ways that the world can be modulated.

Occlusion Capability. Occlusions provide crucial information about the 3D relationship between objects and their relative depth. Due to the nature of typical OSTHMDs designs blending light with the real world, which is at odds with occlusion where light is blocked from objects based on their depth, this is a particular problem for OSTHMDs. Table 3.2 covers the major works in this area. Being able to achieve occlusions would also enhance colour reproduction by reducing the image of colour blending, via occlusion of the real world.

Quintessential to existing solutions to occlusion in OSTHMDs is SLMs such as using and LCD to change the transparency to light rays passing through each pixel. Early examples integrated such an LCD into the optical path (Kiyokawa et al., 2000,

Reference	Occlusion Generation	Adaptive Depth	Varifocal Image	Mask Appearance	Computational Demand	Optical Complexity	Form Factor	Field of View	Portable Prototype
(Kiyokawa et al., 2000)	LCD	no	no	sharp	low	complex	bulky	25°	yes
(Kiyokawa et al., 2003)	LCD	no	no	sharp	low	complex	bulky	30°	yes
(Wilson and Hua, 2017)	LCoS	no	no	sharp	low	complex	thin	30.58°	no
(Gao et al., 2012, 2013)	LCoS	no	no	sharp	low	complex	thin	40°	no
](Cakmakci and Rolland, 2006)	LCD	no	no	sharp	low	moderate	thin	40°	no
(Santos et al., 2008)	LCD	no	no	dull	low	N/A	thin	20°	yes
(Maimone and Fuchs, 2013)	LCD	no	yes	dull	high	moderate	thin	65°	yes
(Maimone et al., 2014)	LCD	yes (in-focus)	yes	dull	moderate	simple	moderate	80°	yes
(Itoh et al., 2017)	LCD	partially	no	sharp	low	simple	thin	70-80°	no
(Yamaguchi and Takaki, 2016)	LCD with Microlenses	yes	yes	sharp	high	moderate	thin	moderate	no
(Uchida et al., 2002)	DMD	no	no	sharp	low	complex	moderate	narrow	no
(Kim et al., 2019b)	DMD	no	no	sharp	low	complex	moderate	narrow	no
(Krajancich et al., 2020)	DMD	yes	no	sharp	high	complex	moderate	8.7°	no
(Hamasaki and Itoh, 2019)	LCD on linear stage	yes	no	sharp	low	moderate	moderate	narrow	no
(Rathinavel et al., 2019)	LCD with refocusable lens	yes	yes	sharp	moderate	complex	moderate	15.3°	no
(Zhang et al., 2021a)	LCD	no	no	sharp	low	complex	bulky	wide	no
(Zhang et al., 2021b)	LCD	no	no	sharp	low	complex	bulky	wide	no

Table 3.2: Reported properties of different occlusion-capable HMDs adapted from (Hamasaki and Itoh, 2019). Designs of moderate optical complexity use 3–4 optical elements (including the SLM), and a moderate FoV is 40–80°.

2003). Whilst the approaches produced a sharp occlusion, their size produced a bulky display, which facilitated further research into methods of achieving occlusion.

Further works utilised near-eye LCDs which introduce an issue of fixed focus occlusion masks, something addressed by (Itoh et al., 2017).

Further approaches utilised integrated LCDs with light-field displays, either single layer (Maimone et al., 2014), or multi-layer (Yamaguchi and Takaki, 2016; Maimone and Fuchs, 2013). Others utilised DMDs as reflective SLMs (Uchida et al., 2002; Kim et al., 2019c; Krajancich et al., 2020) or LCOS (Cakmakci et al., 2004; Wilson and Hua, 2017; Gao et al., 2012, 2013). Further recent works have looked at creating wide FoV displays which enable occlusion (Zhang et al., 2021a,b).

As a final approach, Santos et al. (2008) proposed a binocular approach using an undisclosed optical design.

Varifocal Occlusion. As the aforementioned occlusion methods suffer from the *accommodation–vergence conflict* issue common to OSTHMDs, some research has also looked to overcome this issue with varifocal occlusion (Hamasaki and Itoh, 2019; Rathinavel et al., 2019).

Much like accommodation capabilities, the occlusion capabilities of our Computational Glasses are that of the OSTHMDs they are built on. Occlusion capabilities also impacts the capabilities of Computational Glasses in a similar manner to accommodation, restricting the range of modulations capable of being produced.

Accommodation Capability. The accommodation is an important cue in the natural ability of the HVS to perceive depth, and this extends to viewing virtual content (Cutting, 1997). However, most conventional OSTHMD do not allow for this as they have only a single focal plane. Although not technically a plane as discussed, the focus distance is fixed by the optics, generally at 3-7m. Not only does this hinder the ability to utilise these cues, it actually introduces issues when viewing virtual content that has been placed at a distance different to that of the focal plane based

Reference	Mechanism	Classification	See-Through	Field of View	Resolution	Eye-tracking Needed	Optical Complexity	Form Factor	Computational Demand	Miniature Prototype
(Lanman and Luebke, 2013)	microlenses	light-field	no	33.3°	low	yes	simple	thin	moderate	yes
(Yamaguchi and Takaki, 2016)	microlenses	light-field	yes	4.3°	low	yes	simple	thin	moderate	no
(Otao et al., 2017)	microlenses	light-field	no	narrow	low	yes	simple	bulky	moderate	yes
(Maimone et al., 2014)	pinlight display	light-field	yes	110°	low	recommended	simple	thin	moderate	yes
(Akşit et al., 2015)	pinhole display	light-field	no	83°	low	recommended	simple	thin	moderate	yes
(Song et al., 2014)	pinhole display	light-field	yes	narrow	low	recommended	simple	thin	moderate	yes
(Jang et al., 2017)	laser w/ steering mirror	holographic	yes	68°	moderate	yes	complex	thin	high	yes
(Maimone et al., 2017)	SLM	holographic	yes	80° hori.	high	yes	complex	thin	moderate	yes
(Shi et al., 2017)	SLM	holographic	no	N/A	high	yes	complex	moderate	high	no
(Moon et al., 2014)	SLM	holographic	yes	narrow	N/A	yes	complex	moderate	high	yes
(Gao et al., 2016)	amplitude SLM	holographic	yes	1.7°	high	no	complex	bulky	low	no
(Gao et al., 2017)	SLM	holographic	yes	5.4°	high	no	complex	moderate	low	yes
(Huang et al., 2015)	stacked LCD panels	multi-plane	no	125°	high	no	simple	bulky	high	yes
(Lee et al., 2016)	savart plate	multi-plane	yes	37°	high	no	moderate	moderate	low	yes
(Yoo et al., 2019)	polarised lenses	multi-plane	yes	7.5° hori.	high	no	moderate	moderate	low	no
(Liu et al., 2016)	polymer stabilised scattering shutters	multi-plane	yes	narrow	N/A	no	simple	moderate	low	no
(Maimone and Fuchs, 2013)	stacked LCD panels	multi-plane	yes	65°	low	no	simple	moderate	high	yes
(Rolland et al., 2000)	stacked planar displays	multi-plane	yes	N/A	high	no	simple	moderate	low	no
(Liu et al., 2008)	focus tunable lens	varifocal	yes	28°	moderate	no	complex	moderate	low	no
(Rathinavel et al., 2018)	focus tunable lens	varifocal	yes	15.3°	moderate	no	complex	bulky	low	no
(Xia et al., 2019)	focus tunable lens	varifocal	yes	37°	moderate	no	complex	bulky	low	no
(Dunn et al., 2018, 2017)	focus tunable membrane	varifocal	yes	103°	moderate	yes	simple	bulky	moderate	yes
(Wilson and Hua, 2019)	alvarez lens	varifocal	yes	103°	moderate	no	complex	moderate	moderate	no
(Matsuda et al., 2017)	SLM	focal surface	no	18°	high	yes	complex	bulky	high	yes

Table 3.3: Properties of different display designs taken from the major works that present virtual content at different focal distances. Moderate FoV equals 40–80° and moderate optical complexity requires 3–4 optical elements.

on vergence. This is known as the accommodation–vergence conflict and affects 3D perception, as well as being a cause of VR sickness (Dunn, 2019; Hoffman et al., 2008; Kramida, 2016; Kress and Starner, 2013). Solving this issue requires additional optics or alternative display technologies (Hua, 2017) which can be categorised into four categories: adjustable-focus OSTHMDs; multifocal or varifocal OSTHMDs; OSTHMDs utilising light fields or holographic displays; and in-focus OSTHMD. Table 3.3 covers the major works in this area.

Before covering the alternative designs to allow for accommodation, it is important to note that they come with a need for rendering algorithms that can be computed real time to account for the needs of each design. For example image areas to display on each plane in multifocal displays (Narain et al., 2015), or phase images in holographic displays (Peng et al., 2020). Xiao et al. (2018) covered an overview of this for accommodation-supporting HMDs independent of hardware.

Adjustable-focus OSTHMDs. Potentially the simplest approach conceptually is to adjust the focal plane, either manually or automatically. Whilst not fully mitigating the issue, it would allow context or convergence based adjustment as demonstrated by Willson and Hua using free-form Alvarez lenses (Wilson and Hua, 2017).

Multifocal or Varifocal OSTHMDs. Another approach to providing accommodation is to provide further focal planes either by having multiple concurrent planes (multifocal (Rolland et al., 2000)) or by shifting between focal planes in a time-multiplexed manner (varifocal). Displaying content on the various planes helps reduce the accommodation–vergence conflict, however, still requires set focal planes to which the correct focal point must be approximated, as well as the potential for perceivable flicker commonly found in multifocal displays.

A simple method to achieve this is utilising *half-silvered mirrors*, larger versions of those used in the simplest OSTHMD designs and stacked to create various focal planes was demonstrated by Akeley et al. (2004). However, whilst similar in principle to OSTHMDs their prototype was a large desktop display. In OSTHMDs prototypes several examples have been shown of dual layer multifocal displays. These have been achieved using a geometric phase holographic lens as a waveguide (Lee et al., 2016) and a fast polarisation rotator and Savart plate in conjunction with a micro-display (Lee et al., 2019). The commercial Magic Leap One also demonstrated dual focal planes which are switched based on eye-tracking. [3].

A greater number of planes have been shown using in a bench prototype where layers of liquid crystal that switch scattering properties in a time sequence have an image projected onto them (Liu et al., 2016). An alternative initially designed for VR displays, focal surface displays, utilise phase-only SLMs, displaying multiple focal planes as a continuous focal surface (Matsuda et al., 2017).

Looking at varifocal displays, solutions can again be found in desktop-based solutions. Suyama et al. (2000) used liquid crystal varifocal lenses on a display to give the illusion of different focal planes. Various other methods for creating varifocal displays have also been considered. Liu et al. (2008) created a monoscopic OSTHMD using liquid lenses to vary focus between two planes. Another approach using a varifocal display by Xia et al. (2019) has demonstrated creating a switchable VR, OSTHMD using a focus-tunable lens with a shutter glass. Further applications of focus tuneable lenses have been to create a volumetric OSTHMD to render 280 perceptually simultaneous depth planes (Rathinavel et al., 2018), and to cancel the accommodation effect (Konrad et al., 2017).

A final approach to creating a varifocal display is to use a deformable membrane mirror as the beam combiner as shown which enables a large FoV and accommodation capability (Dunn et al., 2018, 2017).

Light-field and Holographic OSTHMDs. Although using completely different ways of producing images (Lee et al., 2017), light-field displays (Jang et al., 2017; Lanman and Luebke, 2013; Maimone et al., 2014; Masia et al., 2013; Otao et al., 2017) and holographic displays (Cem et al., 2020; Gao et al., 2016, 2017; He et al., 2019; Moon et al., 2014; Shi et al., 2017), allow users to focus correctly at any depth at any time by, theoretically, offering all depth cues simultaneously.

To create light-field displays, where light perceived from virtual objects is typically represented as 4D light rays, two major approaches exist: microlens arrays (Song et al., 2014; Yamaguchi and Takaki, 2016), and stacked transmissive LCDs (Huang et al., 2015; Maimone and Fuchs, 2013). Outside these Jang et al. (2017) proposed a light-field projection display that utilised multiple laser light sources and an eye-tracking camera.

To create a holographic display the amplitude or phase of a collimated laser in controlled by an SLM so that the wavefront of propagated light forms the desired image at the view point (Yaraş et al., 2010, 2011). This has been demonstrated in an OSTHMD by (Maimone et al., 2017). However, to produce high-quality images in real-time in this manner is computationally expensive ((Chakravarthula et al., 2019; Padmanaban et al., 2019b)) so Peng et al. (2020) proposed the use of a neural net-

[3] www.magicleap.com

work method. Other improvements to the FoV (Cem et al., 2020) and étendue (Kuo et al., 2020) of holographics displays have been shown.

In-focus OSTHMDs. The final method for alleviating the accommodation–vergence conflict is to ensure that the display is always in focus, regardless of accommodation. Whilst this can be achieved by a pinlight display using near-eye pinlight projectors (Furness III and Kollin, 1995; Do et al., 2019), more typical approaches use retinal-scanning displays and virtual retinal displays (Maimone et al., 2014; Akşit et al., 2015).

Because we built our prototypes around the OSTHMDs commercially available to us, and with consideration for modifications and colour reproduction, our prototypes are constrained to single planes of focus and have no accommodation capabilities. An added component to user accommodation with our Computational Glasses is that the world cameras also have a singular focus, which is either set, as in our case, or needs adjusted based on the user's accommodation to enable the production of modulations based on the correct accommodation.

Wide Field of View. A more common issue with OSTHMDs and one commonly discussed, to the point of being used to classify them (Cakmakci and Rolland, 2006), is the limited FoV. The human FoV extends to, typically, 180^o-200^0 which far exceeds that of standard OSTHMDs (Cakmakci and Rolland, 2006; Maimone et al., 2014), as such, even minor eye rotations can cause the FoV to be exceeded, causing virtual objects to disappear and breaking coherence.

Conventional designs such as using half-silvered mirrors or prisms allow for large FoV however, include the trade-off of equally large size (Sutherland, 1968). Utilising curved mirrors can however reduce this form factor (Kiyokawa, 2007; Nagahara et al., 2006). Waveguide based solutions also suffer when attempting to create a wide FoV requiring large form factors and causing ghost images and colour bleeding (Mukawa et al., 2008). An alternative design has been proposed in the form of membrane optics (Dunn et al., 2017), however whilst allowing for accommodation capability, are still bulky. Further optically-complex and bulky designs have been used for wide FoV displays capable of enabling occlusions (Zhang et al., 2021a,b).

More recently, free-from optics (Cheng et al., 2011) and meta-surfaces (Lee et al., 2018) have become viable. These allow for expanded of FoV, being demonstrated up to 70-76^o. Another solution is to use a point light array placed inside the typical accommodation range of the eye. This defocused array forms pinlights which the light from can be modulated by a transmissive SLM between the eye and array (Maimone et al., 2014). This has also seen commercial application where instead of using a point light array a contact lens worn by the user that serves the same purpose, and a defocused image is projected from a micro display (Sprague, 2010).

Taking different approaches, (Benko et al., 2015) demonstrated the use of SAR projectors to infill the peripheral vision of the user outside of the OSTHMD, whilst Maimone and Wang (2020) demonstrated a VRHMD that could achieve 90^o and was later extended by Lee et al. (2020) to create an OSTHMD prototype. This design employed a polarisation-based optical folding technique.

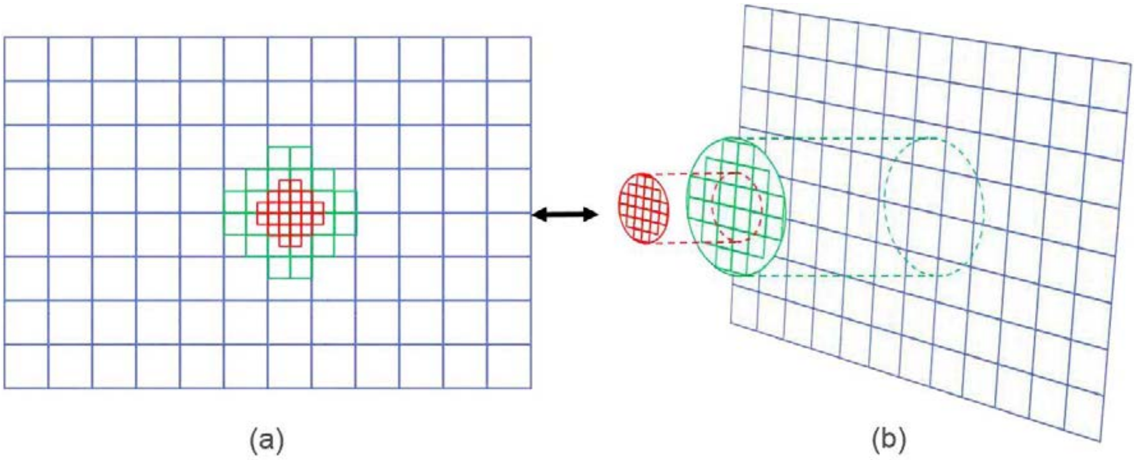


Figure 3.3: Example of foveation of a pixel grid inspired by (Guenter et al., 2012). (a) Grid foveated around a point with the highest detail/acuity in red, reduced in green and low acuity in blue. (b) central "fovea" area with the highest level of detail in red, reducing as eccentricity increases in green and low detail in the blue periphery.

As we look to produce a new means of glasses akin to traditional ones, the resolution and field of view of our devices is of particular importance. Whilst modulations can be achieved using a limited central area, ideal Computational Glasses would act more like traditional glasses. In our research we were only able to use limited FoV devices and as such our modulations were limited to small areas of the central vision of participants. For Computational Glasses to be applied in practice a wider FoV is needed.

Resolution. The requirements to meet the potential resolution of the HVS across its visual field far exceed the 1,920 × 1,080/1,200 for research prototypes (Hua and Javidi, 2014; Itoh et al., 2019a). Whilst the resolution of displays is increasing the problem to process the required information is still pressing. Several techniques to reduce the resolution can be utilised (Spjut et al., 2020). As the resolution of the eye is only high at the fovea, this is the only area that need rendered at high resolution, the majority of the field of view is only perceived in the periphery and can be rendered at low resolution (Figure 3.3). This idea is known as foveated rendering and has been well covered for general displays (Guenter et al., 2012) and HMDs (Patney et al., 2016) including early works on near-eye displays (Howlett, 1992; Rolland et al., 1998). Recent works have demonstrated the application of foveated rendering in OSTHMDs. This has been achieved using two displays, one for the fovea and one for the periphery (Howlett, 1992; Rolland et al., 1998); multi-layer displays (Lee et al., 2018); and *holographic optical elements (HOEs)* (Kim et al., 2019a; Lee et al., 2019). An alternative approach was shown by Akşit et al. (2019) who created foveation via hardware.

The resolution of OSTHMD has a large conceptual impact on Computational Glasses built on them. The idea of modulating the light as seen by the users is constrained to the pixel resolution of the display and variances within the size of a pixel cannot be represented. Smaller details can be lost. Higher resolution will however come with the added computational effort required to produce modulations. Furthermore, whilst modulations can be applied at pixel resolution, determining modulations is restricted to that of the integrated cameras.

3.1.4 Impact

Generally, we can see that whilst there is a large body of work exploring accommodation issues, other issues such as latency, colour reproduction, and resolution have been less explored. Furthermore, these explorations have been largely limited to bench top prototypes and addressing only one issue, with more complete solutions or wearable prototypes being minimal, with the work of (Kim et al., 2019c) standing out as an exception.

When looking to modulate the environment of users the ideal display would do so with complete coherence to the real world. As many of the prototypes presented to help address the issues of coherence are constrained to bench prototypes and require specialist hardware and setup, we choose to remain with commercially available OSTHMDs that we could readily modify for our purposes. We accepted the short comings of these devices and leave the solutions to these issues to future research and other work that we can clearly see is being done. This does however leave us with several of the aforementioned incoherencies to deal with.

Throughout this research we relied on manual methods to calibrate the OSTHMD to the user's view of the screen for use in user studies. Initially we utilised a camera track-able marker on the screen to complete both user eye-display and display-world calibrations simultaneously. To improve the calibration during later user studies with a static head and display position, we adjusted our calibration to present the system output to the user and allowed them to adjust the calibration until the output aligned with their view.

The latency in Computational Glasses needs to be minimal to allow for seamless, continuous modulation. We developed all our systems to run on GPUs to allow for fast processing with the frame rate of the world cameras often being the constraining factor on the processing. However, there is still notable display latency, and for the purposes of our research we looked to use static prototypes and images to mitigate its impact on our results.

Because we cannot fully correct for the colours of our displays, we instead look to mitigate the largest of the impacts by utilising OSTHMDs that cover a large colour gamut, have a relatively good representation of white, and are relatively bright. For example, we used the Epson Moverio series of devices that use OLEDs. We also compensate for a main effect caused by integrated optics through correcting for uniform colour shifts.

When considering the accommodation–vergence conflict we always look to place real world content viewed through the device as close to the virtual image plane as possible. In conjunction with this, we generally looked to use minimal modulation where the user will naturally accommodate to the real world and view the modulations as slightly blurred, rather than accommodating to a different depth than their convergence.

As occlusion is not available on commercial OSTHMDs we worked around this by creating modulations and techniques that only added light to the user's view. This was a key component in our research.

The other issue we had to consider was the limited FoV of the devices. To this end we simply ensure areas of the real world that needed modulation remained within the available FoV. Similarly, the resolution of the displays was limited, which impacts the detail with which we can modulate the world. However, we looked to ensure that our modulations did not rely on users reading fine details, although there were some exceptions when modulating natural scenes and exploring application of our prototypes is more generalised scenarios.

3.2 Implementing Computational Glasses

Whilst not currently able to achieve complete coherence between the real world and virtual content, OSTHMD are developing to be able to better do so. Given the research in this area we believe in the viability of such devices to be used in conjunction with the AR techniques to provide a new means to assist the HVS. In fact, for some application scenarios we believe they are already capable of demonstrating efficacy, so long as we take the issues mentioned into consideration.

In our work we look to use Computational Glasses by directly modulating the light the user's perceive using OSTHMDs. Using these we seamlessly align computer generated overlays with the world seen by the user, thus directly modulating the light and colours they perceive. To this end a precise understanding of the user's view of the world to determine how modulations are required and need to be displayed. Whilst commercial displays for AR, such as *video see-through (VST)* and OSTHMDs have the potential to modulate the light a user perceives and integrate computer generated overlays they are unable to function as Computational Glasses the way they are.

VSTHMDs allow for the necessary precision to seamlessly align graphics with pixel-precision, however they limit the wearer's view of the world to the responses of the cameras used, from the position of cameras, and by the displays used to re-project the world to the user. OSTHMDs avoid these issues, however, are unable to achieve the required pixel precision due to the position of any world cameras not matching the user's view of the world. Even using high precision tracking would not solve these issues as the glasses need to be able to modulate unknown physical environments, something fundamentally different to traditional OSTHMDs.

In keeping with the development of traditional OSTHMD in desiring a form factor and weight to traditional glasses but allowing for integrated computation and battery (either inbuilt or tethered), Computational Glasses differ in their foundational premise of looking to provide new or improved forms of vision assistance. Where OSTHMDs and AR look to integrate 3D virtual content into the real world,

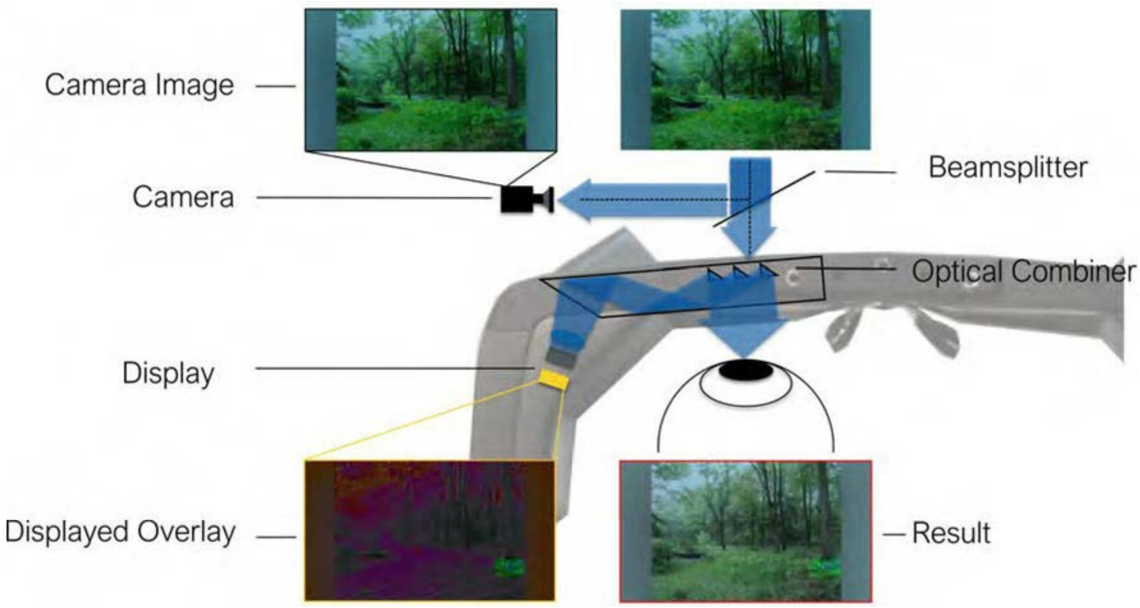


Figure 3.4: Our implementation of Computational Glasses. Light from the scene is redirected to a camera via a beam splitter. This camera image is then analysed and an overlay is produced to be displayed on the OSTHMD. The overlay integrates with the light from the scene being perceived by the user and modulates it.

Computational Glasses look to directly modulate and compensate the user's view of the real world, changing its appearance. However, in keeping with the similarities optical elements and displays used to present virtual images in OSTHMDs and AR can be used for the purpose of modulating vision, and as such is what we build our prototypes on.

The biggest challenge to using OSTHMDs for Computational Glasses is the pixel-precise modulation of the environment. Doing this has three main challenges. Firstly, being able to capture the environment as seen by the user. Secondly, being able to process the captured environment as seen by the user to identify areas of interest. Finally, being able to modulate these areas via the semi-transparent head-mounted display. For our work we took inspiration from earlier work in the field of traditional AR that tried to improve colour rendering in semi-transparent displays by considering the background colour on a per pixel basis (Langlotz et al., 2016). The key is to virtually place a camera at the position of the user's eye. This can be achieved by adding a beamsplitter to off-the-shelf OSTHMDs for each eye. These beamsplitters redirect a portion of the light travelling towards the user's eye to a camera (the scene camera).

Since the cameras can only be moved close to the ideal position and human eyes differ, we also apply a software calibration similar to the Single Point Active Alignment Method (SPAAM) (Tuceryan et al., 2002) to account for potential minor misalignment. We cannot take the traditional SPAAM used for eye-display calibration but have to extend it by incorporating the added scene camera. The details of the calibration differ between the prototypes we developed. However, for all prototypes the resulting eye-display-camera calibration creates transformations allowing for a mapping from camera to display via the eye, consequently allowing a precise modulation from the user's point of view. Thus, for each pixel within the scene camera imagery (representing a part of the physical environment), we can compute the corresponding pixel within the semi-transparent display of the OSTHMD, so that they both align with each other from the perspective of the user's eye. See Figure 3.4 for an overview of our implementation of Computational Glasses.

Whilst the focus of this research is on creating Computational Glasses using commercially available OSTHMDs, as discussed, it is by no means the only method to realise Computational Glasses. For example, parallel to and in conjunction with our research on using OSTHMDs, that utilised near-eye optics based on phase modulation (Itoh et al., 2019a,b, 2021). Whilst in early stage and using optical bench prototypes their work acts as a programmable colour filter with pixel-wise colour control. Other works that could also fall under the category of Computational Glasses implemented by means other than OSTHMDs include the works such as AutoFocals (Padmanaban et al., 2019a) and FocusAR (Chakravarthula et al., 2018).

Chapter 4

Vision Aid Use Case: Colour Vision Deficiency

In this chapter the use of Computational Glasses as a visual aid for colour vision deficiency (CVD) is covered. We present an overview of the visual impairment itself to aid understanding before covering in detail the prior techniques that have been developed to aid CVD computationally. These techniques serve as our basis for developing our Computational Glasses solution, however due to some limitations in the coverage of the techniques, which we cover, we look to test a variety of them. We present a series of studies conducted to test the efficacy of our concept, compare it against alternative approaches, look to explore use in more generalised scenarios, and future directions for creating compensation techniques for CVD in Computational Glasses.
The contributions of this chapter are our implementation of Computational Glasses as a means to aid CVD and our subsequent studies into its efficacy and future directions.
This chapter draws on papers Langlotz et al. (2018) and Sutton et al. (2022a).

Figure 4.1: Examples demonstrating the results from using Computational Glasses to create a spectral wave-length shift for CVD. Top: Pictures presented in front of a user wearing our prototype. Bottom: The pictures when viewed using a compensation technique, insets demonstrate what is displayed on the glasses. Note that these are exaggerated compensations for paper clarity. Right: The large, stable stereoscopic prototype used for controlled user studies. All images were captured through this system and are not simulated. We also demonstrate the development of smaller portable versions.

As our first foray into creating Computational Glasses we look to demonstrate the potential for them to aid visual impairments that are current unaided by looking to compensate for the effects of *colour vision deficiency (CVD)*. CVD is a condition in which those afflicted perceive colours in a manner that is considered abnormal. This condition affects an estimated several hundred million people worldwide (Childs, 1989) and is most commonly seen in the form of a genetic condition under which red or green wavelengths are perceived as being closer to the other than they are under normal vision. However, there are also other more uncommon causes and effects.

Techniques to address CVD through computational means have been explored for over 20 years (Meyer and Greenberg, 1988; Brettel et al., 1997; Viénot et al., 1999) and usually focus on optimising user interfaces (e.g., web pages or navigation interfaces). Most user interface techniques shift the representation of the confused colours by changing either their hue or brightness thus improving the user's ability to distinguish between them. However, only a few works considered applications in everyday activities, such as deploying the system onto a mobile phone (Ananto et al., 2011; Schmitt et al., 2012) or an *optical see-through head-mounted display (OSTHMD)* (Tanuwidjaja et al., 2014) and, prior to our research, none have allowed for direct modulation of the user's perception of the world the colours they perceive.

Problem selection. We choose to investigate CVD as our first exploration of Computational Glasses as it represents a well understood form of visual impairments that effects a large portion of the population and yet has no accepted aid or means of treatment. CVD is also common across all age demographics in our population, making recruitment of participants for user studies viable. Being well understood and common, there is also a large amount of prior works from which we can look to draw upon. We could investigate the related literature to aid in development of compensation techniques, then test the efficacy of our Computational Glasses as a potential solution, before opening up our research and looking to explore future research directions.

In this chapter we cover our work verifying the feasibility of Computational Glasses for CVD mitigation in a series of increasingly realistic studies, ranging from an initial technical validation of a monocular prototype to a more realistic stereoscopic system, to mobile system. Our tests showed our concept was feasible and indicated that a colour shift might provide good discernability when tested on the standard Ishihara test for colourblindness. A comparison with state-of-the-art CVD compensation on Google Glasses showed that our system performed better for small details without increasing a participant's workload. A further study using a mobile prototype explored acceptance of devices and use in more realistic scenarios. In a final study, various techniques are compared to investigate future research directions. Example output from our Computational Glasses for CVD can be seen in Figure 4.1.

4.1 Colour Vision Deficiency

A common issue affecting the *human visual system (HVS)* is CVD. In 1798 Dalton reported on the observed differences in his own colour perception compared to that of others, and found that this was not unique to himself, before speculating as to the nature of this anomalous colour vision (Dalton, 1798). Although his explanations were not complete, subsequent research revealed how our eyes process incoming light and the causes of CVD. It is now known that the human eye has a number of photo receptors on the retina that are sensitive to different wavelengths of incoming light (*cones*) and light intensity (*rods*). As discussed, normal colour vision, also known as *trichromacy*, is the result of the different response functions from cones to incoming light waves. Anomalous cone responses result in various forms of CVD. The mildest form is *anomalous trichromacy* where the eye retains all cone types however the colour response of at least one of them is reduced. Depending on the affected type of cones this condition is referred to as *protanomaly, deuteranomaly, or tritanomaly.* In more extreme forms, the response from one cone type may be missing entirely, resulting in colour vision consisting of variations of only two colours, also known as *dichromacy. Protanopia, deuteranopia, and tritanopia* are all forms of this. Note that whilst dichromacy is considered abnormal in humans due to a lack of cones, it is a common state among some animals, for example dogs and cats (Jacobs, 2009). The commonly used term of "red-green blind" or "red-green weakness" is an overarching term used for: protanopia; protanomaly; deuteranopia; and deuteranopia. This is due to all people within these groups exhibiting similar colour differentiation problems[1]. The rarest forms of CVD lack the response from two, or even all three, cone types resulting in lack of colour vision, or *monochromacy.*

[1]https://www.color-blindness.com/

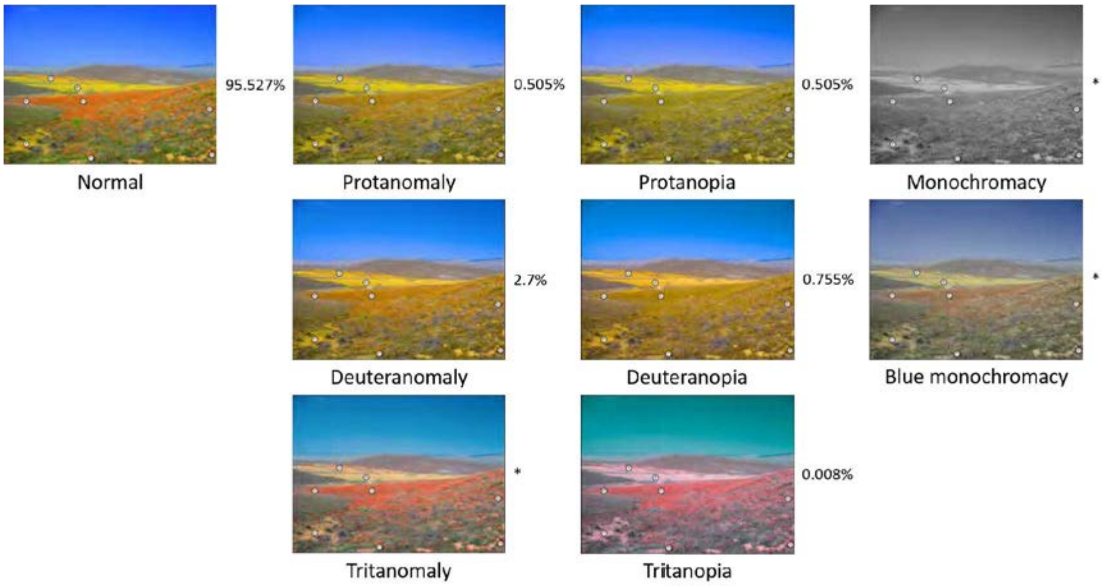


Figure 4.2: Types of CVD and estimated world percentages. Values estimated based on The Perception of Colour (Kalloniatis and Luu, 2007). * Percentages for tritanomaly, and forms of monochromacy are not available due to their scarcity.

See Figure 4.2 for example simulations of the colour confusions seen by those with CVD and their estimated occurrences within the population.

Studies have shown that CVD affects an estimated 4% of the population worldwide (Spalding, 1999). The distribution of those affected varies by gender, with much stronger prevalence among the male population. Some 7.4% of men are estimated to be afflicted, compared to only 0.5% of the female population in northern Europe (Sharpe et al., 1999). The bias towards males is due to deuteranopia and protanopia being caused by a recessive X linked gene making its occurrence in women exceedingly unlikely (Deeb, 2005). Whilst deuteranopia and protanopia (and the associated ~nomalies) are X-linked, tritanopia and monochromacy are not, so do not show the same gender bias, being relatively uncommon across the board. Furthermore, prevalence varies with ethnicity, with Singapore reporting rates closer to 5% of boys and .2% of girls (Chia et al., 2008) and some works reporting values between 8% and 14% of men (Spalding, 1999).

CVD affects the lives of those afflicted in various ways, many of which can be considered mostly inconsequential, however can have a severe effect on home/general life and occupational opportunities, with many professions requiring applicants to have normal colour vision. Previous literature has reported on CVD's impact on one's career choice, such as doctor specialities (Chan et al., 2014) and the military not accepting applicants affected by CVD. Various daily activities such as determining fruit ripeness or cooking (Steward and Cole, 1989b), and even the ability of children to learn (Chan et al., 2014) are also affected. Some evidence also exists for increased trouble driving (Verriest et al., 1980).

Figure 4.3: Visualisation of the prevalence of different CVD compensation techniques. We can see that techniques applying a spectral shift far outnumber other approaches and we thus further broke them down into subcategories.

4.2 Potential Techniques for Compensation

As part of our research into compensating for CVD, we looked at the compensation techniques utilised for screen-based assistance to find potential solutions that could be applied in our Computational Glasses. These techniques can be identified as belonging to four main directions. The vast majority of works aim to compensate for CVD by shifting the spectral wavelength of identified critical image areas. Another direction is to compensate by changing the amplitude or intensity of identified image areas to increase the distinguishability of colours within. Rather than shifting the colours of image areas, some approaches introduce an artificial pattern to visually separate otherwise similar areas. Finally, a few works exploit the characteristics of human visual perception and create visually salient effects such as binocular rivalry to emphasise otherwise similar image areas. Figure 4.3 shows the distribution of these directions.

4.2.1 Spectral Wavelength Adjustment-Colour shifts

The most common approach for compensation techniques is to adjust the colours such that those wrongly perceived as similar by CVD viewers are adjusted to be colours that are perceived as different.

Early works replaced the colour palette used for the figure creation with a palette where colours remain distinguishable by observers with and without CVD. First detailed by Meyer and Greenberg (1988), various palettes have since been proposed for different tasks (Viénot et al., 1999; Troiano et al., 2008; Jenny and Kelso, 2007b,a).

Among the techniques that target general applications we can differentiate between different colour spaces used to adjust the confused colours.

The most regular space to create techniques is RGB. Due to its relative simplicity and that final colour outputs in computer graphics are encoded in this space many researchers choose to remain in this space. The effects of CVD can be loosely approximated to problems identifying red - protanopia, protanomaly -, green - deuteranopia, deuteranomaly -, and blue - tritanopia, tritanomaly. Because of this easy but inaccurate adjustments can be made in RGB. A common approach has been to use a mutating matrix to shift chromaticity values between RGB channels (Anagnostopoulos and Anagnostopoulos, 2007; Tanuwidjaja et al., 2014; Doliotis et al., 2009; Jeong et al., 2011; Melillo et al., 2017; Itoh et al., 2019b). Other techniques to reassign colours between channels have also been created (Ostia et al., 2019; Jeong et al., 2012b; Lee and Dos Santos, 2011), as well as filters (Poret et al., 2009), highlighting chosen colours (Tanuwidjaja et al., 2014), or changing pairs of colours that users identified as difficult to differentiate (Tanuwidjaja et al., 2014). Alternative techniques working in RGB have used networks (Flatla and Gutwin, 2012), bin-based algorithms (Jefferson and Harvey, 2006), and clustering (Milić et al., 2015; Talom et al., 2016).

The HSV, HSI, and HSL colour spaces have been explored in some early techniques as they provide a means to directly manipulate the hue of colours, the range of which is reduced for those affected by CVD. In the HSV colour space, adjustments of all axes have been considered with H being adjusted in isolation the most common approach (Ribeiro and Gomes, 2013; Ching and Sabudin, 2010; Jia-Bin et al., 2008; Ohkubo and Kobayashi, 2008), although other works have included S (Yang and Ro, 2003) or V (Lai et al., 2009) variance. Whilst in HSL space, techniques adjust red and green pixels by their predominance (Iaccarino and Malandrino, 2006; Ananto et al., 2011). Similarly, adjusting S in HSI has also been used (Chen and Liao, 2011). Generally, adjustments in these spaces are rare and the choice between HSV, HSI, and HSL has minimal impact. Recently, techniques have tended to use more perceptually based spaces such as LMS, *Lab*.

LMS is regularly used as a space for adjustments due to its relationship to CVD. Whilst works in RGB and HSV use simple spaces without direct regard to CVD, LMS can directly, and accurately, represent the effects of CVD. Early works for anomalous trichromats transformed from RGB to LMS (Yang and Ro, 2003; Yang et al., 2004, 2005, 2008; Oka et al., 2010) to create compensations. Various other techniques have also been created (Kotera, 2012; Amano and Kato, 2010) including those that allow for user control (Jefferson and Harvey, 2007; Lau et al., 2015).

Efforts have been made to work in the perceptually uniform *Lab* space. The perceptual uniformity of *Lab* enables the effect of CVD on the discernibly of colours to be measured as distances. This space more accurately reflects the effect of changes that RGB or HSV, and enables the quantifying of colour differences to trichromats vs their dichromatic counter parts. This does however presuppose that *Lab* holds relative uniformity of those affected by CVD. Various optimisation techniques have been created (Huang et al., 2007; Kuhn et al., 2008; Tanaka et al., 2009; Jeong et al., 2012a; Wakita and Shimamura, 2005; Flatla et al., 2013), as well as temporally consistent adjustments (MacHado and Oliveira, 2010; Huang et al., 2011). Clustering (Huang et al., 2009) and image segmentation (Park et al., 2011; Choi et al., 2019) have also been considered. For stereoscopic setups optimisations which look to utilise stereoscopy to enhance distinguishability for dichromats whilst making minimal changes for trichromats have been developed (Shen et al., 2016; Hu et al., 2019a,b).

Transforms between higher and lower spaces have also been used as an approach to compensate for CVD (Deng et al., 2007; Bao et al., 2008; Mochizuki et al., 2008, 2011a,b; Oshima et al., 2016).

Alternatively, some techniques work independent of space dependent characteristics. Machine learning has received some attention (Ichikawa et al., 2003, 2004;

Troiano et al., 2008; Bräunl et al., 2016; Zhang et al., 2018) and self-organising maps have been used (Ma et al., 2009; Orii et al., 2014).

A different approach was taken by some researchers who showed a grey-scale version of colour differences to inform viewers (Rumiński et al., 2010; Nigam and Bhattacharya, 2013). The only works specifically designed for monochromats preserves distances between colours in greyscale for monochromats and trichromats (Rasche et al., 2005a,b).

4.2.2 Intensity Based Adjustment-Lightness shifts

Whilst some works in spectral shifts have looked to maintain the effect or meaning of colours, the semantic meaning of colours are still affected by all spectral shifts, so some techniques look to adjust colours without changing the hue. This is generally one of the less explored areas of compensation with some works in HSV (An and Park, 2014) and *Lab* (Tanaka et al., 2010; Meng and Tanaka, 2016; Meng and Tanaka, 2019).

4.2.3 Patterns

Rather than adjusting the spectral wavelength or amplitude of colours, some researchers have investigated how different overlays can help users affected by CVD differentiate otherwise confusing image areas. These techniques ensure the original colour, and its associated name and meaning are retained whilst still aiding those affected by CVD. However, these techniques can introduce occlusions or false textures. Highlighting the edges between potentially confusing areas (Tanuwidjaja et al., 2014), overlaying hatching (angled, coloured lines) (Hung and Hiramatsu, 2013; Sajadi et al., 2013; Herbst and Brinkman, 2014; Flatla et al., 2015), as well as simply overlaying text of common colour names (Flatla et al., 2015) have all been considered.

4.2.4 Visual Effects

As a final style of adjustment that maintains the original colours, some researchers have investigated applying visual effects that take advantages of the HVS to adjust for CVD. Three such effects have been investigated: the Craik-O'Brien effect (Suetake et al., 2012; Bao et al., 2015), binocular rivalry (Chua et al., 2015), and flicker (Besic et al., 2019; Hasana et al., 2019).

4.2.5 Issues in the related work

Analysis of the methods for compensating critical colours revealed several challenges and research gaps that affect us in pursuing our research goals on building Computational Glasses for CVD. These gaps are also of relevance to general researchers working on computational solutions for CVD. More specifically, there is a lack of

Figure 4.4: Overview on used datasets and evaluations in CVD research. Left: Examples of commonly used natural images to showcase effectiveness of compensation algorithms (from (Rasche et al., 2005b; Kuhn et al., 2008; Jefferson and Harvey, 2006)) Right: Showcase of works that reported only visual results (red), results with participants (green), and statistical results of user studies (blue). The corresponding connection shows that two techniques were compared in a publication. The numbers correspond to the papers in the references. We can see that much of the prior work only reported visual results that are not supported by user studies. Only a few studies involved actual users and even less ran comparative studies.

datasets that provide a widely accepted baseline for testing, formal user studies with CVD participants, and comparisons between techniques presented in the literature.

Datasets When looking at the datasets used in the literature there is a lack of a consistent and accepted dataset that can be used as a baseline. The only dataset commonly used are the pseudo-isochromatic plates (Ishihara plates) that provide an easily evaluable technique for testing efficacy however they are a niche use case designed to be extreme. This makes them abstract with limited direct application to the world being designed to be readily understandable. Whilst this provides an easy way to measure a technique's viability, the applicability of the results based on them to wider scenarios is limited. Works that evaluated the efficacy of their technique not only on tests used to determine if someone is affected by CVD usually utilize only a small dataset of less than 10 other images. Furthermore, most works do not provide any details on these images identifying them only as "natural images". We do note that some works used a large number of images (1000-65000) selected from the internet, however, these are commonly used for building a dataset, e.g., to improve image search results for those affected by CVD. An example of commonly used images (and similar variations) is shown in Figure 4.4.

Not having a consistently used and verified dataset for testing CVD techniques that is representative of real-world scenarios, as well as easy baseline tests for efficacy causes problems with testing and comparing techniques in a manner that can be generalised. Throughout this work, we looked to mitigate this problem by using the common Ishihara plates for baseline efficacy tests and by introducing a set of real-world images with related questions to represent real-world scenarios. We verified that people affected by CVD could not readily answer the questions, however the image set is relatively small and not formally verified. As such we do not redress this issue but hope to provide a basis and direction for future work.

User Studies From the reviewed works only slightly more than half (57 out of 101/ 56.4%) had any evaluations with users, and among these only 29 had more than 10 participants. Studies with more than 50 (3) or 100 (1) participants were almost non-existent clearly highlighting the question of the generalisability of the findings to a larger population.

Among the user studies conducted we identified three common tasks. (1) Colour disambiguation that requires users to either judge what is shown in similar colours (e.g., Ishihara patterns) (Lai and Chang, 2009), or whether areas in an image have the same or different colour (Shen et al., 2016). (2) Colour matching that requires participants to decide which colour patch matches a reference colour, or determine matches between a graph and a legend (Flatla and Gutwin, 2012). (3) Identification of colours where participants are asked to name a colour and compare it against a ground truth, e.g., obtained from participants with normal vision (Chua et al., 2015). For subjective evaluations the most common judgement criteria are naturalness, quality, colour preservation, and confidence when choosing colours. It is important to note that whilst we do not doubt the effectiveness of the presented techniques, only 12 of the conducted user studies reported statistical evaluations whilst of the remaining works only 4 report on user feedback and the rest just report measured metrics. As most of the techniques are only compared against the baseline of no modification, these tend to show improved results on the measured metric. A major limitation of the conducted studies is that they were conducted in laboratory settings. While this allows researchers to control confounding factors and reduce the bias in comparisons against the baseline, they do not indicate how well the system would perform in real world scenarios. For example, how often would users utilize the system deployed on a mobile phone vs. a head-mounted display? Even if participants can imagine themselves using the system in an interview, will this really be the case? In what situations do users actually require continuous assistance, and do the systems perform well for that task?

The generally low number of user studies, limited tasks, and low numbers of participants produces problems with determining the generalisability of the research and the general effectiveness of developed techniques. In our work we look to utilise comparatively high numbers of participants (10-19 per study), however still only utilise a limited set of tasks due to the aforementioned lack of available dataset. Furthermore, due to the pioneering nature of our research we were also unable to extend our studies outside of the lab.

Comparisons The third limitation we identified in the literature is a lack of works that compare the newly developed techniques with existing ones. Out of the 101 works we identified, 62 had no comparisons with previous techniques. From those

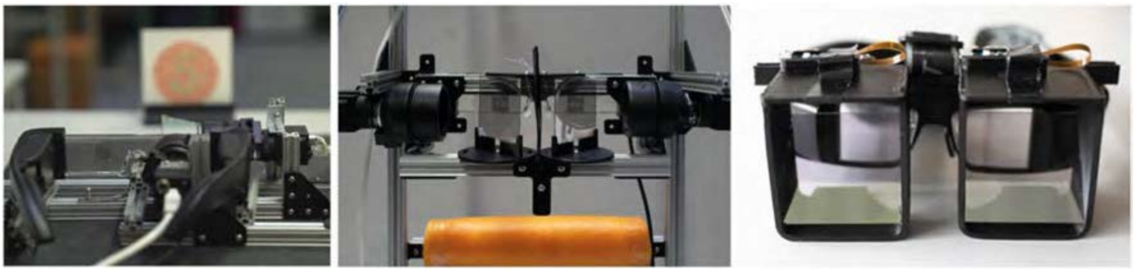


Figure 4.5: The prototypes developed for our research and used in our studies. Left: The bench prototype used in the initial efficacy study. Centre: The stereo prototype used in the efficacy study on user calibrated Computational Glasses and in the study comparing Computational Glasses against an indirect compensation (Google Glass). Right: A version of our portable prototype which was modified for use in an exploratory study.

that had comparisons, the most common was with the technique of Huang et al. (2007) (7), Kuhn et al. (2008) (6), Machado et al. (2010) (4), Rasche et al. (2005a) (3) and (Doliotis et al., 2009)(3). Among the other comparisons 6 techniques were compared against twice, and 17 were compared against once. At the same time, when considering publications that included user studies, we find that only the techniques of Kuhn et al. (2008), Huang et al. (2007), Shen et al. (2016), and Machado et al. (2010) were compared against in two works, and 5 other techniques were compared against only once. Furthermore, whenever algorithms were compared with each other in user studies participants observed differences in "quality", "accessibility", or temporal consistency and found only minor differences in the effectiveness of the algorithms on CVD mitigation. This highlights that although there is a large number of techniques that could be potentially used, there is a lack of knowledge on how they compare to each other, especially when studied with participants.

Figure 4.4 Right shows the comparisons between techniques. It is apparent that there is a lack of comparisons between techniques to show advantages, disadvantages, and improvements. This is exacerbated by the lack of user studies showing the user impacts in these comparisons.

Whilst we look to work around these limitations in this work and provide initial steps to solving them, our primary focus remains to introduce and explore the concept of Computational Glasses for CVD.

4.3 Apparatus

We developed three prototypes that we used to explore the benefits of Computational Glasses for CVD compensation in our experiments. When developing these prototypes, we used commercial OSTHMDs as a base. Here, our focus was on a slim form factor and ease of modification, rather than tracking capabilities that do not play a vital role in Computational Glasses. As such, we did not utilize recent HMDs with in-built tracking capabilities, e.g., the Microsoft HoloLens, Meta 2, or

MagicLeap One.

The first prototype was a monocular bench prototype, where a camera was used to capture the view through the Computational Glasses instead of a user's eye. This prototype was designed to verify that the developed algorithms could indeed assist in discerning colours for those affected by CVD by avoiding any misalignment of the overlaid modulation when users looked through the Computational Glasses.

As the first prototype was only monocular and showed the user the image captured by the camera, not the user's eyes, and to evaluate the performance of Computational Glasses when worn by users, we developed a second benchtop prototype in which users could look at the scene through binocular Computational Glasses. This presented a more realistic testing scenario than the initial prototype.

Our final prototype was a portable version that we used to learn more about what our target users may experience when utilizing Computational Glasses in daily life, as well as their impressions.

In the following we describe in detail the design and the necessary calibration steps of each of our prototypes, followed by the processing pipeline common to all the prototypes.

4.3.1 Bench Prototype

The first prototype we created was a monocular bench prototype (Figure 4.5 Left). The bench prototype consists of an Epson Moverio BT 100 OSTHMD (960×540 pixel, 23° fov) mounted on a table with a PointGrey Blackfly camera (BFLY-PGE-14S2C, with Sony Pregius IMX249 sensor) placed at a right angle to the viewing axis of the *head-mounted display (HMD)*.

We refer to this camera as the *scene camera*. A second PointGrey Blackfly BFLY-PGE-14S2C camera is placed behind the Computational Glasses to capture the view a user would see. We thus refer to it as the *user-perspective camera*. We aligned the views captured by the scene and the user-perspective camera with a half silvered 50/50 mirror, directing equal amounts of light towards each camera. Due to the optics of the Computational Glasses, the amount of light reaching the user-perspective camera is less than for the scene perspective camera and we accounted for that during the calibration step described below. Images captured by the scene camera were sent to a computer to determine the content to be displayed on the Computational Glasses. The combination of the scene and the overlaid modulation was captured by the user-perspective camera and shown on a monitor.

Calibration: To correct colour ambiguities we performed two types of calibration, first a spatial alignment calibration, and second a colour correction calibration to determine the colour response of the cameras and the HMD. We followed the approach described by (Langlotz et al., 2016). First, we estimated the spatial alignment between the two views as a homography (Figure 4.6 Left). While this did not account for any non-uniform distortions due to the optics of the Computational Glasses (Itoh and Klinker, 2015b), these have only a small impact on the alignment quality. We also performed a radiometric calibration of the colours captured by the user-perspective camera to account for optical effects, e.g., screen transparency and colour response, and vignetting (Figure 4.6 Centre). This correction was implemented in a shader and was applied to the computed correction of images captured by the scene camera. We utilized pre-computed lookup-tables for the radiometric and spatial correction for improved runtime efficiency.

4.3.2 Stereoscopic Prototype

As our first prototype provided only monocular correction, we designed a stereoscopic prototype (Figure 4.5 Centre) that provided correction for both eyes. This prototype was built on top of the Epson Moverio BT-300 OSTHMD (1280×720 pixel, 23° fov), as it provided a better colour range and a larger field of view than the BT-100. We used two scene cameras (PointGrey Blackfly BFLY-U3-23S6C) that were mounted similarly to the monoscopic prototype, one camera for each eye. As our goal for this prototype was to evaluate how users would see the modifications with their own eyes, we did not mount a user-perspective camera. Instead, the prototype was mounted on a support structure with a chinrest that users could put their head on for support. This allowed us to ensure that the user's viewpoint does not change when they looked through the Computational Glasses. As the user's viewpoint through the Computational Glasses changed depending on the eye location, we performed a user-specific calibration routine whenever they looked through the prototype.

Calibration: The most common calibration approach for spatial alignment in OSTHMDs is *single-point active alignment method (SPAAM)* (Tuceryan et al., 2002). In our implementation we asked users to sequentially align 8 patterns shown on the Computational Glasses with corresponding matches shown on a monitor in front of them (Figure 4.6 Right). We also detect the location of the patterns on the monitor in the images captured by the scene cameras. From the detected matches, we could compute a homography that aligns the view captured by the scene cameras with the user's view. We performed this process for each eye. We opted for this approach as it is easy to do and generally results in very small alignment errors in the range of 0-5 pixels (Itoh and Klinker, 2014b; Moser et al., 2015). This is approximately 0-0.078° error for our OSTHMD in the optimal case. However, this is subject to user movement and error, and in practice errors can be greater than this. Automatic alignment techniques would require additional hardware such as eye tracking cameras and are less accurate than manual alignment (Itoh and Klinker, 2014a; Plopski et al., 2015; Moser et al., 2015).

4.3.3 Mobile Prototype

We created a portable prototype (Figure 4.5 Right) using a Lumus DK-52 OSTHMD (1280×720 pixel, 40° fov) as the base with custom 3D printed mounts attached in front of each eye. Each mount housed a half-silvered mirror and a Sony IU233N2-Z

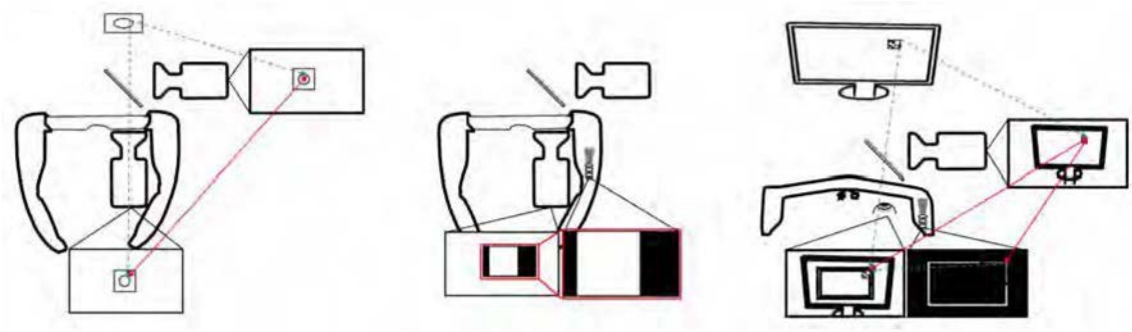


Figure 4.6: Different calibration steps of our prototypes. For the camera-based bench prototype calibrations left and centre were used. For the user based stereo prototype the right calibration was used. Left: The calibration used to align the user-perspective camera with the world camera. This was achieved via feature matching in camera space. Centre: To align the virtual content in the display to the user-perspective camera structured light was used to generate a look up table. Right: To align the virtual content with the user's eye, the user aligned a marker with the virtual content. This marker was also tracked by the world camera and was used to align the world camera with the virtual content and the user's eye.

camera so that the camera could capture an image of the scene from the same viewpoint as the user. As shown in Figure 4.16 we attached our prototype to a head band to counter the increased weight pushing onto the user's nose and improve the stability. The prototype was tethered to a computer via several cables for the purposes of the study, however it could also be connected to a mobile computing device similar to commercial products by NReal and MagicLeap, or be designed to incorporate a processing unit on the device itself, similar to the Microsoft HoloLens V1 and V2.

Calibration: As the calibration we used in the stereoscopic prototype is quite time consuming, we used a simplified approach for our portable prototype. We initially placed a user-perspective camera behind the Computational Glasses and manually aligned the modulation rendered on the Computational Glasses with the pattern shown on the monitor. This step was performed for each screen separately. This provided us with an initial good estimate that was consistent for all users. Similar approaches have also been explored to simplify the SPAAM calibration (Makibuchi et al., 2013). When users put on the portable prototype, they were asked to adjust the overlay so that it correctly overlaid the image of four squares shown on a monitor in front of them. This routine was once again repeated for both eyes, and users were asked to confirm that the overlay remained consistent when shown for both eyes at the same time. We did not expect users to perfectly align the virtual content due to the unstable nature of the headset and limited ability to calibrate the system with a simple calibration. We understand there to be some error we could not expect to control for and consider this when using this prototype.

4.3.4 Processing Pipeline

Once calibrated, the overall processing pipeline for all the prototypes can be summarised by the following. We grab the input from the scene camera. In this image we sense the critical areas that are characterised by having colours that are critical for people affected by CVD (e.g., could be confusing or lacking contrast). We achieve this by running a real-time CVD simulation on the input image from the scene camera. This CVD simulation requires the specific type of CVD as input and the resulting image represents a simulation of how the image is perceived by that person. This technique for simulation is outlined in detail by Brettel et al. (1997) and commonly used when compensating CVD in user interfaces. The difference between the original Trichromatic view (the input image) and simulated Dichromatic view (the CVD simulation) results in an error mask I_{error} that identifies all pixels within the image that are critical in the sense that they are perceived differently for someone having that specific form of CVD. In the next step we compensate the pixels given by the error mask by computing an overlay that changes the colour of the pixels identified in the error mask. There are different compensation techniques that affect the modulation which we will lay out later but most of them change the appearance of the pixel by shifting them towards blue. Finally, the initial calibration gives us the per-pixel transformation that allows us to compute the position at which the overlay needs to be presented to the user via the Computational Glasses so that it aligns with the background. This general overview can also be seen in Figure 4.7. It is worth to mention that our pipeline considers the parameters of the optics (e.g., transparency of the beamsplitters) and colour response of the semi-transparent display to optimise the colour reproduction.

4.4 Study 1: Initial Efficacy

As this was the first study to explore the viability of direct scene modulation for compensating *colour vision deficiency (CVD)* by overlaying graphical cues using Computational Glasses, it was important for us investigate our approach in a well-controlled study, reducing confounding factors as much as possible. We conducted an initial study using the monocular prototype to verify that our algorithms could indeed improve the user's ability to distinguish colours in an Ishihara test. Using this setup allowed us to exclude possible confounding factors such as the quality of the eye-display calibration and because we could always see the visual result presented to the participants, we could assure that the system is working as intended.

Hypotheses: For this initial study we had two hypotheses about the general efficacy.

- H4.1: Using the Computational Glasses, participants would improve their ability to pass a set of colourblind test cases (would not be detected as colourblind while seeing the corrected view).

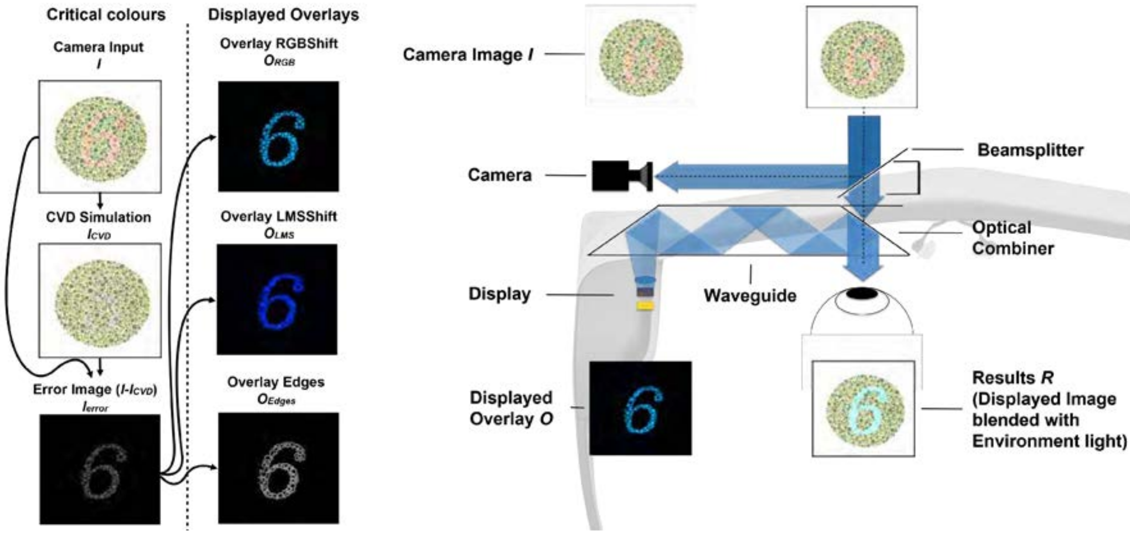


Figure 4.7: Conceptual overview of Computational Glasses for compensating CVD. Left: Overview of our correction algorithm. The input camera image I in simulated as CVD I_{CVD} (here protanopia) and an error image is created I_{Error}. This is then used to create the various modulation overlays used in our studies (here O_{RGB}, O_{LMS}, O_{Edges}). Right: Overview of our hardware setup. The scene camera views the world from the same axis as the user via a beamsplitter. The Computational Glasses are then used to display an overlay that aligns with the user's view and compensates it.

- H4.2: Using the Computational Glasses, participants would feel more confident when recognising the correct content on the plates.

Design: To investigate these hypotheses, we designed a within-subjects experiment with participants affected by CVD. The task required participants to judge the content of four Ishihara plates commonly used for detecting CVD. These plates show figures, e.g., numbers and animals, that are perceived differently by those with normal vision and those with a particular type of CVD. While highly artificial, Ishihara plates provide a readily available measure to test for efficacy. Without a properly verified dataset that can accurately show efficacy on more natural cases, as previously described, we opted for a standardized, widely used baseline. Due to the characteristics of the test, it is commonly used to determine if someone has Protanopia, Protanomaly, Deuteranopia, and/or Deuteranomaly. Hereby, the degree of CVD is determined by the ratio of correct answers, ranging from 0 (all wrong) to 1 (all correct). As it would be too time consuming to administer the test for all techniques, we chose a subset of four plates targeted for the participant's type of CVD (determined beforehand). During the test we measured the ratio of correct answers (0 for all wrong, 1 for all correct), as well as the participants' confidence in the answers on a 5-point semantically anchored scale. This is because it is known that people with CVD can have a different degrees of severity that in some cases allow them to perceive the shown figures, but it becomes much more challenging.

As such, our experiment had one independent variable (Correction technique) with 6 levels (*None*, *RGBShift*, *RGBShiftAdjust*, *LMSShift*, *LMSShiftAdjust*, *Edges*) and 2 dependent variables (ratio of correct answers and weighted confidence).

Apparatus: In this experiment we utilised the monocular prototype where instead of looking directly through the monocular bench prototype, the participants saw the camera feed from a user-perspective camera placed at the position normally intended for the user's eye. The camera feed was shown on a standard 24-inch monitor. The Ishihara plates for each test were placed in such a way that they were not directly visible to the participants.

Compensation Techniques: As our work is the first to investigate CVD compensation by directly modulating the scene with Computational Glasses, we first implemented techniques explored by Tanuwidjaja et al. (2014) as they also utilized an AR device, although they directly modified images captured by the camera effectively creating a HUD. We also implemented a technique based on the work of Jefferson and Harvey (2007) that works in the commonly used *LMS* space. This technique works by shifting points away from confusion lines using a rotation. These techniques can be seen in Figure 4.8. As previous work showed that depending on the severity of the user's CVD the optimal modification may vary, we also implemented a variation of each of the colour shift techniques that allowed users to adjust the degree of the compensation.

RGBShift: Colour adjustments in the *RGB* space are commonly used with screens and this technique was also demonstrated by Tanuwidjaja et al. (2014). After calculating the confusion map I_{error} we applied a linear shift in the RGB space to improve the differentiability. The amount of colour shift was set to values used by Tanuwidjaja et al. (2014). As in our system users perceived a displayed colour as a mix of the background and the colour shown on screen of the Computational Glasses, we computed the colour to be displayed on the screen so it matched the target colour as closely as possible based on Langlotz et al. (2016).

LMSShift: As shown by our literature review, shifting colours in *LMS* space is the most common approach to address CVD. While *RGBShift* adjusts the colours directly in the *RGB* space, *LMSShift* first converts the original image into the *LMS* colour space and computes the adjustment within that space, before transforming the resulting image back into the *RGB* space to be displayed. If the *L, M, and S* components of the colour space are plotted on 3 axes in space, adjustments in the LMS colour space are computed as rotations around the axes. In a similar manner to *RGBShift*, we did not manipulate the original image, but rather create an overlay to optically combine with the background producing the compensation.

Edges: The final correction technique we implemented was to provide an outline of the areas with critical colours as shown by Tanuwidjaja et al. (2014). Whilst the originally proposed technique used black outlines to show edges, we utilised white outlines as the current generation of Computational Glasses cannot darken the user's view. We detected the edges to be shown with a Sobel filter on the error image to

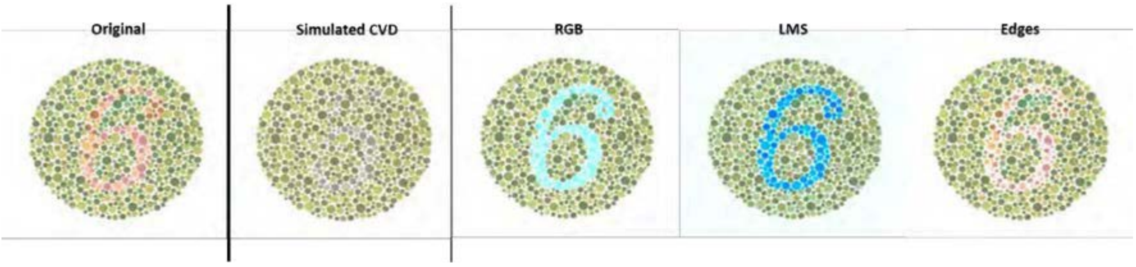


Figure 4.8: Example of compensations applied to an Ishihara plate.

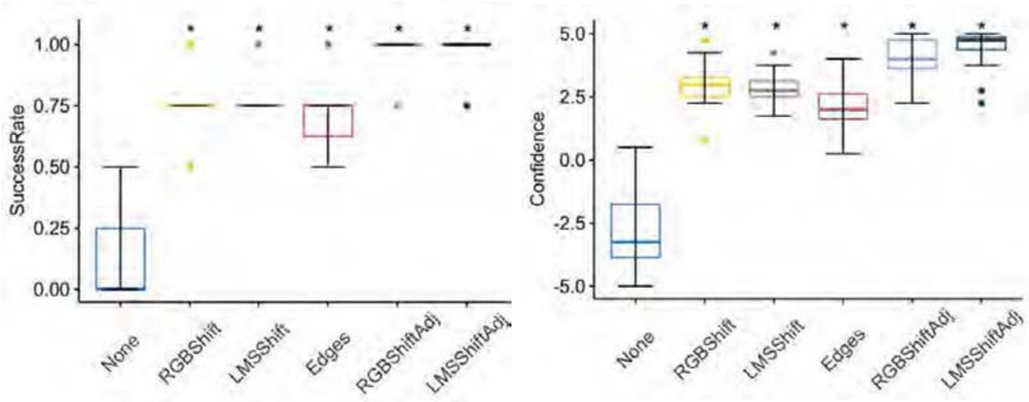
only show edges within areas of confusion. This technique represents a different approach to the *RGBShift* and *LMSShift* by outlining the contours of confusing areas whilst the other approaches look to alleviate the issue directly by adjusting the colour.

RGBShiftAdjust and LMSShiftAdjust: Following previous findings on the user-dependent amounts of required colour shift, we created a modified version of *RGBShift* and *LMSShift* where participants could adjust the parameters of the compensation to better suit themselves. For *RGBShift* the adjustment was the amount of each colour channel that was used in the compensation, and for *LMSShift* it was the angle of rotation around the axes. This was achieved by giving the participants a keypad with 3 pairs of buttons that allowed for one button to increase and one to decrease each parameter respectively. The functionality of this was explained to the participants whenever they were used and also displayed to the user on the monitor.

Task: The task for this experiment was to identify the number on an Ishihara. Looking at the output from a user-perspective camera looking through our monocular prototype the participant was asked to identify if they saw anything on a plate placed in front of the prototype, and if so what. They were then asked to rate how confident they were in their answer given what they could see on the plate. This was measured on a 5-point semantically anchored scale.

Procedure: After reading and signing a consent form each participant completed a demographic questionnaire with information on age, occupation, gender, presence of CVD and type if known, and other uncorrected vision ailments. They were then screened for CVD using Ishihara plates not used in the subsequent study. This was an easy test to use for a screen and was introduced/demonstrated to us by our colleagues in Ophthalmology. It gave us an indication of the form and severity of the CVD. Based on this, the system was set to use a Protanopia or Deuteranopia simulation. The plates used in the study were excluded from this test.

We chose a semi-randomized order of the correction techniques to avoid biasing the results as participants could adjust the degree of correction for *RGBAdjust* and *LMSAdjust*, potentially memorising the correct answer for subsequent tests. During the study, participants completed the task for each plate under all of the conditions, with the order of the conditions randomised. To collect a baseline the first completed the task without any compensation, followed by *RGBShift*, *LMSShift*, and *Edges*



	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.002	-	-	-	-
RGBShiftAdj	0.002	0.003	-	-	-
LMSShift	0.002	1	0.002	-	-
LMSShiftAdj	0.002	0.002	1	0.002	-
Edges	0.002	0.286	0.003	0.283	0.002

Figure 4.9: Results of the bench-top efficacy evaluation with boxplots for success rate (Left) and confidence rates (Right). Significance is marked against the baseline condition ('None'). The full breakdown of p-values is on the bottom with significant values highlighted in grey.

in random order. Finally, to understand if the results can be further improved with a personalized correction, participants were given the chance to adjust the correction strength whilst completing the task in the *RGBAdjust* and *LMSAdjust* conditions.They were allowed as much time as they wanted to experiment with the adjustments and select a personalized adjustment. The study took roughly 30 minutes and participants were compensated for their time with a voucher worth $14USD ($20NZD).

The procedure of this and all subsequent studies was approved by the University of Otago Human Ethics Committee.

Participants: We recruited 19 participants from the student body and staff of the University of Otago through advertisements in lectures and mailing lists. All participants were male (mean age = 24, σ=10.04) and were affected by red-green blindness. We had expected this participant composition due to the higher prevalence of CVD, especially red-green blindness, in the male population.

Results: For each correction technique we computed the success rate as the ratio of correct answers. A Shapiro-Wilk test showed that our data was not normally distributed. A subsequent Friedman test showed significant differences in the success rate ($\chi^2(5) = 79.154$; $p < 0.001$). A post-hoc Wilcoxon signed-rank test (Holm correction) showed that all correction techniques on the Computational Glasses have significantly higher success rates than the uncorrected condition. *LMSAd-*

justed performed best (mean 0.97) and significant differences can be seen between all techniques except *RGBShift* - *LMSShift*, *RGBShift* - *Edges*, *LMSShift* - *Edges* and *RGBShiftAdj* - *LMSShiftAdj*. These results are shown in Figure 4.9 Left.

A Friedman test showed that there was a significant difference between the weighted confidence scores ($\chi^2(5) = 75.332$, $p < 0.001$). A post-hoc Wilcoxon signed-rank test (Holm correction) showed significant differences between all techniques except *RGBShift* and *LMSShift*, with all correction techniques performing significantly better than the uncorrected baseline. Once again *LMSAdjusted* performed best (mean 4.43). It is of note that the confidence score for *Edges* was lower than that of the other techniques (mean 1.99) with a statistically significant difference (highest p value of 0.0195 for *RGBShift* - *Edges*. See Figure Figure 4.9 Right for the results.

We considered a potential correlation between different adjustment values and the selected plates, basically we considered finding perfect parameters for the adjustment strength based on the adjustment by the participants, but we found no consistent results indicating a correlation.

Discussion: The results of our technical validation confirm the efficacy of the approach shown on our bench prototype. More specifically, we showed that we improved the participants' ability to pass colourblindness test cases (H4.1) and make them more confident when answering the questions (H4.2).

One point to note with the results of our study is the learning effects created by presenting participants with all modulations on each image. Any positive or negative effects on the colour information presented to the participant will have subsequently influenced the number they are looking for, and their confidence in what they see. Depending on the magnitude of these effects the magnitude of the improvement demonstrated by a technique will be inversely exaggerated. When observing participants, we did see them return to incorrect answers when compensations did not offer them sufficient aid after successfully correctly answering question , as we did ask them to base their answers on what they could currently see. As such we believe any learning effects would have been more prominent when considering confidence. Regardless, we do not stress an overall winning technique, rather the general efficacy of Computational Glasses.

We saw that personalisation of the correction further improved the participants' ability to discern the patterns, with *LMSAdjust* giving the best results in both metrics. It is important to note that the correction strength participants selected for *LMSAdjust* and *RGBAdjust* was not consistent between the different plates, suggesting that it is important to consider not only inter-personal differences, but also situational and environmental compositions. As the participants were able to adjust the values for *LMSAdjust* and *RGBAdjust*, we did expected that if the techniques could work the participants would be able to learn the correct answer. This means that the participants could increase their confidence in their answers. However, as these results show a significant difference between not only them and the baseline, but also the unadjusted techniques there is evidence to support the improvements gained by customisation of compensations to and individuals needs and situation.

Testing our concept in the monocular prototype allowed us to control for the confounding variable of the eye-display calibration, the quality of which is hard to estimate beyond that reported by the participants. Although this ensured that all participants saw correctly modulated views of the plates, it constrained the colours presented to the participants by the camera response and the display on which the final image was presented.

4.5 Study 2: Efficacy Validation

Whilst our first efficacy evaluation showed a clear improvement in the participants ability to pass a colourblindness test, it was still constrained to a bench prototype where the user's eye was replaced with a camera and the participants only saw the camera feed instead of directly looking through the prototype. As a user's perception of the overlay when looking through the Computational Glasses can substantially differ from what is captured by a user-perspective camera, we conducted a second study to further test the applicability of our correction techniques when participants were actually able to see directly see through our prototype. Furthermore, instead of it being a monocular prototype, we added support for actual stereo vision by utilising our approach for each eye.

Hypotheses: Based on the encouraging observations in the technical validation study, we had the following hypotheses that basically aim to replicate earlier results using a more realistic prototype but still in a well-controlled environment.

- H4.3: Using the Stereoscopic Computational Glasses, participants would improve their ability to pass a set of colourblind test cases (would not be detected as colourblind while seeing the corrected view).
- H4.4: Using the Stereoscopic Computational Glasses, participants would feel more confident when recognising the correct content on the plates.

Design: The general design for this study followed that of our first efficacy evaluation to test the effectiveness of the compensation techniques when observed directly by participants. The independent variables were again the different techniques *(None, RGBShift, RGBShift Adjust, LMS Shift, LMS Shift Adjust, Edges)*, and the dependent variables were success rate and weighted confidence score. Due to the small cohort of potential participants, we used a different set of Ishihara plates in this experiment, allowing a partial overlap of the participant pool with the first study without biasing the results. Overall, we talked directly to several thousand students to recruit the participants for this study.

Apparatus: For this study we used the stereoscopic prototype that allowed the user to directly look through the glasses and perceive the compensation. An LCD screen represented the physical world and displayed content to the participants. The Computational Glasses were mounted in a fixed rig to stabilise the prototype and maintain the relationship to the screen so that participants had a good view through the glasses onto the 27" LCD screen placed at approximately 2 meters away. A chin rest was included in the rig to allow participants to maintain a constant head position more comfortably over an extended period.

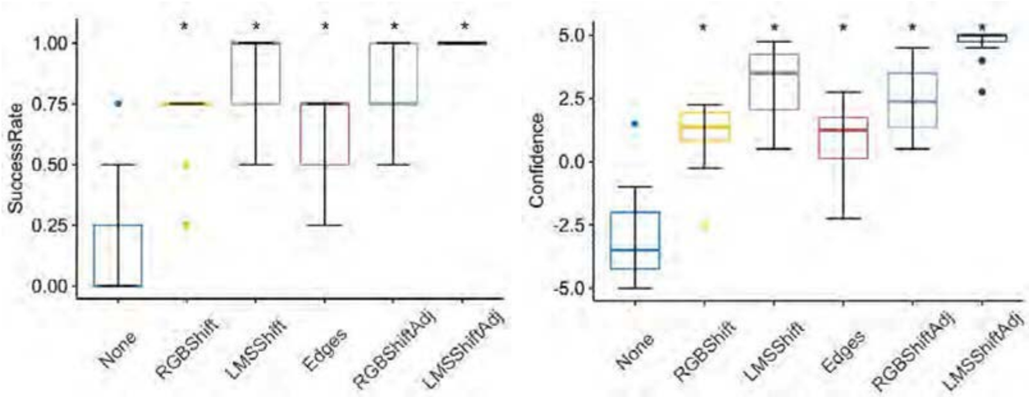
Compensation Techniques: As this was a replication of our prior study, we utilised the same techniques in this study as we detailed and implemented previously.

Task: The task for this study was essentially the same as that used in the Bench-Top Efficacy Validation study. Once again, the participants were presented with an Ishihara plate and were asked to identify what, if anything, they could read on the plate. They were then again asked to provide a confidence rating for their answer. However, for this study, the user was looking directly through our stereoscopic prototype, and the Ishihara plates were presented on a monitor in front of the participant.

Participants: We once again recruited 19 participants (1 female; mean age=23.2, σ=9.23) from the student body and staff of the University of Otago through advertisements in lectures and mailing lists. One participant showed complete colour-blindness which was also medically diagnosed and is extremely rare, and thus was treated separately. The remaining participants showed forms of red-green blindness. Participants were compensated for their time with a voucher worth the equivalent of $14USD ($20NZD).

Procedure: The procedure for this study followed that of the Bench-Top Efficacy Validation study. As the alignment of the scene camera's view with the user's view in the stereoscopic prototype is user-dependent, participants had to perform the described calibration routine before completing the task in the uncompensated condition. As in the Bench-Top Efficacy Validation study, the order of the plates was randomised for each participant, and the techniques were presented in the same semi-randomised order. The study took 45 minutes.

Results: As the success rate was again not normally distributed, we compared the results with a Friedman's Test that showed significant differences between the conditions ($\chi^2(5) = 73.446$, $p < 0.001$). A post-hoc Wilcoxon signed-rank test (Holm correction) showed significant differences between all techniques except *RGBShift - Edges* and *LMSShift - RGBShiftAdj* (Figure 4.10 Left). For the weighted confidence scores, a Friedman's test showed significant differences between the conditions



	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.002	-	-	-	-
RGBShiftAdj	0.002	0.001	-	-	-
LMSShift	0.002	0.591	0.004	-	-
LMSShiftAdj	0.002	0.004	0.010	0.002	-
Edges	0.002	0.020	0.002	0.011	0.002

Figure 4.10: Results for the stereo efficacy evaluation with boxplots for success rate (Left) and confidence rates (Right). Significance is marked against the baseline condition ('None'). The full breakdown of p-values is on the bottom with significant values highlighted in grey.

($\chi^2(5) = 78.926$, $p < 0.001$). A post-hoc Wilcoxon signed-rank test (Holm correction) showed significant differences between all techniques except the aforementioned *RGBShift* - *Edges*, and *LMSShift* - *RGBShiftAdj* (Figure 4.10 Right).

Discussion: The results of our study support our hypotheses H4.3 and H4.4. More importantly, these results support the internal validity of our findings as we not only replicated them with a new set of Ishihara test plates and a partially different cohort of participants, but also in an unmitigated view of the scene with participants viewing the compensation with their own eyes.

While the results supported our initial observation that all techniques improved the participant's ability to distinguish shapes shown on the plate, we also found differences from our first bench-top efficacy study. First, *Edges* performed significantly worse than the other techniques. One possible explanation is that this study utilized a stereoscopic screen. Minute errors in the alignment of the rendered content with the scene could have made the edge outline more difficult for participants to fuse to a coherent image. Another possible explanation is the abstract nature of the task with many edges being highlighted in a small space making it difficult to discern details. It is important to note that this is in line with the lower confidence participants had when using the edge compensation compared to the colour compensation techniques in the first bench-top efficacy study.

We also found that *LMSShift* significantly outperformed *RGBShift*. One possible reason for this is that LMS is a better representation of how humans perceive

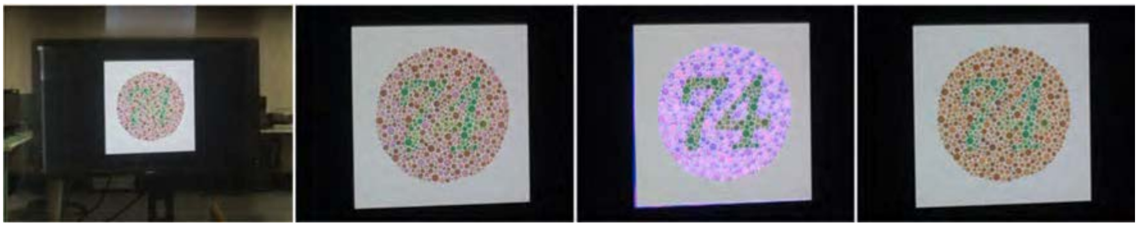


Figure 4.11: From left: view of an Ishihara plate viewed through our prototype with sufficient compensation for a mild case to read a '74', maximum compensation on the same plate to clearly demonstrate the '74', the uncompensated view of the plate where a CVD viewer would read '21'.

colours. Another reason could be that humans perceive colours differently from a camera, resulting in potentially insufficient adjustment. This is supported by the improved performance of the *RGBShiftAdjust* over *RGBShift*.

As with our first efficacy study, due to the replicated study design, we must point out the potential learning effects affecting our results. Due to them viewing the same image under repeated conditions, effects by the compensations on the participants ability to identify what was on the plates will have affected subsequent compensations.

Some notable anecdotal evidence voiced by several participants was that the adjustments used were overly strong. This was unsurprising as we utilised a generic strong shift that could cater to the most severe cases of CVD. For less severe cases a reduced adjustment could suffice, as demonstrated in Figure 4.11 which is verified to be sufficient for a mild case. This demonstrates a need for context aware compensations that adjust to the user's needs.

We also garnered feedback from one participant showing total colourblindness (who we treated separately). Whilst we only had simulations to create compensations for protanopia and deuteranopia we still tested the system's ability to aid them. Unsurprisingly the default shifts we utilised were of no assistance, however utilising custom shifts they were able to achieve a nearly 100% success rate with high confidence.

As discussed with the prior study, individual eye-display calibrations are difficult to verify under the conditions presented by the stereoscopic prototype and we had to rely on participants reporting for the accuracy of their calibrations. Whilst this introduces a confounding variable, the corroboration between our initial study, where it is removed, and this study increases the internal validity of this study. Furthermore, the corroboration between the studies increases the validity of our initial study. The impact of confounding factors such as the compressed colour gamut caused by the camera and display used to present images to the participants are removed.

4.6 Study 3: Comparison to Google Glass

In our previous studies, we demonstrated the efficacy of our approach using different prototypes and by exploring different compensation techniques. As a next step, we were interested in comparing our approach against prior research in the field. To the best of our knowledge, our work is the first to directly compensate the appearance of the scene when seen through an OSTHMD as shown in Figure 4.12 (Top). This makes it difficult to compare our findings with other assistive devices supporting people affected by CVD. The closest to our approach is the work by Tanuwidjaja et al. (2014). In that work the authors used a Google Glass to support people with CVD. It is important to stress here that while a Google Glass is sometimes referred to as an OSTHMD, their approach is fundamentally different. Instead of directly compensating the world by seeing through the glasses, Tanuwidjaja et al. (2014) showed in the Google Glass a camera image that is compensated but this overall visual representation does not align with the real world (is off axis) as shown in Figure 4.12 (Bottom). It can be compared to constantly holding a phone next to one's eye and gazing at it would show a compensated version of the physical world. Given that it is the most related prior research, we decided to compare our approach with that of Tanuwidjaja et al. (2014) to determine whether directly compensating the scene increased the user's mental workload or negatively affected the efficiency of the system.

Hypotheses: Given the positive feedback on our system from our previous studies as well the positive feedback on the Google Glass as reported (Tanuwidjaja et al., 2014) we assumed both solutions would be viable for compensating CVD. Thus, we focused this study on exploring the mental workload and effectiveness for each approach with our initial hypotheses that our approach using Computational Glasses for CVD has advantages when compared to a Google Glass based approach in terms of effectiveness and mental workload mainly because of the direct compensation. However, during pilot tests we realised that we did not consider several factors. Firstly, Google Glass might be more effective because participants have the corrected view as shown in the Google Glass as well as their normal view while our approach always shows a compensation. Secondly and probably more importantly, we had to give the Google Glass approach an unfair advantage. As the camera quality is poor and the display is very small even larger details in the captured environment are completely lost and objects have to be held very close to the eye to be able to compensate for them in the Google Glass (See Figure 4.12 (Bottom-Center)). Once we knew this, we also saw this in the pattern in the original paper (Tanuwidjaja et al., 2014). So instead of using the low-quality camera feed of the Google Glass we decided to load the actual test images in full resolution on the device (See Figure 4.12 (Bottom-Right)). This certainly introduces a co-founding factor as it gives the Google Glass an advantage of having the best possible image quality available. Consequently, we see this study more of an exploratory study with the initial hypotheses of performing on par.

- H4.5: Computational Glasses will show similar mental workload to the state-

Figure 4.12: We compare two compensation styles, our direct overlays (Top) vs the peripheral compensation suggested by Tanuwidjaja et al. (2014) on a Google Glass (Bottom). The camera lens, image quality, and display size when shown on the Google Glass (Bottom-Centre) prevented a correction of small scene elements as they were almost unnoticeable when presented to the wearer. For the comparison study we gave the Google Glass approach an advantage by showing the compensated original image (also shown in the screen) instead of the camera image (Bottom-Right) simulating a high quality zoomed in camera.

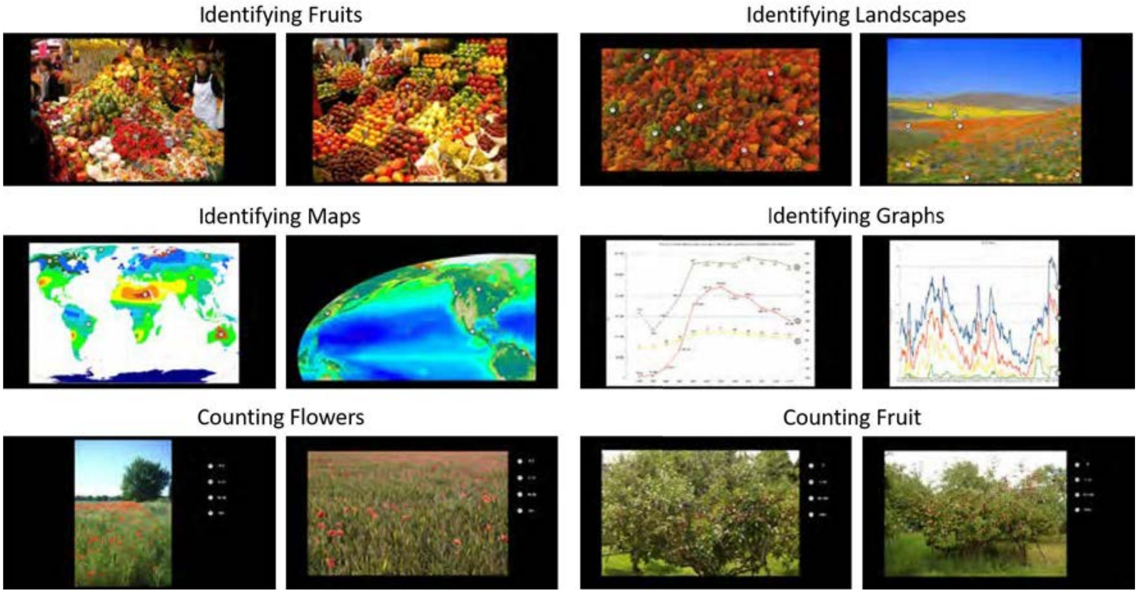


Figure 4.13: Natural image material as used in Study 3.

of-the-art Google Glass.

- H4.6: Computational Glasses will show similar efficiency to the state-of-the-art Google Glass.

Design: We designed a within-subjects experiment with the goal of investigating the workload and efficiency of using Computational Glasses compared to the state-of-the-art system deployed on the Google Glass. To test the differences, we prepared six tasks that users could encounter when using the system in their daily life. Each task had two images associated with it. These had been tested with people affected by CVD beforehand to ensure they would present similarly difficult situations. The images used can be seen in Figure 4.13.

Our independent variable was the task and the display device (Computational Glasses and Google Glass) the correction was shown on. The dependent variables were the workload, measured using the results of the *NASA Task Load Index* (NASA-TLX) questionnaire, and the efficiency, which was measured with a questionnaire created by sub-setting the *System Usability Scale* (SUS). We excluded from the traditional SUS the questions "need support of a technical person to be able to use this system", "various functions in this system were well integrated", and "too much inconsistency in this system" as we felt they were not applicable to our investigation.

Apparatus: In this study, we used the same setup as in Section 4.5. Besides our stereoscopic prototype of Computational Glasses, we added a Google Glass-based solution that replicated the system of Tanuwidjaja et al. (2014). In our pre-tests we identified a significant limitation when utilizing the Google Glass. Displaying an image captured by the built-in camera on the display resulted in details of the environment being very difficult to pick out due to the comparably small field of view of the HMD compared to that of the camera and the general low image quality. As such, target objects would need to be held up very closely to the participant's eyes (about 30 cm away) to ensure discernibility of the details. To address this issue, we advantaged the Google Glass condition by directly showing the modified version of the image in the display at maximum size and resolution. As our studies had shown that $LMSShift$ resulted in best performance and our focus was on comparing user experiences with the different devices, we only presented participants with images compensated using this mode (Figure 4.12). For the study the Google Glass was mounted on the stabilising frame for the Computational Glasses. This ensured it remained stable for use and was in a constant position relative to the display.

Task: The tasks created for this study were: 1) identifying red fruits, 2) identifying red areas within a landscape, 3) identify red area in a complex graph, 4) identify red areas on maps, and approximate the number of 5) red flowers or 6) fruits. For tasks that required participants to identify red-coloured objects, participants were shown a number of labelled options and had to select their answers from among these. For tasks that required participants to estimate the number of objects in the image, we asked participants to select a suitable estimate among four options. At the end of each task the participant was asked to fill out a Raw NASA-TLX and our subset of the SUS. These were completed on the monitor being used to display the images in the task.

Participants: All participants for this study were recruited as the 18 participants with forms of red-green blindness who participated in Study 2 (Section 4.5) because we already had calibrations for them, reducing the time needed for the system calibration.

Procedure: Before beginning the study, each participant read and signed a consent form. The participant then calibrated the Computational Glasses for their eyes. The positioning of the Google Glass was also adjusted such that the display suited each participant. We randomized the order of the devices the participants experienced the correction on. Each participant completed each of the six tasks in a randomized order using one of the two images associated with each task. The participant then changed to the other device and again completed each of the tasks using the alternative image. The participants were allowed as much time as desired.

Results: We compared the NASA-TLX scores of the devices for each task with a paired Wilcoxon Signed Rank test and found significant difference in the answers for graph ($p=0.015$) and map ($p=0.049$) (Figure 4.14 Top). In both cases, Computational Glasses (Graphs: mean=21.76, Maps: mean=33.87) received a lower score than Google Glass (Graphs: mean=34.26, Maps: mean=40.2). A paired Wilcoxon

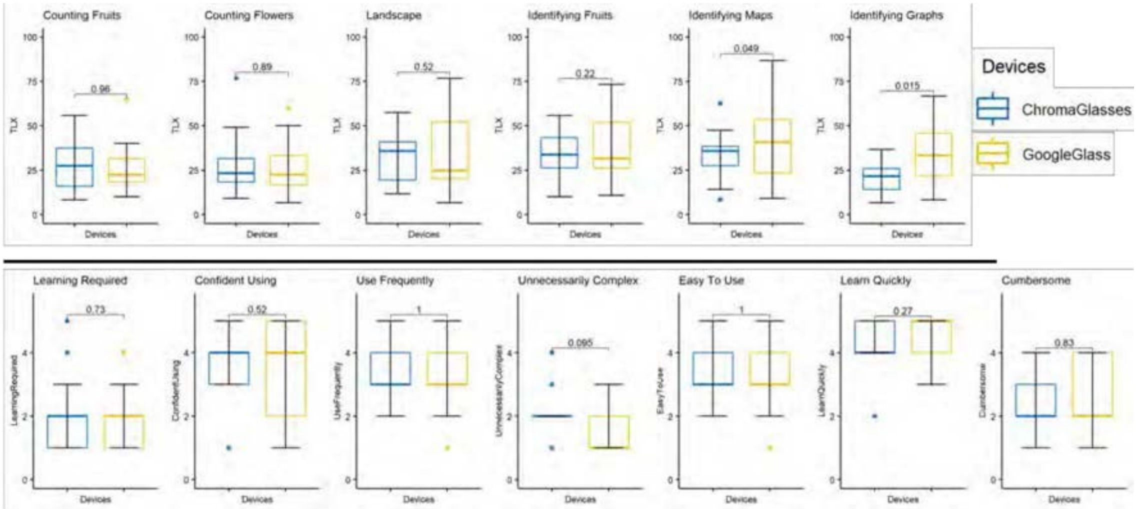


Figure 4.14: We compare the subjective NASA-TLX (Top) scores for each task and questionnaire scores for the two approaches.

Signed Rank did not show any significant differences between the questionnaire scores of the devices for any of the tasks (Figure 4.14 Bottom).

Discussion: Overall this study took a different direction than initially anticipated. While planned as a comparison against an indirect compensation (Google Glass) we realised that the study required changes that are likely to have affected the overall outcome. As such we consider this comparison study to have more of an exploratory character than initially anticipated. Foremost, as we pointed out earlier, we realised that because of the camera and display of the Google Glass a compensation of CVD for small scene elements is not possible unless they are held very close to the user's eye (and consequently the Google Glass camera). This is confirmed when looking at the figures and examples in the original paper (Tanuwidjaja et al., 2014) but was not obvious from the writing or original experiments. The fact that this effect does not arise in our approach and thus we are able to also correct scene details because the display sits in front of the users can already be seen as one of the findings of this study.

To still be able to explore our initial hypothesis we gave the Google Glass an advantage. With respect to these hypotheses, we must reject our hypothesis H4.5. While Computational Glasses performed on par with Google Glass for some examples we also saw a significant reduction in mental workload for others. Given the advantages already given to Google Glass we would expect this effect to be even more pronounced in a fair comparison. When looking at those conditions where our approach performed better by showing a lower workload, we realised that those were tasks that benefited more from not needing to map from real world to the display due to the direct overlay in the user's vision using our approach. This was especially evident in the cases where participants had to pay attention to small details. We confirmed our hypothesis H4.6 that we would not find significant differences compared to the state-of-the-art Google Glass implementation indicating that

Figure 4.15: The images used on the poster used in our explorative study. Participants were asked to read the plate, identify what the image contained, and correctly identify the segments of the pie chart.

both approaches could be similarly efficient in the presented situations. While one could argue this is surprising, we observed that the Google Glass based approach offered wearers the ability to rely on their remaining colour vision and any tricks or subtle cues they are accustomed to. This would reduce the need for reliance on the Google Glass, as well as allowing it to be ignored if desired. Together with the varying severity of CVD in the participants users of the Google Glass based approach have basically several cues they can utilise to compensate for their CVD. In contrast, when using our approach participants had to rely on the compensations provided which they had only recently been introduced to. As such we would expect the Google Glass to have advantages in efficiency when unneeded, whilst the Computational Glasses to be more efficient when needed.

4.7 Study 4: Mobile Evaluation

While our previous studies showed that the participant's ability to discern confusing colours can be improved using the Computational Glasses, the tests were performed in a controlled stationary setting only partially replicating the envisioned context of use. To further explore a user's experience using Computational Glasses in a more natural context, and to explore and identify issues not revealed by our efficacy evaluations, we conducted an exploratory qualitative study with our portable prototype.

Design: The goal of our study was to explore how participants could experience the system in everyday life, gathering information on potential scenarios they could imagine using it, and their general experiences with CVD. As we had already shown efficacy, we opted not to empirically compare the results against another visual augmentation being presented on the display. We considered a potential situation where participants could benefit from a correction via our Computational Glasses and chose a viewing task in which they had to visually explore a presented poster.

The presented poster contained information about CVD, and several images (Figure 4.15) and graphs to support the text. The text also featured questions about each image that participants were asked to answer, alongside stating how confident they were answering the question on a 5-point semantically anchored scale (1, not

Figure 4.16: Left: Our portable prototype based on the Lumus DK-52 modified to increase stability. Right: Participant wearing the prototype in front of the poster.

being sure at all, and 5, being 100% confident). The images shown on the poster were designed to present difficulties for anyone affected by red-green blindness. The images were an Ishihara plate showing a 3 for those with normal vision and a 5 for those affected by deuteranopia, a pie-chart showing a modified distribution of different colourblindness types within the population and small rectangles above the chart colour coding the displayed information, and a field with red poppies in it (Figure 4.15). The field image was also used in our comparison studies. While observing the poster, participants were encouraged to think aloud, describing what they saw through the Computational Glasses. After observing the poster and answering all questions participants took part in a semi-structured interview about their experience, what they saw, their experience with CVD, and thoughts on the benefits of our approach. We provide the questions of the interview in the appendix (B).

Hypotheses: As this was an explorative study we did not have any formal hypotheses. We instead formulated a research question.

- RQ4.1: How would sufferers of CVD perceive having their vision aided by a Computational Glasses technology?

Apparatus: For this study we used our portable prototype described in Section 4.3 which was tethered to a desktop computer for the computational aspects (Figure 4.16 Left) with the goal of collecting general impressions of using a device like ours.

Techniques: We opted for a one-suits-all approach to simplify the procedure, and because we learned from prior studies that this provides good results. Thus, we only used the $LMSShift$ technique for correction.

Task: The task for this study consisted of two parts. Firstly, the participants were asked to explore a poster (Figure 4.16) situated on a wall whilst using our wearable prototype. They were asked to read the text and answer the questions therein that related to images on the poster. Afterward, for the second part of the study the participants took part in an interview with the researcher.

Participants: We recruited 10 participants (average age=27.0, σ=13.3) from the student body and staff of the University of Otago through advertisement by email. Nine participants were male, and one was female. All but one participant had normal or corrected to normal vision. Two of the participants had taken part in at least one of the previous studies, while the remaining eight had no prior experience with the system.

Procedure: After participants entered the experiment room they received an explanation of the experiment procedure, the goal of the experiment, and signed the consent form. After consenting to the study, participants filled out a demographic questionnaire, and a Covid-19 tracking document, and were tested for CVD using Ishihara plates. After that, participants were shown how to wear and adjust the glasses and received a quick introduction. Participants were then seated in front of a monitor at a distance similar to that which they would later complete the first part of the task from. Participants performed the geometric calibration routine described in Section 4.3. Once calibrated, participants were free to walk around with our wearable prototype and complete the first part of the task. After doing so participants were seated and took part in the second part of the task. The entire procedure took about 30-40 minutes, and participants received a voucher worth 13 USD (20NZD) to compensate for their time. Everyone taking part in the experiment wore masks and the equipment was disinfected after each participant in accordance with the Covid-19 requirements and guidelines at the time.

Results: Results gathered from the study pertain to their reactions and feedback whilst viewing the poster, their reflections on their compensated vision afterwards, their thoughts on the prototypical device in its current state, as well as general reflections on being able to correct for CVD and how it affects their lives.

Feedback when viewing the poster: There was a high variation in the answers given for the poster. All participants correctly identified the Ishihara plate, most noting the aid of the blue overlay increases ease and therefore confidence. The modified areas of the chart were noted to only show slight variations in blue when compensated, which coupled with the small legend size, made correct identification difficult, making it still challenging for participants to answer the question correctly or with high confidence. For the poppy field scene participants either noted the individual flowers or realised that something was being highlighted but for the latter they did not immediately identify individual poppies.

Reflections on their compensated vision: Several participants noted the system's great assistance for the plate, but mentioned the confusion that is generated when compensated results in colours already present (e.g., graph). This once more highlights the practical limitations of recolouring compensations such as when not utilising Ishihara plates but utilising realistic scenarios. Shifting colours towards colours already present not only affect the cues those suffering from CVD have come to rely upon but also introduce new confusions when observing a diverse color pallet. While this limitation has been noted before, e.g., (Hasana et al., 2019; Kuhn et al., 2008), it remains an unsolved problem. Participants suggested that context-specific compensations could address some of the issues and that different compensations could be less obtrusive or more suited to specific situations. Utilizing patters or textures was suggested. Alternatively, one participant noted that simple, quick, identification of confusing colours could be sufficient.

Thoughts on the prototypical device: Most participants commented on some of the limitations we know to be current limitations of our prototype but not the general concept. Latency between head movements and the display were mentioned, creating visually incoherent or unstable overlays whenever participants moved their head too quick. While the latency made the system currently ill-suited for non-stop utilization, some participants mentioned it could already help in specific static scenarios. It was also noted that whilst the compensation could be seen on the poppies, in particular when further away from the poster the low resolution of the cameras (further reduced during eye-camera-display adjustment) made individual flowers hard to interpret as high frequency structures got lost ("blobs of blue"). Some participants mentioned that there is a perceptual variance between the overlay display for both eyes. We attribute this variance to the eye position within the eye-box, and to variances in the brightness of the cameras which can easily change perceived brightness. When queried as to who could most benefit from the envisioned approach, most participants noted the effects on career choices (e.g., architecture, medicine, military) and hobbies (e.g., sailing). Three participants raised problems for children and potential bullying, with one recalling such events due to wearing an incorrectly coloured shirt, and another referencing teasing (although not so far as to be considered bullying). Participants thought the glasses could help children distinguish colours or identify when they needed help (a sentiment also echoed by participants for their own use). One participant detailed asking friends for help when using colours at school and almost all highlighted here the issues in their childhood and early youth. While encouraging, it is important to highlight the need to verify these observations with the target audience. The answers could have been affected by biases when recalling ones past situations or imagining another person's needs. Another bias could have been introduced from the participants feeling the need to give some answer to the question posed by the interviewer.

General feedback on their experience of CVD: We asked our participants if they would consider correcting their CVD through surgery, or glasses similar to our prototype. Although most participants felt that CVD did not significantly affect their daily lives, most would consider correction. All bar one would consider glasses, either for daily or occasional use. They did note however that form-factor was very important and would have to resemble current glasses. Specific benefit was questioned, and participants noted that they would want to evaluate resulting corrections. Only two participants would definitely consider surgery, a further two would consider it if it was safe and cheap. One participant had looked into using traditional lenses to compensate for CVD in the form of EnChroma[2]. They preferred our system as according to the participant, the colour shift in EnChroma was of little apparent practical effect. Computational Glasses enabled clearer understanding of confusions. Another participant noted they had explored the same option but cost and time requirements for potential benefits prohibited them from proceeding.

The participants reported common issues in which they perceive CVD to affect their lives. A common problem was colour-coded indicators. One participant mentioned trouble with car warning lights that could be critical or cautionary. Participants also noted a need to double check traffic lights and problems when lights were placed horizontally as opposed to vertically. Other participants recalled problems when colour was used as an identifier, e.g., "green stickers", "green glass bottles". This issue was not restricted to specific applications and is critical particularly in unknown environments. Participants noted that education tests and posters would sometimes utilise colours for graphs and images that they could not distinguish. Even when teachers were aware of the problem, they did not adjust the educational material. Participants also mentioned videos and computer games not supporting CVD adjustment and losing golf balls in grass due to inability to distinguish them. Cooking was also an issue as participants could not easily determine if things were cooked properly (e.g., chicken) or the freshness of ingredients. Interestingly, although absolute colour identification was raised as problematic, i.e., being asked to bring the red bucket, colour matching did not tend to be so as participants could compare the hues to do so.

Discussion: Drawing from the results of our study we find several points of discussion with relevance to future developments of Computational Glasses, AR compensations for CVD, and CVD aid in general. From our study we see the need for considering context-aware compensation for CVD. For example, whilst compensation was beneficial in some areas of the poster it caused problems in others, and more targeted solutions were requested by the participants. The same is true not only for the technique used but also for the overall strength of the compensation so as not to create visual obtrusion thus, needs to be context aware. Whilst the efficacy of using patterns was less than spectral-based shifts in our prior studies, participants expressed an interest in the use of these in certain areas, so using patterns and textures warrant further investigation, in particular the more commonly demonstrated hatched overlays (Hung and Hiramatsu, 2013; Sajadi et al., 2013; Herbst and Brinkman, 2014; Flatla et al., 2015).

The feedback highlights the known issues to perceived latency. This is a general

[2]https://enchroma.com/

challenge in Augmented Reality and *optical see-through head-mounted displays (OSTHMDs)* (Itoh et al., 2021; Lincoln et al., 2016a) and thus anticipated. We argue that higher integration of the components would easily allow for acceptable latency that is on par with existing OSTHMDs used in AR but would be an issue for the engineering and cost of the prototype as we used off-the shelf hardware where possible. Similarly, using higher resolution cameras that are available but costly when bought in small numbers would alleviate the identified issues when preserving or highlighting small details. An interesting point from the interviews was the potential of only highlighting the areas where issues for people affected by CVD might arise so they can deal with them as needed. Whilst many of the works in the literature highlight areas of difficulty inherently by changing them, future work might consider to only highlight but not change or compensate which opens interesting research questions.

Finally, the exploratory study provided encouraging feedback when moving forward with compensating CVD using Computational Glasses. Most participants did view themselves as people who could most benefit from glasses despite the mild nature of CVD. Participants pointed towards careers that they would not be able to participate in and issues early in their lives. We generally noted a consensus that glasses style compensation for CVD would be of interest to participants, either as specific glasses or as an additional feature to current glasses even when compared to a hypothetical surgery. Such surgery does not currently exist, although gene therapy has shown promising results in introducing new colours to red-green colour-blind monkeys[3]. However, the desirability of such solutions was very low and participants were much more open to the idea of a glasses based solution.

4.8 Study 5: Method Comparison

Whilst our initial studies had shown that Computational Glasses could indeed improve the participant's ability to distinguish between colours, we saw a need to further develop techniques that can be properly applied to long term usage and are less obtrusive. This was supported by the feedback from our exploratory study in which participants expressed their interest in having different compensations depending on the current context. So far, we had only explored a small subset of possible compensation techniques. We thus decided to investigate which of the techniques currently present in literature could be adapted for use with Computational Glasses. We saw this as a required step to explore promising future research directions for CVD compensation techniques as well as to have a set of known and working techniques that can be applied depending on user preference or context. Consequently, as our initial tests only investigated user performance in a standard colourblindness test (Ishihara plate), we also aimed to incorporate natural images in our evaluation of different compensation techniques for CVD.

Design: As with the efficacy studies, we designed a within-subjects experiment with different images to view for each algorithm. Rather than simply showing the

[3]http://www.neitzvision.com/research/gene-therapy/

Ishihara plates as before, we also included natural scene images, resulting in a total of 4 tasks (2 on Ishihara plates and 2 on natural scenes) that participants would complete for each condition. The two natural scenes represented different scenarios participants may encounter in real life and participants were asked to either count the number of red flowers in the image, or identify which of the labelled fruits was red. These images were the same as those used in Study 3 (Section 4.6), *"Identifying Fruit"* and *"Counting Flowers"* (Figure 4.13). As our survey identified, these were among the most common tasks in user studies. For the plates, we included those from two different palettes that needed to be read. For each task, we included two images that a researcher who was affected by CVD judged to be of similar difficulty. The correct answers for the natural images were determined by researchers with colour vision measured as normal.

As our dependent variables we collected the success rates and confidence scores when giving an answer. The success rate for the plates and the counting task was measured as in the previous studies. The identification task success rate was measured as the ratio of correctly identified items (0.0 no items identified correctly, 1.0 if all items correctly identified). For the plates and counting tasks we measured the confidence scores in the same manner as in the prior efficacy studies. For the identification task it was measured as the confidence scaled by the success rate. As our review and first study showed that all systems would likely perform well compared to the baseline, we also asked participants to judge the obtrusion of each technique on a 5-point semantically anchored scale (5=most obtrusive).

Hypotheses: As this was an explorative study, we did not have any formal hypotheses.

Apparatus: We used the stereoscopic prototype with the same environment setup as in Study 2 (Section 4.5).

Techniques: We have showed in our discussion on related works and background technologies that there are many techniques that produce a viable compensation for users affected by CVD. For our final study on compensation techniques, we opted not to test all the presented techniques but selected a representative subset of the three main directions we identified in our survey (4.2): Colour adjustments (Spectral and Amplitude), Pattern effects, and Visual effects.

Selection: To select suitable techniques for our comparison, we initially considered all presented techniques and selected a subset that we 1) believed can be implemented in real time based on what was reported in the original works, 2) could be implemented in an additive manner as required to achieve an effect in Computational Glasses, 3) produced sufficiently different output to those already included, and 4) had sufficient implementation details provided for us to replicate them. We applied constraints to the techniques; being able to run interactively on a stream of camera frames, and be implemented in an additive manner as our Computational

Glasses can only add light. We also decided to exclude techniques that relied on user adjustments due to the desired pervasive nature of Computational Glasses and existence of non-computational filters. Whenever possible we only looked at the most current forms of the techniques, with techniques that have been shown to update and outperform older techniques being selected. Techniques that were designed for scenarios outside of Computational Glasses use, such as working for both CVD and normal viewers were also excluded due to the drastic differences in their design goals. We also avoided flicker techniques as we were concerned about their suitability for non-static scenes and potential adverse effects on users. The final constraint we applied was to only include techniques that were designed to work for all forms of red-green colourblindness as this would be the pool of participants that we expected our recruits to belong to, and we wanted to be able to utilise all of them.

This left us with 28 works presenting 15 techniques. As several of the remaining 15 techniques performed similarly, we selected a subset of 9 unique techniques to avoid participant fatigue from comparing repetitive techniques that present little to no new insights.

Final Techniques: Unless otherwise stated, we implemented the remaining techniques using the equations provided by the authors and readers should refer to the referenced works for full implementation details.

RGBShift, LMSShift, Edges: These techniques were already implemented in our existing prototypes as previously detailed in Section 4.4.

HSVShift: The selected HSV effect uses a skillet adjustment to shift H. If colours are considered "red" H is adjusted based on H and maximum S, V as detailed by Ribeiro and Gomes (2013).

Craik-O'Brien Effect: We implemented a modification of the original technique presented by Bao et al. (2015). This technique looks to ramp pixel intensity up around pairs of colours which are perceived as different to normal sighted viewers but similar to CVD, creating the Craik-O'Brien effect. However, during implementation and initial pilot tests, we found that we were unable to reproduce the effect on the Computational Glasses given that we only were able to add light intensities and wanted to be able to compensate smaller details (the original work by Bao et al. (2015) changed larger image areas). Instead, we included the variation of the *LMSShift* technique *LMSReduced* where we reduced the effect size, as we knew this technique to work well based on our previous studies and it had been noted that the techniques could be considered excessive. We thus looked to investigate its effectiveness at reduced levels to see if it would better achieve the goal of an unobtrusive compensation for ubiquitous use.

Contrast: Whilst none of the presented intensity adjustment techniques completely met our criteria, we chose to include an example from this category of techniques for completeness and modified the work of (An and Park, 2014) only including the additive components. Based on the H value we modified S, V using the input S, S_{max} and $S_{average}$.

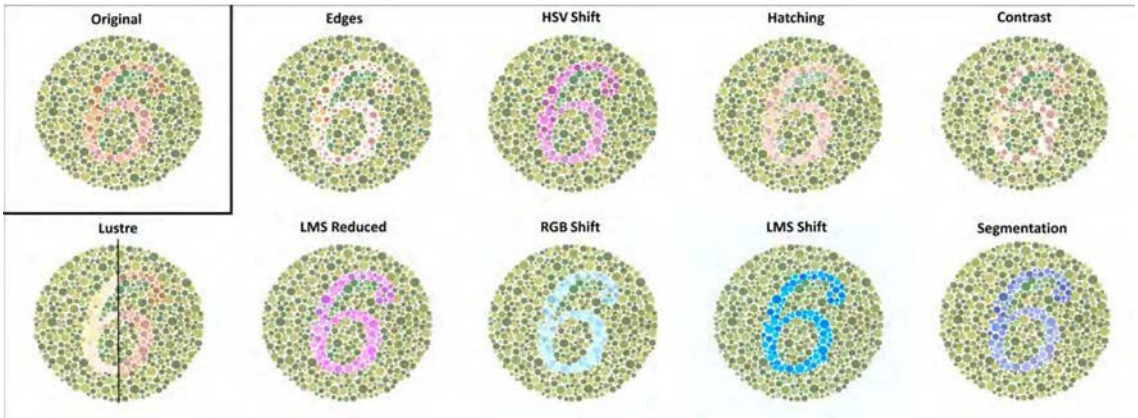


Figure 4.17: The various techniques implemented for the comparison study. Top Left: the original image before modification. Bottom Left: *Lustre* is split to show the modification presented to each eye respectively.

Segmentation: We also included a segmentation algorithm as these algorithms pay attention to the current colours in the image to avoid introducing new issues. To this end we included the work by (Park et al., 2011) whose algorithm works by clustering all colours in an image into patches and testing for colour patches that lie on confusion lines. One colour in any pair that lies on a confusion line is adjusted in *Lab*.

Hatching: We implemented the hatching technique presented by Sajadi et al. (2013) that presents hatching line angles based on the input colour and scaled based on the perceptual difference between dichromats and trichromats. *RGB* values are adjusted towards white based on α, the colour difference for dichromats. Unlike the original work where a grid search was used to create optimal pattern layouts, we utilised a *sin* function to ensure real time operation.

Lustre: This technique was presented by Chua et al. (2015) and works by showing a different image to each eye. If these images are significantly different the brain will be unable to fuse the images and a lustre is seen on affected areas. This is seen as a shimmering effect where the areas of difference shift between the colours present in each of the two images. Whilst the original technique was designed for 3D shutter glasses and used both positive and negative components on multiple levels, we adapted it to account for not being able to subtract light intensities in Computational Glasses. We increase the brightness of affected areas for one eye to produce this effect and used only one level. The level was set using the optimal delta report by Chua et al. (2015).

Our final selection of techniques was: *RGBShift, HSVShift, LMSShift, Edges, Contrast, Segmentation, Hatching, Lustre,* and *LMSReduced.* Example output of each of these techniques can be seen in Figure 4.17 with an uncompensated Ishihara plate for reference.

Participants: For this study, we recruited 10 participants from around the University using the same approach as before. We also reached out to participants from our prior studies and included 4 of them. We only included 10 participants in this study because we felt that our limited recruiting pool was being overtaxed through repeated recruitment, reducing the efficacy of each attempt. This reduces the impact of results but based on our previous results and observations from related work, we expected readily working algorithms to be clearly apparent even with reduced numbers of participants. We removed one participant because their CVD was limited to a level that they could answer plates without aid. We also noted that all, bar one, of the participants for this study presented only minor symptoms, which is in contrast to our prior studies where we had a distribution of both mild and more severe cases.

Procedure: As with the previous studies, each participant read and signed a consent form then completed a non-identifying demographic questionnaire before beginning the study. We screened participants for CVD using Ishihara plates not included in the study. For the actual study the participants were seated in the apparatus and instructed on the calibration of the system. We completed the calibration only once at the start of the study. To familiarize participants with the techniques used in our experiment, we showed them a coloured grid and then applied each technique in turn explaining the effect to them. This enabled them to see how the techniques would or would not affect various colours. Subsequently the participants were shown the first task. The order of the conditions the participant completed each task in was randomized such that the task is completed on each image used for the task in the uncorrected condition first. This allowed for the collection of a baseline and provided participants a baseline for the obtrusiveness of compensations. The task was then completed in each condition on one of the two images in a random order, with the occurrence of each image balanced. The order in which the tasks was completed was randomised. After all the tasks had been completed, the participant was then asked to rate the overall obtrusiveness of each compensation technique on a final image. During this stage the participants were shown the compensation on each of the previous images and allowed to switch between images and compensations to compare them. Throughout the study participants were encouraged to voice their thoughts and these were recorded.

Results To evaluate the efficacy of the techniques we calculated the ratio of correct answers given for each task. Confidence was then calculated per our prior studies. We then treated obtrusion as a semantically anchored 5-point scale. Shapiro Wilks tests showed our data to be non-normally distributed so we tested for differences in score and confidence between each technique and the unmodified condition using Wilcoxon signed rank tests. We also tested obtrusion between each technique using Wilcoxon signed rank tests.

Score: Results for the success rate showed significant improvements for *LMS*, *RGB*, *LMSReduced*, *Segmentation*, and *Lustre* (Figure 4.18 Left). Notably scores were lower in this study than in the prior efficacy ones, and *Edges* no longer showed a significant improvement. To investigate this effect, we split the results for the Ishihara

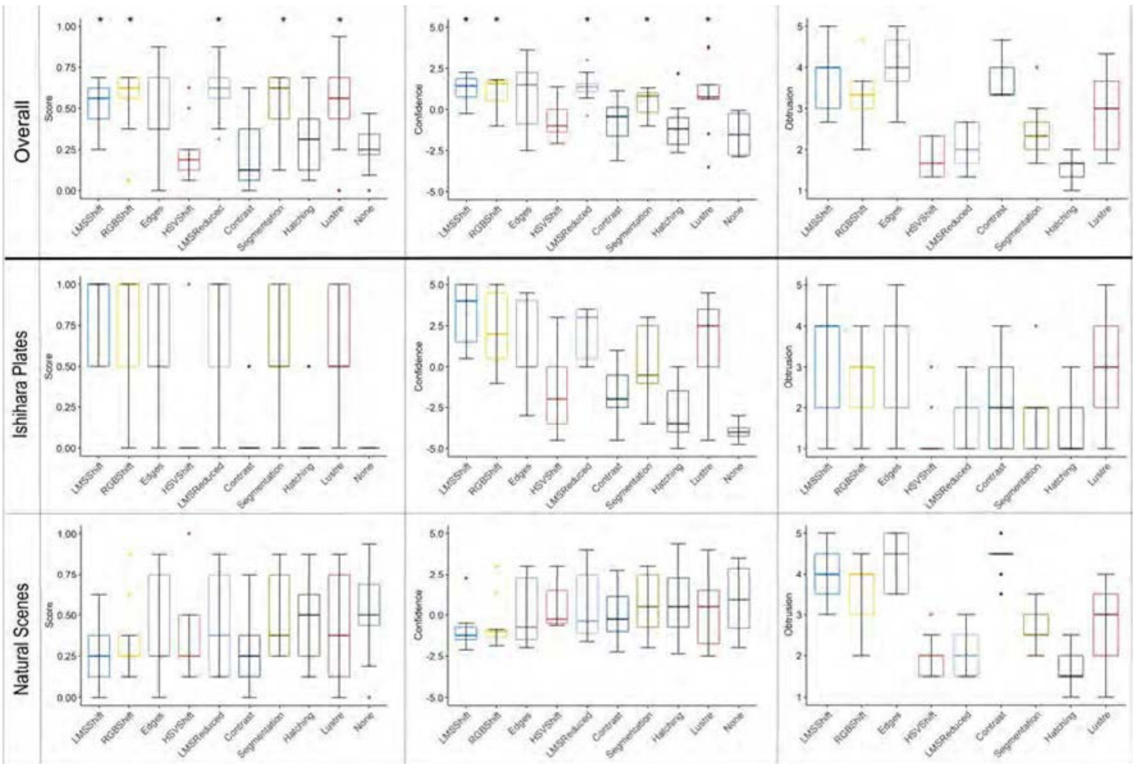


Figure 4.18: Boxplots of the results. From Left: Score, Confidence, and Obtrusion for each of the display compensation techniques. The top row shows the overall results. The subsequent rows split the results by answers for the Ishihara plates and the natural scene images respectively.

Technique	Full (p,W)	Ishihara (p,W)	Natural (p,W)
LMSShift	0.0124, 16.5	0.0005, 07.0	0.0808, 73.5
RGBShift	0.0225, 19.5	0.0016, 11.5	0.1283, 70.5
Edges	0.1598, 31.0	0.0064, 17.0	0.2870, 64.5
HSVShift	0.6762, 56.0	0.5828, 45.0	0.3420, 63.0
LMSReduced	0.0071, 14.0	0.0016, 11.5	0.8791, 52.5
Contrast	0.4947, 59.5	0.3573, 41.0	0.1831, 68.0
Segmentation	0.0187, 18.5	0.0054, 16.5	0.7324, 55.0
Hatching	0.7329, 45.0	0.3573, 41.0	1.0000, 50.0
Lustre	0.0406, 22.5	0.0054, 16.5	0.7326, 55.0

Technique	Full (p,W)	Ishihara (p,W)	Natural (p,W)
LMSShift	0.0045, 12.0	0.0005, 04.0	0.1037, 72.0
RGBShift	0.0112, 16.0	0.0007, 05.0	0.1208, 71.0
Edges	0.0537, 24.0	0.0011, 06.5	0.2563, 65.5
HSVShift	0.1402, 30.0	0.0487, 23.5	0.9397, 51.5
LMSReduced	0.0046, 12.0	0.0021, 09.0	0.9096, 52.0
Contrast	0.4720, 40.0	0.0189, 18.5	0.3843, 62.0
Segmentation	0.0090, 15.0	0.0031, 10.5	0.8796, 52.5
Hatching	0.6229, 43.0	0.2710, 35.0	0.6500, 56.5
Lustre	0.0210, 19.0	0.0535, 24.0	0.9698, 51.0

Table 4.1: Statistical results for Wilcoxon signed rank tests comparing techniques against the unmodified condition for success rate (Left) and confidence (Right). Significant values are in grey. The first columns show the results when evaluating on the full result set, the second if we only consider answers regarding the synthetic Ishihara plates, and the third for only the natural images.

task and the natural image tasks. Here we see an interesting effect. Alongside the previously mentioned techniques, *Edges* also significantly improved the participants success. However, when looking at the natural images we see that no technique has a significant effect on the scores. We show a detailed breakdown of the statistical results in Table 4.1 Left.

Confidence: The results for the confidence again showed significant improvements for *LMS*, *RGB*, *LMSReduced*, *Segmentation*, and *Lustre* (Figure 4.18 Centre). We also looked at the separated results for the confidence and saw significant improvements for *Contrast* and *HSVShift*, however, did not see it for *Lustre*. For the natural tasks we once again saw no significant effects. See Table 4.1 Right for relevant statistics.

Obtrusion: As we were interested in exploring the potential for long term use and unobtrusive compensation, to compare between the techniques we looked at the obtrusion of each technique (Figure 4.18 Right). We found that the *HSVShift*, *LMSReduced*, *Segmentation*, and *Hatching* were all rated lowest with mean values less than three $(1.8, 2.1, 2.4, 1.5$ respectively). After a Holm-Bonferroni correction *Hatching* was considered significantly less obtrusive than all the techniques except *HSVShift* and *LMSReduced*. *HSVShift*, *LMSReduced* and *Segmentation* were not considered significantly different. The original three techniques tested; *LMSShift*, *RGBShift*, and *Edges*, alongside *Contrast* were all considered the most obtrusive, with no significant differences found between them. See Table 4.2 for relevant statistics.

Discussion: Overall, our study comparing compensation techniques on Computational Glasses for CVD delivered mixed findings. Most importantly, we showed that compensation techniques such as *LMSShift, RGBShift, LMSReduced, Segmentation,* and *Lustre* were able to aid people affected by CVD. In general, that would make them suitable candidates for a context-aware compensation of CVD in Computa-

Technique	LMSShift (p,W)	RGBShift (p,W)	Edges (p,W)	HSVShift (p,W)	LMSReduced (p,W)	Contrast (p,W)	Segmentation (p,W)	Hatching (p,W)
RGBShift	0.3025, 64.0	-	-	-	-	-	-	-
Edges	0.4438, 39.5	0.0732, 26.0	-	-	-	-	-	-
HSVShift	0.0002, 100	0.0004, 97.0	0.0002, 100	-	-	-	-	-
LMSReduced	0.0002, 98.5	0.0012, 93.0	0.0002, 98.5	0.1984, 33.0	-	-	-	-
Contrast	0.9077, 48.0	0.2147, 33.5	0.2837, 64.5	0.0001, 00.0	0.0002, 00.0	-	-	-
Segmentation	0.0019, 91.0	0.0119, 83.5	0.0012, 93.0	0.0252, 20.5	0.3561, 37.5	0.0010, 93.5	-	-
Hatching	0.0001, 100	0.0002, 99.5	0.0002, 100	0.2950, 63.5	0.0351, 77.5	0.0001, 100	0.0014, 91.5	-
Lustre	0.0468, 76.5	0.2081, 67.0	0.0122, 83.5	0.0189, 19.0	0.1469, 30.5	0.0378, 77.5	0.5379, 41.5	0.0015, 08.5

Table 4.2: Results for Wilcoxon signed rank tests between each technique. Grey results show significant effects after a Holm-Bonferroni correction ($p = 0.05$)

tional Glasses. However, we need to state here that those overall positive results were biased by the Ishihara plates and do not hold true when looking at the natural scenes in isolation. In fact, when only looking at the selected natural scenes we could not see these positive effects.

There is probably a good explanation for these results. Foremost, the Ishihara plates are designed to be extreme examples. Even those with mild forms of CVD usually struggle with correctly identifying those plates. Thus, it is encouraging to see that the techniques *LMSShift*, *RGBShift*, *LMSReduced*, *Segmentation*, *Lustre*, and *Edges*, performed well when considering Ishihara plates in isolation. Given that our participants overwhelmingly presented as mild cases, we argue that they had less problems with the natural scenes which is evident in the number of correct answers without any compensation technique. Thus, there is less opportunity for improvement. We saw further evidence by looking at the natural scene where participants struggled initially, where we could still see that they improved but the datapoints are too sparse for proper analysis. For example, we saw a change in the items identified in the identification task with one item ('e' in image one) only being identified by one participant unaided but identified by all but one under one or more of the compensations. A reduction in the subjective evaluations of techniques for milder cases of CVD has been previously reported (Wang et al., 2020).

Of the techniques that did not appear to have an effect on the user scores for Ishihara plates (*HSVShift*, *Contrast*, *Hatching*), we noted potential reasons for each. *HSVShift* represents one of the smallest colour shifts, and was one of the least obtrusive shifts indicating it may have been under-tuned for Computational Glasses. Likewise, *Hatching* was considered to be amongst the least obtrusive and has a large parameter space that may warrant further exploration when used in Computational Glasses. Finally, *Contrast* was included for completeness and did not completely match the requirements of Computational Glasses, therefore custom intensity adjustments for Computational Glasses need to be developed before this avenue can be properly explored. That said, a further interesting point on the efficacy of these techniques is the application of *Lustre*, which proved to be an effective visual effect at compensating CVD on Ishihara plates. It must be noted that we did not formally verify whether binocular fusion was being avoided, in which case *Lustre* would have presented as an intensity shift.

When looking at the confidence scores we again saw the same techniques significantly affect the participants' confidence. This would indicate that using these algorithms enables participants to give correct answers with a higher confidence while being less confident in incorrect ones. Interestingly, whilst we do not see a significant effect for *HSVShift* or *Contrast* in the success rate on Ishihara plates, we do see one in the confidence. This would indicate that whilst the techniques were unable to properly aid participants in reading the plates, they did reduce confidence on incorrect answers and increase confidence in correct ones. Lustre showed a significant effect in success however did not show one in confidence, indicating that whilst participants were able to better answer the plates, they were not more confident in presenting correct answers or less confident in presenting incorrect ones.

The final component of our results provided our second important finding from this study. When looking for an unobtrusive technique to compensate for CVD we found that reduced or image dependant colour shifts (*LMSReduced*, *Segmentation*) seemed to prove to be both effective on Ishihara plates and to have low obtrusion. Interestingly, *HSVShift*, whilst not being highly obtrusive, was ineffective on Ishihara plates, whilst colour effects such as *LMSShift* and *RGBShift* were effective but also rated amongst the most obtrusive. We take this to indicate that there exists a range in which colour shift techniques can be utilised to compensate for CVD over which they are effective, however, they must be appropriately tuned. We also saw the only visual effect (*Lustre*) implemented to show promising results in being effective whilst there being no evidence to support them being more obtrusive than the colour shifts, and so view this as a promising research direction. We reiterate again that this result may speak more to the promise of intensity adjustment, rather than that of the intended visual effect, as we did not verify we achieved a lustre. Furthermore, although not reported by our participants, the original work (Chua et al., 2015) observed that lustre could result in discomfort.

Overall, this study represents, so far as we are aware, the first time these methods have been tested for efficacy using the same test, and that all styles of CVD compensation (*Spectral, Intensity, Patterns, Visual Effects*) have been tested on the same test. Despite most works looking to create spectral based adjustment techniques, as we see several of them working, we also show other approaches are viable avenues for creating compensation techniques. We saw that pattern-based techniques, although only in extreme cases, and visual effect techniques also worked. Intensity based techniques have been rarely explored and we saw no evidence that current approaches based on this are viable. As a further point we tested for first time *Contrast*, *Segmentation*, or *HSVShift*, which were not tested for efficacy when originally published. Of these techniques we found that only one of them proved to work as an aid in our Computational Glasses. Whilst this might not be the initially intended environment for these techniques, we believe it reinforces the need for techniques to be properly tested for efficacy using common tests.

We would note that *Segmentation* provided little benefit when compared to other colour shifts, in particular *LMSReduced*, despite high performance costs and while still generating similar shifts. We would propose that for pervasive compensations, further developments in efficiency and effect should be investigated before *Segmentation* can be utilised over a technique like *LMSReduced*.

In conclusion, our study on compensation techniques for Computational Glasses for CVD has shown several viable approaches that can be drawn from current literature (*LMSShift*, *RGBShift*, *LMSReduced*, *Segmentation*, and *Lustre*), and identified those of which were considered to be the least obtrusive and therefore most viable for long term use (*LMSReduced*, *Segmentation*, and *Lustre*). In doing so we have provided the first instance in which several of these techniques have been tested with CVD participants (*Contrast*, *Segmentation*, and *HSVShift*). We can see from our results research gaps in creating effective intensity-based shifts, and a reiteration of the need to make techniques context aware and tailored to users. At the very least user control, as demonstrated in our prior studies, and shown by (Jefferson and Harvey, 2007) is required to enable users to adapt compensations to their needs. Further research is needed into the efficacy of techniques in more natural settings with a larger, fresh pool of participants.

4.9 Discussion

Prescription glasses have been used as a fashionable and simple way to correct for visual impairments caused by refractive issues. While glasses are widely socially accepted there remains a large number of visual impairments that cannot be fully corrected with traditional glasses or optics alone. In this work we showed the potential of utilising OSTHMDs traditionally used for *augmented reality (AR)* to create Computational Glasses that can compensate for one such impairment, CVD. In this work we provided a detailed view into our works including the developed prototypes and studies providing encouraging insights into this technology showing that this technology works and is effective. Our later studies also provided feedback on next steps that would be required to make this approach more mature including addressing latency and context awareness to support continuous usage.

4.9.1 Contribution 1 - Literature review

Given the lack of surveys or meta-studies on techniques compensating CVD, we completed an exhaustive literature review. In doing so, we not only were able to compile a list of the techniques present, but also noted some limitations in the literature that we presented here. These include a lack of comparative studies allowing for the various techniques to be compared, either generally or within the given design space utilised by individual techniques. There is also a lack of a consistent, or even publicly available, dataset usable for testing and verifying techniques that extends beyond the highly tailored and unreal pseudo-isochromatic plates. Whilst we did not overcome the problems presented by these limitations in our subsequent studies as that was beyond the scope of this work, these current limitations are important to understanding some design decisions taken during our research. More importantly, we think the survey is important for other researchers in the field.

4.9.2 Contribution 2 - Computational Glasses for compensating CVD

We proposed the idea of utilizing Computational Glasses built on top of OSTHMDs, enabling a pixel precise compensation of the perception of colours in the environment, to aid CVD. Critical colours are identified in the camera view and the mapping from the camera to the semi-transparent display allows for those colours to be adjusted in the user's point of view. To demonstrate Computational Glasses we created a series of prototypes. These prototypes allowed for a varying trade-off between control of external factors and real-world applicability. An initial monocular benchtop prototype, where the user's view through the glasses, and therefore of the compensation, was mediated by a camera, allowed for complete control of the system. This meant that we could know exactly what the user was seeing and ensuring pixel precise compensation. This allowed for the removal of various confounding factors such as user-specific calibration errors, however, forced the world to be mediated by the camera, reducing the perception of the world to that facilitated by the camera. The second prototype we created was a stereoscopic setup that enabled users to look directly through the system while still being stationary. This removed the limitations of the camera when users viewed the compensation, allowing for proper interaction between the user's vision and the compensation, however, required user-specific calibrations, and removed the ability of knowing exactly what the user is seeing. The final prototype created was a portable prototype. Whilst tethered to a computer and limited by cable performance, this prototype enabled the testing of users that could freely move around and look where they desired. The prototype introduced issues that are likely to happen in the real world such as a more unstable calibration (because someone scratched their nose or moved the glasses) as well as issues that are caused by current hardware such as latency. It did however allow for exploration in a less restricted and more open space and shows the potential for miniaturisation similar to existing glasses.

4.9.3 Contribution 3 - Studies on CVD compensation

We provided a series of studies to prove the efficacy of Computational Glasses. These studies were overall positive in showing the efficacy of the system in several studies compensating for different external parameters and with different participants, contributing to the high internal validity of our findings.

4.9.3.1 Efficacy

With the two efficacy studies we were able to show that using Computational Glasses significantly improved the ability of those with CVD to correctly read example material from a standard test in the form of Ishihara plates in regard to both success rate and confidence in answering. All techniques tested during our efficacy tests showed significant improvements and allowed participants to read the plates more confidently. We found the colour shifts tested to be more effective than outlining in this situation and saw further improvements in participant's scores when we allowed them to adjust the compensation to best allow themselves to read the plates. There was however no correlation between the shifts. We also showed that Computational Glasses were able to assist participants affected by CVD with default compensations and this effect could be further improved with user-customised compensations. The lack of statistically significant improvements on natural scene images in our final study indicates a need for further efficacy testing in more general scenarios but also reflects on a more general issue: The interpersonal differences in severity of CVD and strategies used to cope with CVD allows participants deal with natural scenes better than with specific patterns developed to emphasise CVD. As such we do not think the lack of significance is a shortcoming of our Computational Glasses (as highlighted with the Ishihara plates) but reflects on the improved performance even without any compensation for CVD. Better datasets of natural scenes and long-term usage of Computational Glasses in realistic environments would provide a more realistic picture of the performance in practical scenarios. Also, a more focused recruitment of participants, e.g., screening for severe forms of CVD could have further improved the results for natural scenes. Finally, as we saw in our efficacy studies, individual improvements could be found by adjusting techniques to each participants optimal settings for each test plate. Without any evidence of consistency in optimisation this indicates a need for inter-personal, alongside situational and environmental, factors to be considered. Whilst the Computational Glasses demonstrated efficacy, compensation techniques still require further development regarding when to employ them and how they are optimised.

4.9.3.2 Comparisons

Our findings showed that many developed techniques could be deployed on Computational Glasses, although issues arose when considering how well the techniques work on natural images, and the obtrusion of compensations as perceived by the participants. We also tested our approach against a current state-of-the-art device and found Computational Glasses performed on par, if not better than, the comparison.

To Google Glass: When comparing our system with an off-axis display that shows a compensated image we found that directly compensating the scene did not result in an increased workload and in cases where participants had to distinguish between small details even required less mental demand, once again highlighting its potential for spontaneous or even potentially uninterrupted continuous use. We did not find significant differences in the usability between corrections by Computational Glasses and Google Glass, which we hypothesised due to the fact we had to modify the content shown on the Google Glass to make it more accessible, and that such a system allowed retention of a personal cues used to differentiate colours whilst participants had limited time with Computational Glasses to learn the new ones. If participants had to zoom in manually or could use it only with objects held in their hand the results could have been different. An alternative explanation for there being no significant difference between the two conditions can be found in our final study. There we saw that the compensation used did not in fact provide statistically significant assistance on the natural images, only being assistive in anecdotal cases, and therefore it may have increased the workload and hindered usability when not providing significant benefit.

Techniques: We explored existing compensation techniques addressing CVD for their suitability to be applied within Computational Glasses and eventually context-aware Computational Glasses. In general, we found most techniques applying a spectral shift perform well. This is also true for the technique using a visual effect (Luster effect). However, as previously pointed out we saw the biggest gains mainly in extreme scenarios such as all the Ishihara plates. Expanding on the lack-of notable gains on natural images, we attribute much of this to our participant pool, and the smaller opportunity for improvement. However, many of these techniques have been demonstrated to work previously, including in our own studies. This disparity in results further highlights the need for natural image datasets for CVD research that show realistic everyday issues for people affected by CVD that are also validated. Many of the purposed techniques in literature rely on spectral shifts or have significant impact on the original scene. This is fine in extreme cases, and when the semantic many of colours is not required, however produces issues in more natural scenes. Techniques also rely on implicit understanding of compensations, such as the angles of hatching patterns. Further research is needed optimizing effectiveness together with minimal obtrusion.

4.9.3.3 Portability

We also conducted an exploratory study with our portable prototype showing the potential for Computational Glasses to be extended beyond the restricted lab prototypes previously used. Whilst the results of this study were intentionally of a qualitative and exploratory character, we still saw similar effects of the compensation on the participant's ability to discern previously indiscernible colours. Viewing pseudo-isochromatic plates was greatly assisted, whilst other images saw lesser effects. Latency was still a problem with the portable prototype, exacerbated by the current constrained to tether the prototype to a computer. The primary component of this study was the qualitative data given by participants about the desire to correct for CVD and how it affects their lives. To this end we received very positive feedback to the notion of compensating for CVD with devices with a form factor similar to glasses, a feature of Computational Glasses. We also produced further findings regarding how those affected by CVD view its effects on their lives and who might benefit from assistive technologies. Further points raised show the need for customised, application cognisant compensations.

4.9.4 Limitations

Several limitations for this research should be mentioned.

Lab prototypes: Although we tested with varying prototypes which were increasingly closer to a real-world device, we only tested Computational Glasses within static environments so it remains to be explored whether our findings can be replicated in dynamic scenarios, e.g., while driving or walking down a street. Nevertheless, if latency is minimised, we are confident that Computational Glasses could be used in these scenarios as well as we did not restrict the participant's head movement in our evaluation with the portable prototype. Deploying Computational Glasses in daily life would require advances not only in the form factor, weight, and battery life of the device, but also with respect to lower camera latency (Itoh et al., 2021) and integration into an untethered version of the system. As such, the developed prototypes are still not viable for longer term everyday use with problems such as size and comfort, as well as with currently remaining tethered to a static computer rather than a mobile computing unit.

Number of participants: We managed to recruit between 10-19 participants for each study which can be considered as a limited sample size. However, we must mention that we were limited here by the general prevalence of CVD that is between 5-10% of the male population. Additionally, we observed a general hesitation of acknowledging any form of impairments which can be easily explained by prior experiences such as those mentioned during study interviews. Anecdotally, we had to personally talk to more than 1200 students (visiting their lectures) in addition to university-wide advertisement via email to recruit the participants for the first two studies demonstrating the practical hurdles. Furthermore, whilst targeted towards general CVD, we limited participants to those with less sensitivity in the red-green spectrum of CVD as this represents the vast majority of cases (we accepted more rarer forms of CVD into the study but did not use them when reporting on the study, only using them for their exploratory feedback). It should still be emphasized that to our knowledge, our studies are among the largest exploring compensating techniques of CVD. Lastly, as we repeated similar studies with mostly different cohorts and prototypes, we think we further increased internal validity beyond what the numbers of individual studies indicate.

Limited exposure time: While we performed an initial exploration of potential situations where participants could benefit from the system, and postulated its long-term use, we did not deploy it over a long duration to explore how users would utilize it in daily situations.

Specific CVD calibrations: Throughout our work we did not look to fully diagnose participants, nor look to explore the various CVD simulation calibration techniques so utilised the simpler, and computationally efficient simulations provided by Brettel et al. (1997).

4.9.5 Future Work

There are various potential areas of research that could be explored building on from this research.

Further miniaturisation/mobilisation: From a practical position, whilst we showed first steps towards a mobilised and fully functional prototype, further work towards a fully portable and longer-term deployable prototype is needed. Doing so would help evaluate techniques in more realistic scenarios and longitudinal studies, improving generalisability of the research. It would also help to further research in the general field of providing vision augmentations. Here we hope in particular to benefit from industry efforts in miniaturising and improving OSTHMDs used for AR.

Creation of a dataset: As noted several times, a comprehensive dataset for testing compensations still needs to be created. We saw in our final study that compensation techniques struggled to provide notable assistance on the natural scene images. We believe this to be due to the limited room for improvement on these images for our participants. In order to better develop and tune compensations techniques such that they might provide assistance in such scenarios, a finely tuned and verified dataset is needed to enable evaluation of techniques as to their applicability to the real world.

Comprehensive comparisons: Whilst we looked at various kinds of techniques to be applied on the Computational Glasses, a more comprehensive set of comparisons between the techniques on some of the various metrics supplied, such as maintaining naturalness, efficiency, or effectiveness, would better enable future works to draw on prior studies when looking to create new and improved compensations for CVD.

Context awareness and continuous use: On the note of improved compensations, during our research we clearly see a need for further development of context aware techniques for compensation. As evidenced by our portable and comparison study, introducing context awareness to the system should improve generalisability and use in real-world scenarios. However, this remains an emerging research topic even in the field of AR (Grubert et al., 2016) but the use of these devices as vision aid demonstrates the need for it. An adaptive user interface is one that "remains well designed even as its world changes" (Browne et al., 1990) and for Computational Glasses to be used in a generalised scenario of continuous use they need to provide such an adaptive interface so that they can adjust to the needs of the user. A simple solution to the need for adaptation in algorithms is to provide simple user control (adaptability) (Grubert et al., 2016). As mentioned in the survey, several papers have covered the introduction of adaptable parameters for techniques (Jefferson and Harvey, 2007; Lau et al., 2015). In our own studies we also saw the advantages that this could provide. Enabling users to adjust compensation degrees, styles, and enable/disable compensations (particularly if other assistive aids are also to be provided on the Computational Glasses) as desired would help alleviate the aforementioned issues. This does however place the onus of use on the user. As with the use of phones to aid CVD this produces problems for a constant persistive aid and requires users to know when they require assistant, even when their needs preclude them from knowing so. As such, we believe that ultimately a context aware solution for Computational Glasses that adjusts the interface implicitly (adaptivity) (Grubert et al., 2016), whilst still allowing for adaptability by the user, is needed.

Individual vs combination techniques: Another largely unexplored component of the field is the ability to utilise multiple styles of compensation such as patterns and colour shifts to enable better compensations in conjunction with utilising different styles within different contexts. As noted by the participants in our explorative evaluation, some wanted stronger shifts on the plate whilst also asking for different styles of shift on the rest to reduce obtrusion.

User-specific compensation: To reduce the potential impact of poor personal calibrations for simulations and compensation levels we largely looked to avoid this throughout our research. This allowed us to ensure that all participants were exposed to the same conditions, allowed for easier replication of our studies and results, and avoided the need to individually optimise everything without a baseline to optimise against. However, as noted in our efficacy studies, individualisation of compensations does have an effect on the results of compensations and furthermore CVD is inherently variable between people so adjusting simulations to suit individuals may change the results. How best to individualise masks/simulations and produce tuned parameters for individual preferences and requirements is still an open and interesting area of exploration, although some works have made inroads in this direction.

Exploration of simulations: As a final note, alongside user-specific compensations, an exploration of simulations and tailoring these to compensations/users/situations is needed. For our studies we stuck with the simplest forms of simulation to ensure real-time operation and to prevent a confounding variable.

In summary, we have demonstrated the potential for Computational Glasses to be used as visual aid for CVD by compensating the view of users with augmentations. Whilst further work is needed for long term use, our results show promise. We saw that the Computational Glasses could assist those with CVD, however, saw issues with the compensation of natural scenes where improvements were not as apparent. Given the generic settings used for our compensations and the limited dataset available for testing natural scenes, we firmly believe that the creation of tailored, user-specific, adaptive, compensations evaluated using a comprehensive dataset is needed to alleviate these issues. This research has strong implications for the ability of Computational Glasses and Augmented Reality to further aid human vision and we would even go so far as to indicate implications for augmentation of unimpaired vision, or contribute to an Augmented Human with real-time compensations based on image analysis providing additional visual information.

Chapter 5

Vision Augmentation Use Case: Visual Guidance

In this chapter, we discuss the use of Computational Glasses as a means of visual augmentation to guide users to areas of interest. We present a design space for visual guidance in augmented reality (AR) that covers the various styles of guidance that can be created, the devices on which they are employed, and how much they look to stand out to the user. To cover a gap in this design space that is of particular relevance to Computational Glasses we created our own guidance method which we then verified via a user study. Finally, we look to explore our design space to find future research directions for Computational Glasses.
The contributions of this chapter are our design for on-screen visual guidance in AR, our implementation of a saliency modulation for visual guidance designed for optical see-through head-mounted displays (OSTHMDs), and our studies validating our guidance technique and exploring our design space.
This chapter draws on papers (Sutton et al., 2022b) and (Sutton et al., TBD).

Having successfully demonstrated the ability of Computational Glasses to modulate users' perception of colours to compensate for *colour vision deficiency (CVD)* and providing aid to the *human visual system (HVS)*, we changed our focus to providing more general visual assistance. Our prior results on providing aid indicated that providing visual augmentation to unimpaired vision should also be achievable. To investigate this we looked at how Computational Glasses could be used to provide visual guidance. Visual guidance has been demonstrated as means to support the HVS and to assist with many tasks. It has been used for training (Yoshimura et al., 2019), task assistance such as search (McNamara et al., 2008), as well as having benefits for focus (Koshi et al., 2019), and memory (Veas et al., 2011).

Many different methods for providing visual guidance have been explored. Traditionally, methods using geometric cues such as circling targets, pointing to them with arrows, or more complex guides have been proven to be effective and simple means of providing guidance in *augmented reality (AR)*. However, other methods have also been developed, such as using temporal variances, for example flickering the luminance at a target (Bailey et al., 2009), or modifying the natural scene and its saliency to provide subtler means of guidance that do no occlude the original view (Mendez et al., 2010). Outside of AR, diegetic cues have also been shown as an effective means of guidance, and whilst being effective for story telling, they are less applicable to general and spontaneous AR.

Problem selection. As a means of providing visual guidance we wanted to investigate using subtle saliency modulations to provide an adjustment to the user's view of the world and draw their attention to target areas. In a manner similar to CVD research, there is a strong understanding of the role of saliency in attention, and prior works pitched this as an interesting means to provide visual guidance with great potential and that suits the prospects of Computational Glasses. This prior understanding and research provides us with a basis to draw upon. We found saliency based guidance to be an interesting approach for visual guidance as, whilst fitting with the notion of direct vision modulation of Computational Glasses, saliency has been shown to be a powerful means of drawing attention, as mentioned in Chapter 2 (Figure 2.2) that can compete with geometric techniques. More importantly, and interestingly, the interaction with the HVS means that it has an effect even in more subtle ways and the utilisation of this has been previously demonstrated (Veas et al., 2011). Being able to exploit this more subtle means of guidance would enable a method of visual guidance that could be used as an alternative to, or in conjunction with, more extreme, grabbing geometric methods such as circling targets. Such techniques are known to cause tunnelling effects on the target area, preventing proper scene awareness. They also contribute to visual clutter, particularly in scenarios where other AR content is being added or multiple instances of guidance target, and can occlude areas of view, interfering with context and scene understanding. A subtle saliency guidance would avoid these issues and have applications to areas where occlusions can be critical, such as medical procedures, as well as times where

Figure 5.1: Examples demonstrating the results from using Computational Glasses to adjust the saliency of a scene. Right: Exemplar mobile prototype. Note that this prototype was for demonstrative purposes only and not used for studies or image generation. Top: Pictures used in our studies for testing the efficacy of our techniques. Bottom: The pictures adjusted using our saliency modulation technique to provide varying degrees of visual guidance. All images were captures through our bench prototype. Targets are, from left: a branch of the tree in the top left, the small building in between the two skyscrapers, the computer under the table on the far right.

over-reliance on visual guidance can be counter productive, such as training. As visual guidance is generally a common area of research in AR we investigated this overarching field, before looking to create modulation techniques for Computational Glasses. We again could look to verify the capabilities of Computational Glasses by testing we could effectively provide visual guidance, before looking to open up future research directions and possibilities.

In this chapter, we cover the design space we created for AR visual guidance as a means to inform creating our own methods of visual guidance for Computational Glasses. Building on our success with spectral modulations we filled a gap in the design space for creating saliency modulations on *optical see-through head-mounted displays (OSTHMDs)*, and our Computational Glasses built on them, by implementing a custom saliency modulation technique. We did that as we considered saliency-based guidance to be an interesting means of guidance with a lot of potential for Computational Glasses. We verified the efficacy of our saliency modulation, focusing on the generally defined use for such algorithms as subtle alternatives to geometric guidance. Having completed this technique, we then looked at evaluating more of our space opening up future research directions, comparing our saliency technique alongside a geometric technique from AR guidance, and a temporal technique from AR which is applicable to Computational Glasses. Our results show that both saliency and temporal techniques can provide effective techniques for visual guidance in Computational Glasses, with saliency providing more subtle means of guidance, and temporal being able to perform on par with traditional OSTHMDs techniques. Example output from our Computational Glasses for visual guidance can be seen in Figure 5.1.

5.1 Visual Guidance design space

Visual attention guidance as a general concept has application in many areas of research. It has been used in various applications such as: training; assisting search; reducing distraction. In our work we looked to create a design space for exploring the attention guidance in Computational Glasses. As these devices are new and techniques are yet to be explored on them, we expanded our space to encompass firstly OSTHMD techniques and then to general AR as these are both areas lacking in research. We followed prior demonstrations of the efficacy of design spaces and covered the various AR devices detailed for use in prior literature, summarised the various cues demonstrated as they relate against the mechanisms of human attention, and the overtness goals of the techniques. Whilst there are existing frameworks and comparisons of guidance techniques, they fall short of providing a complete, precise, and clear design space for AR based visual guidance.

5.1.1 Defining design spaces

Design spaces have been shown to be an effective way to systematically explore existing interface components and identify potential areas for further development by filtering and traversing the space (Hirzle et al., 2019; Goel and Pirolli, 1992).

5.1.2 Style of AR

Various devices have been used to produce visual guidance in AR. Many of the proposed techniques have been demonstrated on computer screens using images taken from cameras (Veas et al., 2011; Lu et al., 2012; Wagner and Rozenblit, 2017; Dorr et al., 2004). Whilst this may not be AR in implementation, in practice this model of guidance is implementable in *video see-through head-mounted displays (VSTHMDs)*. Alternatively, techniques have been demonstrated to work with *spatial augmented reality (SAR)* (Booth et al., 2013; Miyamoto et al., 2018; Volmer et al., 2018; Khan et al., 2005; Ueda et al., 2020) and OSTHMDs simulated demonstration in *virtual reality (VR)* environments (Biocca et al., 2006; Renner and Pfeiffer, 2017) or directly in OSTHMDs (Schwerdtfeger and Klinker, 2008; Biocca et al., 2007). As previously covered, OSTHMDs are currently considered more practical for AR usage as they do not constrain the user's vision to that presented by cameras like VSTHMDs. They do however introduce issues with only being able to add light and having limited light control. Projectors also bring further constraints of limited mobility and affecting the light for anyone present.

5.1.3 Cues

During our research we have identified three cues that can encompass the previously demonstrated techniques in the design space, and correspond to the means on affecting human: Geometric Cues, Temporal Cues, and Inherent Salience Cues. These cues encompass the means given in psychology to affect attention, namely;

the visual salience, scene schema, and motion (Goldstein and Brockmole, 2016). Whilst here we detail the techniques that demonstrate each cue individually, many of the techniques incorporate other cues to greater or lesser degrees, as shown when viewing the space.

Geometric Cues. We define geometric cues as:

> *Any modulation or alteration that introduces new geometrical shapes into the user in order to draw attention.*

This includes techniques commonly used to introduce overt geometry such as circling areas of interest of pointing with arrows, as well as when techniques introduce subtle adjustments in the geometry, such as demonstrated by the technique Subtle Gaze Direction (Bailey et al., 2009) where the use of oval masks introduces a subtle oval shape into the scene alongside other primary cues.

The only axis that is commonly applied in current AR is the use of geometric primitives and geometry manipulation. One early method of achieving this was to outline target areas with boxes (Biocca et al., 2006, 2007; Schwerdtfeger and Klinker, 2008; Yoshimura et al., 2019; Schwerdtfeger et al., 2009), which also utilise tunnels or arches to guide users to targets outside the displays field of view. Simpler methods such as placing a coloured dot at the point of interest have also been demonstrated (Dorr et al., 2004). Various geometrical constructs have been shown with lines, circles, squares, arrows, and cubes being popular demonstrated in SAR (Volmer et al., 2018), on-screens (Wagner and Rozenblit, 2017; Waldner et al., 2014), VR (Yoshimura et al., 2019), and simulated AR (Renner and Pfeiffer, 2017). Recently the combination of arrows and attention tunnels has also been demonstrated (Hein et al., 2020).

Geometric cues have been shown to work well in AR (Renner and Pfeiffer, 2017; Volmer et al., 2018), with the added benefit of being able to direct the user to points outside of the field of view (Renner and Pfeiffer, 2017).

Due to the use of VR headsets as part of VSTHMDs, many approaches used for guidance in VR find application in AR and vice versa, particularly those using geometric cues (Yoshimura et al., 2019; Renner and Pfeiffer, 2017).

Temporal Cues. As an alternative to geometric cues, temporal cues such as flicker have been used. We define temporal cues as:

> *Any modulation or alternation that relies on changes over time in order to attract attention.*

One of the most common temporal cues is a flicker effect that is created by changing the luminance or colour in an area between dark and light to draw attention. This has been detailed for use in AR as *Subtle Gaze Direction (SGD)* (McNamara et al., 2008; Bailey et al., 2009), which has seen application and testing in VR, Dome, and SAR environments (Grogorick et al., 2017, 2018; Booth et al., 2013). SGD with both colour and luminance variances have been tested, however luminance was found to be more effective than colour flickering. One issue noted with utilising a flicker effect is the nuisance it can create. In order to overcome this SGD works have introduced eye tracking to remove the flicker before being directly observed. Interestingly, (Grogorick et al., 2017) noted that they found SGD to show limited improvement in a general search task, however increased improvement when searching for partially occluded targets. SAR has also been used to produce a similar blink effect to SGD (Volmer et al., 2018). Another way to reduce the annoyance caused by flicker is to utilise flicker at such a frequency that it is only perceivable in the periphery, due to the non-uniform ability of the retina to perceive such an effect (Mateescu and Bajić, 2014; Cheadle et al., 2011). Alternatively, the flicker can be rapidly reduced from a high frequency intense flicker to a lower frequency, subtle one (Waldner et al., 2014; Waldin et al., 2017). One of the few temporal cue alternatives to flicker that has been demonstrated is a gradual zoom designed to emulate an approaching object and trigger a shift in attention (Dorr et al., 2004; Grogorick et al., 2018). Alternatively, a wave propagating towards the target has been demonstrated in a VR simulation of AR (Renner and Pfeiffer, 2017; Yoshimura et al., 2019).

Inherent Saliency Cues. Finally, saliency of views has been manipulated to direct visual attention, recently being extended to use in AR. We define saliency cues as:

> *Any modulation that looks to affect the characteristics of a scene in order to adjust their saliency.*

Techniques have been developed with the intention of being used in AR under full control conditions, such as that provided by VSTHMDs (Mendez et al., 2010; Veas et al., 2011; Lu et al., 2012, 2013, 2014). These techniques are generally not tested in VSTHMDs with limited cameras, but rather are tested on monitors. Directly working in VSTHMDs, recent work has demonstrated the use of a saliency adjustment in the form of a blur to be used to reduced background noise and increase attention to a task (Koshi et al., 2019). The only work to be applied to, or demonstrated in, OSTHMDs looked to increase the contrast of a virtual screen (Google Glass) to increase readability (Ahn et al., 2018). Alternatively, SAR has been used to create inherent saliency cues, and is the most commonly demonstrated in practical/actual use. They have been used as simple luminance increases (Khan et al., 2005), as well as being used to generate blurs (Ueda et al., 2020) and colour modulation (Takimoto et al., 2019; Volmer et al., 2018). Notably, one of the primary goals of saliency modulation in AR has been to be subtle or unnoticeable (Veas et al., 2011), however when applied in practical scenarios, such as SAR or OSTHMD scenarios this has not been attained (Ueda et al., 2020).

Whilst using inherent saliency cues in AR presents a limited set of techniques, previously developed techniques for scene manipulation in image material can also be applied to VSTHMD settings. In fact, even the techniques proposed for VSTHMDs have been tested on screen-based image manipulations (Veas et al., 2011). As such we included image manipulations as possible techniques for visual guidance in AR. The most common form of these saliency adjustments is to change the colour of target objects (Kokui et al., 2013; Mateescu and Bajic, 2014; Azuma and Koike, 2018). This provides a simple and effective way of changing saliency however has problems when considering the inherent meaning applied to certain colours. Other common methods for saliency adjustment are focus adjustments using blur (Hata et al., 2016; Danieau et al., 2017), brightness adjustments (Pal and Roy, 2018; Shi and Sugimoto, 2014; Danieau et al., 2017; Waldner et al., 2014), and saturation reduction (Danieau et al., 2017). More holistic shifts of saliency that work on colour, saturation and intensity have also been shown. These methods utilise both addition of colour and brightness, as well as subtraction, with many running iteratively (Takimoto et al., 2015; Mechrez et al., 2019; Wong and Low, 2011; Hagiwara et al., 2011).

One concept that does not fit well within our space is the use of binocular rivalry to create visual guidance (Grogorick et al., 2019). Although developed for VR this method does have application to AR. Whilst it does utilise saliency variances to generate the guidance, the underlying principle of the guidance is not the saliency modulation but rather the visual effect incurred by it.

5.1.4 Goal/Overtness

Commonly covered as justification for non-geometric cues (Veas et al., 2011; Lu et al., 2014; McNamara et al., 2008; Bailey et al., 2009), the overtness of cues can be varied to make them more subtle. Overt cues using geometric cues have been shown to work (Renner and Pfeiffer, 2017; Volmer et al., 2018), however are noted for creating negative effects such as attention tunnelling, occlusions, and contributing to visual clutter (Veas et al., 2011; Lu et al., 2014). To counteract some of these effects researchers have proposed the use of subtler techniques (Veas et al., 2011; Lu et al., 2014; McNamara et al., 2008; Bailey et al., 2009). This notion of varied overtness is self-evident in the cue background research, and has also been previously identified by researchers looking into VR, and here in particular *cinematic virtual reality (CVR)* (Schmitz et al., 2020; Rothe et al., 2019). Whilst prior techniques have self-identified as subtle, unnoticeable, or subliminal, and frameworks/taxonomies have used various terms (subtle vs noticeable, overt, obtrusive), for the purposes of our design space we denoted the overarching scale as the *Overtness* which varies from *subtle* (the most extreme of which would be imperceivable and make no changes) to *overt* (the most extreme of which would be immediately apparent and unignorable).

5.1.5 Comparisons/Frameworks of Visual Guidance techniques

Several works have looked to compare some of the relevant styles of visual guidance. In VR and 3D Immersive environments various techniques have been evaluated with comparisons between overt and subtle techniques in the form of arrows and flicker respectively (Schmitz et al., 2020), as well as the comparisons between various unobtrusive techniques (Grogorick et al., 2018), namely: a coloured dot, flicker, zooming, and blur. A selection of obtrusive geometric methods has also been compared for restoring attention (Yoshimura et al., 2019). Dillman et al. (2018) presented a framework for visually guiding people to locations and on paths in AR based on techniques used in games, however, did not look at directly guiding attention within the field of view.

An area of research which has some similarities to the concepts covered here is investigations into guidance to content in cinematic visual experiences (Rothe et al., 2019).

This work can be considered the closest to our investigation in that it provides a taxonomy of guidance techniques for on-screen and off-screen CVR (Rothe et al., 2019). It covers many of the works citepd in this research, as it looked to summarise the possible techniques for guidance in CVR, covering both on-screen and off-screen, alongside diegetic, visual, forced rotation and haptic cues. Compared to the generic and wide encompassing approach taken by these works to provide exploration of possible means to provide guidance in CVR, we look to provide a precise and clear space that allows for explorations and development of means for providing on-screen visual guidance in AR. In this vain, subtle techniques for guidance in an immersive dome have been explored (Grogorick et al., 2018). However, this work only looked to compare a set of techniques, without considering the wider space, and varying levels of overtness.

The use of design spaces to enable effective and efficient development of new techniques and systems is well understood (Goel and Pirolli, 1992). However, whilst there have been comparisons between methods of compensation in VR (Rothe et al., 2019) and considerations of the styles of techniques (Grogorick et al., 2018) there is as of yet no design space to encompass and enhance research into visual guidance in AR that we are aware of. Therefore, we look to overcome this by presenting our design space for visual guidance, demonstrating where current techniques exist within this space. Utilising our space, we looked to explore the properties of it as they apply to OSTHMD, AR, and to our Computational Glasses.

5.1.6 Viewing the Space

By encoding the *Cues* used to visually guide as axes of a 3D space, encoding the *Overtness Goal* as the distance along any one axis to the origin, and finally, encoding the AR device as the point colour, we can visualise our design space for visual guidance in on-screen AR. See Figures 5.2 and 5.3 for visualisations of the design space. We used the colour scheme of red for OSTHMD designed techniques, blue for SAR, green for techniques designed to work in VSTHMDs, dark green for those shown in VR and should be able to transfer to VSTHMDs, and black for screen based ones. As there is no set measure for the overtness of techniques, we chose to set the overtness primarily based on the interpreted intent of the authors of the respective works, and secondarily, we adjusted based on subjective viewing of the

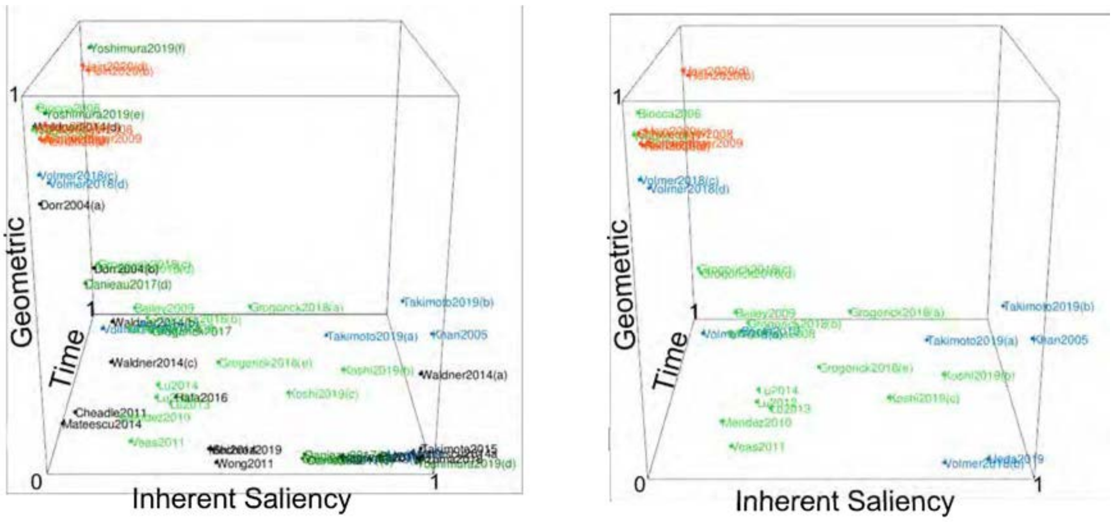


Figure 5.2: Visualisations of current research in our design space. Left: the 3D representation of our space with the overtness of each cue effect encoding on each axis and the AR type encoded as the colour. Red points represent OSTHMD solutions, blue represent SAR, green represent VSTHMD ones, dark green VR and black for screen-based solutions. Right: The same space with only the AR techniques shown. Note that a jitter has been applied to separate methods and we allow for duplicate methods published in separate papers or tested at differing levels.

presented output to generate variance. Therefore, the view of the space may look very different if thorough analysis of the various techniques presented was conducted and more precise values were assigned.

Looking at the space we can see that generally, techniques are constrained to one axis of the space with few works looking to combine styles of effects. Looking at the works that do, in many cases this is not as an intentional combination, but rather due to the process of applying modulations on one axis also creating an effect on another. One noteable effect that stands out from the rest is a modulation which works by applying a saliency adjusting geometric dot over the target area and then removing and returning it temporally (Grogorick et al., 2018). Looking further at the space we can see that there are no saliency modulations developed for OSTHMDs, or indeed subtle techniques for use in OSTHMDs. We can also see that the OSTHMD techniques by Hein et al. (2020) stand out from the rest as they looked to integrate multiple cues into their techniques due to a lack of doing so in the related literature. When looking at the techniques that induce saliency changes, most also have some effect on the geometry in the scene, often due to the masks used for the modulation area being primitives, rather than the exact shape of the target.

We can also view our space in a more simplified view (Figure 5.4). By segmenting the space into subtle and overt, and only considering the primary cue that is used for creating the guidance we can present a simplified 2D matrix of the space. Note that the aforementioned work by Hein et al. (2020) is ill suited to being represented

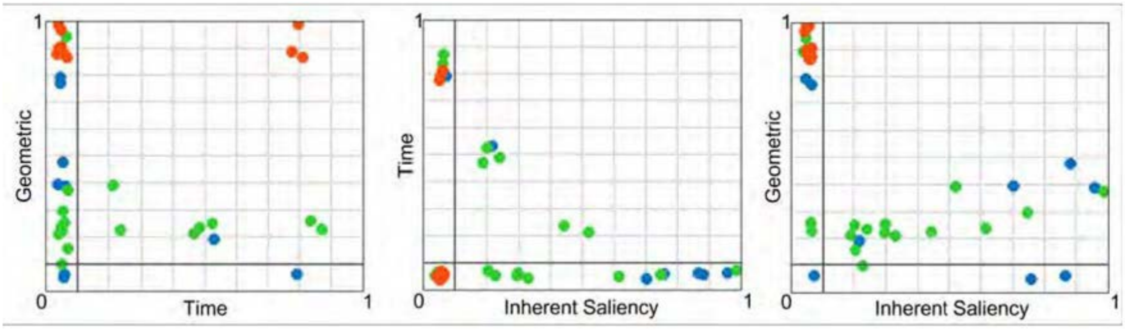


Figure 5.3: Visualisations of current research in our design space. Axis pairs of the space with only the AR techniques shown for easier reading. Red points represent OSTHMD solutions, blue represent SAR, and green represent VSTHMD ones. Points below or right of the inset black lines have no effect using the cue on the given axis. Note that a jitter has been applied to separate methods and we allow for duplicate methods published in separate papers or tested at differing levels.

by this space, as would be other works that look to use multiple cues at similar overtness. Looking at this view of the space we can see that no subtle techniques look to work primarily with geometric cues. We also see a heavy tendency towards overt cues working on Saliency, however Temporal cues have a much more even distribution between subtle and overt. It is also noteable that only Booth et al. (2013) have demonstrated a subtle temporal cue in SAR and again note the lack of OSTHMD subtle techniques. In fact, OSTHMDs have only been directly used for overt geometric guidance.

We again stress that the locations of methods within our space are decided based on the original papers with subjective variances, and if we consider results which have indicated the subtly proclaimed by works may not always be achievable (Grogorick et al., 2017, 2018), then distribution of works in our space may look different.

5.2 Saliency Guidance

One area of the space that was lacking was an inherent saliency modulation technique applicable for OSTHMDs. Whilst there are prior methods for adjusting saliency as detailed in the following, none are applicable to our implementation of Computational Glasses, and inherent saliency-based methods look to overcome the focus of other styles to draw maximum attention. This maximum attention drawing often comes at costs in scene understanding as the techniques often hide or distort other important scene elements.

As we wanted to utilise saliency modulation as a means to provide visual guidance in Computational Glasses due to its indicated potential (Veas et al., 2011), we needed to create our own technique that conformed to our constraints. We also looked to ensure we created a saliency modulation technique that could adhere to the notion of providing subtle guidance.

Space			
	Primary Cue		
Overtness	Geometric	Temporal	Saliency
Overt	Dorr2004(a), Waldner2014(d), Volmer2018(c), Volmer2018(d), Yoshimura2019(a), Yoshimura2019(b), Yoshimura2019(c), Yoshimura2019(e), Yoshimura2019(f), Biocca2006, Wagner2017, Biocca2007, Schwerdtger2008, Schwerdtger2009, Hein2020(a), Hein2020(b), Hein2020(c), Hein2020(d), Hein2020(e), Renner2017(a), Renner2017(b)	Dorr2004(b), Waldner2014(b), Volmer2018(a), Danieau2017(d), Grogorick2018(c), Grogorick2018(d)	Waldner2014(a), Kokui2013, Mateescu2014a, Azuma2018, Pal2018, Takimoto2015, Hagiwara2011, Volmer2018(b), Khan2005, Takimoto2019(a), Takimoto2019(b), Ueda2019, Yoshimura2019(d), Danieau2017(a), Danieau2017(b), Danieau2017(c), Koshi2019(a), Koshi2019(b), Koshi2019(c)
Subtle		Waldner2014(c), Mateescu2014, Cheadle2011, Booth2013, Grogorick2017, McNamara2008, Bailey2009, Grogorick2018(b)	Hata2016, Shi2014, Mechrez2019, Wong2011, Grogorick2018(a), Grogorick2018(e), Veas2011, Mendez2010, Lu2012, Lu2013, Lu2014

Figure 5.4: A simplified view of our design space with overtness split into overt and subtle. We again use the same colour encoding to denote the device developed for. Red areas represent OSTHMD solutions, blue represent SAR, green represent VSTHMD ones, dark green VR and black for screen-based solutions.

When looking to create our own saliency modulation we first have to cover the basis of using saliency for visual guidance, prior works we can build on for creating our own techniques, and the prior techniques proposed for use in AR and why they are not applicable to our OSTHMD based Computational Glasses.

Visual Saliency: As previously discussed in Chapter 2 (2.1.1), psychology has demonstrated the role that the properties of a scene have on the attention of a viewer in the forms bottom-up (inherent features of the scene) and top-down (conscious effort and goals) saliency.

Koch and Ullman proposed a biological model in which the various features being processed in parallel are combined into a singular map that shows the impact of the individual features as a saliency map (Koch and Ullman, 1987). Computer vision techniques have been used to model and compute saliency maps to better understand and utilise saliency. Itti et al. (1998) proposed one of the first and most well-known maps that builds on the biologically plausible model of Koch and Ullman (1987) but since then other saliency maps have been proposed (Hou and Zhang, 2007; Parkhurst et al., 2002; Torralba et al., 2006). More recent approaches integrated gaze maps into their models for computing saliency maps with gaze maps showing how much attention will be placed on various areas in images and videos. Examples include SAM-ResNet (Cornia et al., 2018) and MSI-Net (Kroner et al., 2019) for images or networks for videos (Wang et al., 2018). These approaches were tested against datasets of real users' attention, such as CAT2000 (Borji and Itti,

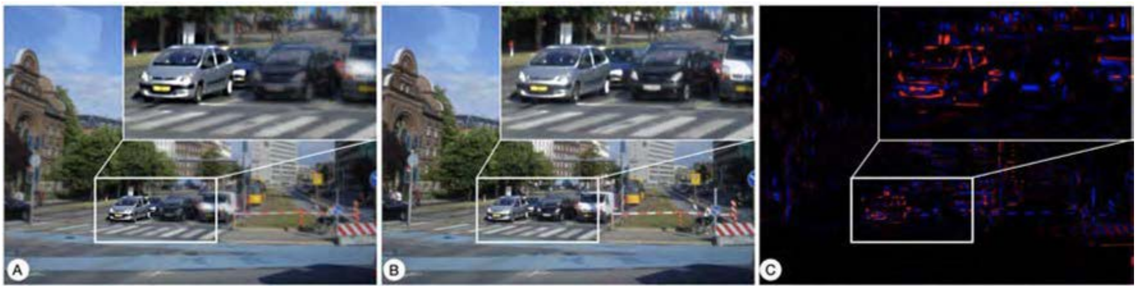


Figure 5.5: Differences in saliency modulation methods not considering additive only properties of OSTHMDs exemplified by our implementation of the algorithm by (Mendez et al., 2010) (insets for details): (A) Full colour control allows increasing the saliency of some image elements (e.g., silver car) while decreasing the saliency of other elements (e.g., black car). (B) Modulation using only the additive components fails to demonstrate the full saliency modulation. (C) Visualisation of required additive (red) and subtractive (blue) modulation necessary to highlight the silver car showing the majority of image areas are modulated by subtraction not possible in OSTHMDs.

2015). All these models for saliency and visual attention are bottom-up approaches and do not consider conscious influence.

Given the relevance of saliency to visual attention and our ideas for providing guidance using Computational Glasses, a core idea of our work is to explore visual guidance by modulating the saliency of the physical environment using OSTHMDs normally used for AR.

Visual Saliency Modulation: So far, existing techniques for saliency modulation that we looked to draw on have only been demonstrated on video images that have been displayed on standard monitors.

Existing works looked into different parameters for modulating the saliency of a scene. For example, Kokui et al. (2013) and Takimoto et al. (2015) applied colour shifts based on saliency maps whilst other approaches looked at modulating the spatial frequency (Takimoto et al., 2017) or texture power maps (Su et al., 2005). We have also seen approaches combining colour and intensity modulation using a saliency maps (Hagiwara et al., 2011; Shi and Sugimoto, 2014), the introduction of subtle blur effects to modulate the visual saliency (Hata et al., 2016), or the usage of genetic algorithms (Pal and Roy, 2018). More recent approaches combine different parameters affecting visual saliency including, blurring, intensity, saturation, and contrast (Suzuki and Nakada, 2018).

Visual saliency modulation has also been considered for AR. Research has also used saliency analysis to place visual information, so it does not distract the user from important scene elements (Grasset et al., 2012), but most existing works focus again on guiding the user. For example, the work by Lu et al. (2012, 2013, 2014) demonstrated how the visual saliency of AR content could be increased to aid users searching for it. However, the idea here was to have virtual objects stand out even more which is relatively simple but not to change the saliency of the real world through AR as aimed at in our work. Recent work has also looked at using SAR systems to adjust saliency (Takimoto et al., 2019). For instance, Ueda et al. (2020) proposed an approach using projectors and synchronised shutter glasses showing that they can both blur and focus scene elements selectively to guide attention.

The most related work to ours is that by Veas et al. (2011) and Mendez et al. (2010) who looked at achieving subtle visual guidance using saliency modulation on images and videos. However, whilst their general aim did align with ours in modulating the saliency of the physical environment, they eventually only demonstrated manipulation of video footage falling short of any AR scenario and as such putting it closer to prior work in visual attention guidance.

While providing interesting insight into human perception, these works are far from applicable for envisioned real-world application scenarios. More importantly, these approaches would not work on OSTHMDs because of the differences in how the overlay is displayed to the users (e.g., optically blending virtual overlays, inability to subtract light). In OSTHMDs, we can only show information by adding light (additive) while interactive state-of-the-art saliency modulation techniques shown on videos (e.g., (Veas et al., 2011)) always assume full light control (additive and subtractive). The consequence of not being able to subtract light diminishes the effect when using the original approaches (See Figure 5.5).

To conclude, there is generally a good understanding of the relevance of bottom-up saliency and its importance for scene understanding. This has been shown with works modulating the visual saliency in images or videos displayed on a screen. Studies exploring practical usage of saliency modulation or even saliency modulation of the real world to guide users are mainly in a conceptual stage or use projectors to change the appearance of the real world (SAR). For our work, we needed a saliency modulation of the real-world using OSTHMDs.

5.2.1 Saliency Modulation via OSTHMDs

As stated before, existing approaches for saliency modulation do not consider OSTHMDs but mainly manipulated video images giving full control of each pixel. Instead, our approach interactively computes a modulation that when displayed in the OSTHMD aligns with the real-world view and modulates its saliency (See Figure 5.6 Left).

While even naïve overlays, such as a rendered frame, affect the saliency of a scene, we are looking for a more precise effect that uses all parameters affecting visual saliency such as colour, orientation, size, motion, and depth (Wolfe and Horowitz, 2004). But, because we can only modulate the light entering the eye from the environment, we cannot change some other aspects (e.g., size of an object or motion). Related works usually apply a selection of changes in contrast (blur or sharpening), (de)saturation, or changing the lightness of objects (Veas et al., 2011). Whilst some researchers have looked to directly change the hue of an object entirely, we chose not to utilise this effect as it can cause confusions and clutter (Mateescu

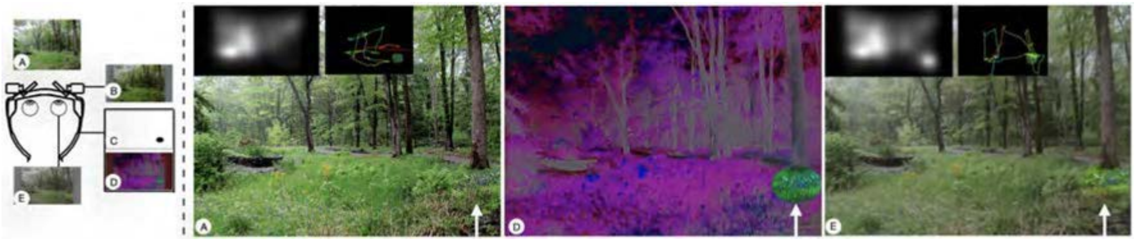


Figure 5.6: Left: Our approach to providing visual guidance via saliency modulation in Computational Glasses. Right: Example output from our approach.

and Bajić, 2014; Nguyen et al., 2013).

Our implementation used an algorithm combining a set of components based on those described to work in the literature, that can influence saliency whilst considering that they can only be applied by adding light. In the following we assume that the reader is familiar with the conceptual properties of colour spaces such as RGB, HSV, and Lab.

The earlier work by (Veas et al., 2011) utilised the method of (Mendez et al., 2010) which only looked to adjust lightness L then shift opponent colours, RG, BY to effect conspicuities in Lab space thereby simply creating a contrast shift. Unfortunately, our experiments showed that their algorithm strongly relies on darkening image regions that are not of interest and consequently fails to achieve the desire effect in OSTHMDs where this is not possible (see Figure 5.5). As our modulation has to be purely additive, we modulated several components in our algorithm. To this end we included saturation increasing and decreasing, a contrast increasing and decreasing, blurring, and sharpening. Each of these components has been used previously for modulating saliency within images shown on normal screens (Suzuki and Nakada, 2018; Lu et al., 2014; Hata et al., 2016; Takimoto et al., 2015), combining them to achieve saliency modulation when reducing environment light per pixel is not possible (e.g., OSTHMDs or Heads-up displays) is novel. We are basically replacing saliency modulations that mainly adjust lightness by utilising other factors that affect visual saliency.

Within our implementation, for each component we defined a parameter p used to adjust the degree of modulation. For the saturation component of our algorithm, we set the S component of the colour in HSV colour space to max for increasing saturation, and min for decreasing. C_{hsv} = Colour in HSV, C_n = component N of colour C.

$$C_{hsv} = (C_h, C_{smod}, C_v), C_{smod} = \begin{cases} 1.0 & Increase \\ 0.0 & Decrease \end{cases} \tag{5.1}$$

For contrast we used a sigmoidal contrast function with $\beta = 10$ and $\alpha = 0.5$ to increase contrast. To decrease it we then used the inverse of the function with the same parameters. To simplify computation we also simplified the formula based on our input beta and alpha.

$$C_{rgb} = \left(\frac{1.0}{(1.0 + e^{10.0(0.5 - C_{rgb})}} - 0.00669 \right) * 1.00925 \tag{5.2}$$

$$C_{rgb} = 0.5 - \frac{\log\left(\frac{1.00925}{C_{rgb}+0.00669} - 1.0\right)}{10.0} \tag{5.3}$$

To adjust the output of saturation and contrast modulations to the parameter level p we subtracted the original value from the modified to get a Δ vector which we then scaled by p. C_M = Modified Colour, C_O = Original Colour, p = input parameter defining level of adjustment.

$$C_M = C_O + (C_M - C_O) * p \tag{5.4}$$

To blur the image, we applied a common Gaussian blur with $\sigma = p$, and in order to sharpen we use an unsharp filter scaled by p.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, C_{rgb} = G(x,y) * C_{rgb} \tag{5.5}$$

$$C_{rgb} = p * \begin{bmatrix} 0.0 & -1.0 & 0.0 \\ -1.0 & 5.0 & -1.0 \\ 0.0 & -1.0 & 0.0 \end{bmatrix} * C_{rgb} \tag{5.6}$$

As all components can still produce negative values, which will have no effect within OSTHMDs but will be relevant when simulating the effect on normal screens, we also took Δs for the blur and sharpening components and clamp all Δs to ensure they contain no negative values before being added to the original image for simulation or displayed on the OSTHMDs.

Overall, our implementation utilises GLSL and the performance is only limited by the camera update rate (41fps). In fact, first experiments showed sufficient performance even on mobile hardware with a less capable GPU. This compares to 15fps by Mendez et al. (2010) which is still too slow for AR while many existing approaches are even slower and running in an offline process.

5.3 Apparatus

We implemented our approach for real-world saliency modulation using Computational Glasses in several prototypes. These were created for use in our later user studies.

5.3.1 Stereoscopic prototype:

In order to enable participants to perceive modulations directly through the Computational Glasses, we built a non-mobile stereoscopic prototype. This prototype utilised an Epson Moverio BT-300 and integrated a 50/50 beamsplitter in front of each eye (Figure 5.7 Left). Two Point Grey Blackfly cameras were used as scene cameras. We decided for the BT-300 because its OLED display is known to cover almost the entire RGB colour space unlike devices such as the MS HoloLens. This

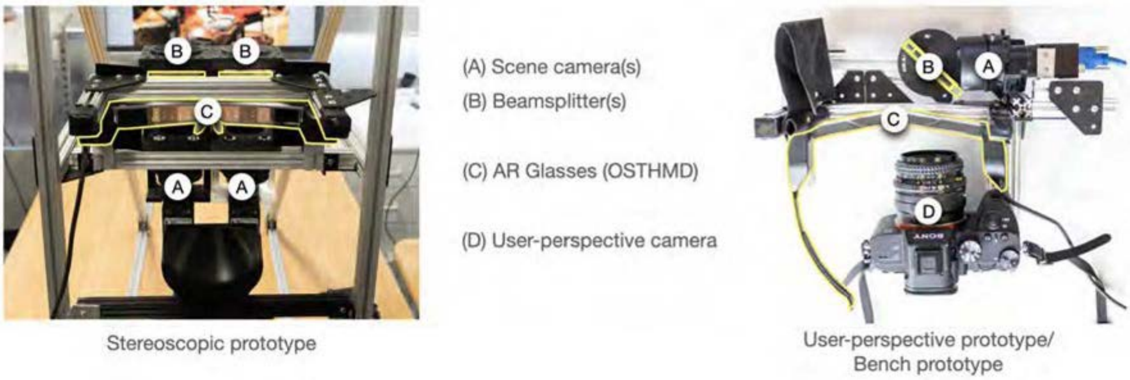


Figure 5.7: Two of the prototypes developed for this project. (Left) Stereoscopic prototype with the main components highlighted. (Right) Bench prototype (mono) with user perspective camera capturing the image through the Computational Glasses also showing the beamsplitter and scene camera. This allows for a controlled study environment.

prototype was mounted in a stabilising frame and a chin rest was included to enable participants to maintain a comfortable head positioning. We integrated the saliency modulation approach described in the previous section into our stereoscopic prototype and were able to successfully modulate saliency. In order to evaluate the efficacy of our saliency modulation approach, we added a Pupil Labs eye tracker to be able to track the user's gaze. While achieving good visual results, upon initial testing we found that due to our specific setup (e.g., additional optics) and the nature of using a retrofitted eye tracker we were unable to ensure a sufficiently accurate eye-tracking as needed for our planned studies. We considered using alternative OSTHMDs that include inbuilt eye tracking, e.g., HoloLens 2 or Magic Leap One. Unfortunately, their integrated displays have a very limited colour space and suffer from low colour correctness and chromatic aberrations[1] which became immediately obvious in our tests. Consequently, using those devices was not an option either, as they would not allow to display correct colours as required for the environment modulation.

5.3.2 User-perspective bench prototype:

Our second prototype addressed the identified shortcomings of the earlier prototype, mainly the lack of reliable eye tracking (Please note that this is only needed for performing user studies using eye tracking). This prototype again used a modified Epson Moverio BT-300 as an OSTHMD and integrated a 50/50 beamsplitter to reflect a portion of the incoming environment light towards the scene camera (Point Grey Blackfly). However, instead of the user looking directly through the setup, we placed a camera at the position of the user's eye (we used a Sony A7M3) as a user-perspective camera. The output of the camera can either be saved or directly be further processed by accessing its output via HDMI (Figure 5.7 Right). To address

[1]https://kguttag.com/2020/07/08/hololens-2-display-evaluation-part-2-comparison-to-hololens-1/

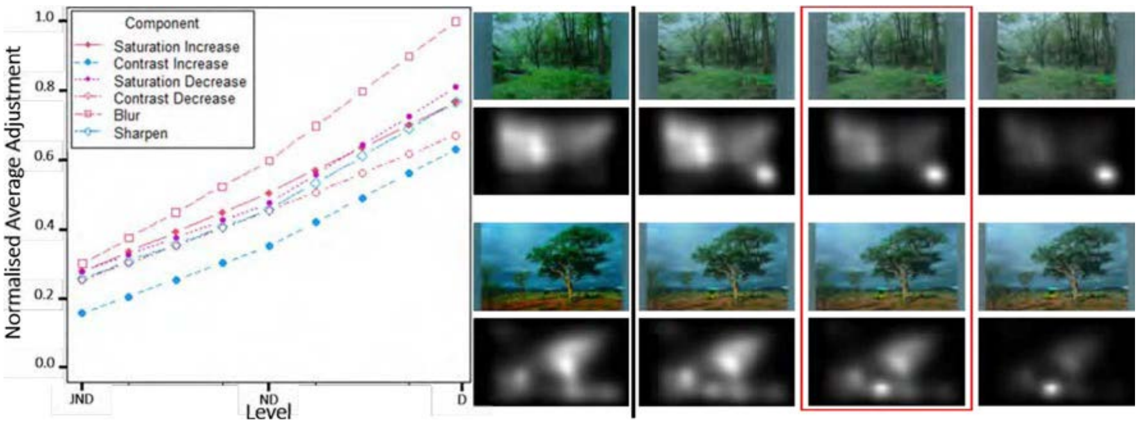


Figure 5.8: Selection of the appropriate modification level for the primary study. Left: Average selected adjustment levels participants rated as just noticeable (JND), notable (ND), and distracting (D). We define several steps between each rated level and interpolate the corresponding adjustment level. Right: Appearance and saliency of an image (left) and results (right) that provide a good balance between colour and saliency change that were selected for the primary study.

the lack of eye tracking in our first prototype we streamed the actual view of the user-perspective camera through the OSTHMD to a VR head-mounted display with integrated eye tracking (HTC Vive Pro Eye). VR has often been used in the past to simulate AR interfaces (Ragan et al., 2009) but here we used it to show the output from actual OSTHMD Computational Glasses instead of a VR simulation of AR. In a similar fashion, user-perspective cameras have been commonly used for evaluating optical devices (Chakravarthula et al., 2020; Itoh et al., 2015).

The advantage of this approach is that we could utilise the high-quality eye tracking of the used immersive VR headset while also guaranteeing a good calibration (because we calibrate for the camera) and guaranteeing the same quality of calibration for all participants. This eliminates a large number of possible confounding variables. In particular the calibration of OSTHMDs for actual user's eyes often proves problematic as calibrations are very individual and verification of the quality is hard, introducing at least one confounding variable. Similar prototypes are commonly used for that reason (Langlotz et al., 2018; Chakravarthula et al., 2018; Itoh et al., 2017; Gabbard et al., 2000).

5.4 Parameter setting for Subtle Saliency

Validating saliency modulation is traditionally done by analysing gaze data. However, as for most other existing approaches of saliency modulation not aimed at OSTHMDs (Mendez et al., 2010; Veas et al., 2011), we needed to first identify a suitable level of saliency modulation, so we explored the parameter space of our algorithm and looked at the effect according to study participants.

In the following, we describe this study with the main purpose of finding suitable parameters. For this study, we looked at normalising the effect of each individual component affecting saliency (i.e., saturation, contrast, blur and sharpening) as each could work drastically differently. We choose to do this based on users' responses and so devised a study where participants set levels for each component, looking for the adjustment levels where the differences first became noticeable (Just Noticeable Difference, JND), once it was having a notable effect (Notable Difference, ND) and the point at which it became distracting to the user (Distracting Difference, D). This provides adjustment levels for our algorithm at which each component would start to have an effect, where the effect was clearly working and where it was over-tuned and causing a detrimental effect. We then linearly combined these results to create one modulation parameter. One could argue that this linear combination is not representing the complex interactions between the components. However, building better models to describe the complex interactions would be a research topic on its own and is beyond the scope of this work.

Design: We designed an experiment to test the effect of manipulating each component (i.e., saturation, contrast, blur, and sharpening) within a certain min-max range [0-1]. Our goal was to identify a parameter range for each component representing three different levels: Just Noticeable, Notable, and Distracting. We did this for each component separately. The study design was approved within the regulations of the human ethics committee of the University of Otago.

Apparatus: For this study, we seated users in front of a monitor, where the image modified by the parameter was shown. We placed a dial, that was used to adjust the level of parameters, and a numpad in front of them. The number pad had labels placed on the relevant keys for running the study (reset, set, none) and the rest of the keys were covered with a single cover to prevent their use.

Task: The task for this study was to adjust a dial until the effect reached a given level (JND, ND, or D) then press the 'set' key, or the 'none' key if they could not find a value that they believed met the definition.

Procedure: After signing a consent form and completing a demographic questionnaire collecting information on age, gender, and vision impairments (colour and refractive), each participant was seated in front of the monitor and was informed about the study procedure. They were also given the instructions and relevant definitions as text which they were asked to read. Once the participant understood the procedure and had no questions the study was started.

For the actual study the participant was shown a random image with a random component selected. They were asked to complete the task. The image was then removed and they were asked to reset the dial. This was then repeated for level. The user could reset the values for all levels on that image at any stage. Each participant was asked to set values for each component twice.

Participants: For this study we recruited 10 participants (6 male, 4 female, 20-51 years old $\overline{x} = 29.5$). We excluded participants having colour vision deficiency or vision not corrected to normal.

Outcome: This study allowed us to empirically determine the range of values under which the effects of each component can be compared/normalised, and we can expect behaviour to be similar so we could use it to select a singular, combined level across all parameters. To define our final, singular level, we took the mean JND and ND for each parameter (See Fig. 5.8 (Left)) and linearly interpolated between the levels. We used 5 steps. Using our bench prototype, we took images of each of our test images and simulated the levels of modulation on them. We then looked at the changes in the saliency maps and colour shifts and selected the level of our parameters providing a strong response in the corresponding saliency maps (See Fig. 5.8 (Right)) and used them in our primary study .

Parameter Levels Transfer to OSTHMD: While we estimated our parameters on a simulation, this provided us with a saliency map for each degree of noticeability. Meaning, when applied on a different display in the same additive manner, we expected that similar saliency maps will be rated on a similar level of noticeability, allowing scaling of the displayed output to match the desired results.

Due to health and safety restrictions for COVID-19 we were limited in our ability to run studies. As such, rather than run further studies to set levels in OSTHMD we choose to utilise these values for our OSTHMD settings. In order to ensure that our values could be as accurate as possible on the OSTHMD we compared the saliency maps from the simulated images on the screen to images taken through the bench prototype. We looked to scale the output on the OSTHMD due to the different contrast produced, until we found the closest output. A similar adjustment of the output would be needed to create levels for use in a direct viewing experience as the contrast perceived by the use would be greatly different to that of the camera.

5.5 Study 1: Subtle Saliency Efficacy

Before exploring our space, we looked to investigate the efficacy of our saliency modulation technique and confirm that it would work.

Hypotheses: Primarily interested in the effect saliency modulation has on gaze patterns and on the subjective evaluations of the images, we formulated three hypotheses:

- H5.1: Real-world saliency modulation via Computational Glasses alters gaze pattern when compared to an unmodified scene with respect to time until the first fixation on a target area and number of participants who fixated.

- H5.2: Real-world saliency modulation via Computational Glasses is not rated as less natural, more obtrusive, or lower quality compared to unmodulated scenes.
- H5.3: Real-world saliency modulation via Computational Glasses is less visually distracting when compared to augmenting geometrical primitives (circles).

Design: We designed a within-subject study to investigate the effect of saliency modulation using our approach. Participants observed views of the images from the image dataset in each of two initial conditions (condition 1:'unmodulated' and condition 2: 'saliency modulation via Computational Glasses') in randomised order.

For each image, we collected the user gaze data and asked participants to rate the image's *Naturalness*, *Obtrusion*, and *Quality* on a Likert-like scale from 1-7. In these conditions, we evaluated the *time until the first fixation* on a target area, the *explored area of the image*, and the *number of participants who fixated* on said area, and the answers to the questionnaires.

After participants viewed and rated all images in both conditions, they were again shown views of the images in a third condition (condition 3: 'circles'). In this condition, the images had a circle overlay displayed on the Computational Glasses around the highlighted area, as an example of traditional guidance in AR. We included this condition to explore how our modulation affects the user's gaze behaviour compared to an AR overlay (showing a circle to highlight an area of interest). In our preliminary tests, we observed that this condition had a strong anchoring effect on participants, affecting their gaze patterns whenever they were exposed to the scene again. We thus opted to show this condition last instead of fully randomising the order in which the participants experienced the three conditions to avoid biasing the results. In this condition, we only collected the participants' gaze data and evaluated the *explored area of the image*.

Our independent variable was the modulation state of the image with the three conditions: 1) "Unmodulated", 2) "Saliency modulation via Computational Glasses", and 3) "Circles". See Figure 5.10 for examples of each condition.

When asking about the naturalness and obtrusion we provided definitions for each word that steered participants towards the requisite measures. This was due to the variability of the definition of such words and to prevent participants from taking drastically different views. We, however, left quality undefined to prevent biasing towards certain aspects of the images, preventing reporting on others. Our provided definition for naturalness was: "having undergone little or no processing" and for obtrusion was: "noticeable or prominent in an unwelcome or unwanted way". The study design was approved within the regulations of the human ethics committee of the University of Otago.

Apparatus: For this study, we faced the challenge of recording gaze data from eye trackers. Due to the limitations of the initial stereoscopic prototype in this aspect we opted to use our bench prototype showing the actual view through our OSTHMD

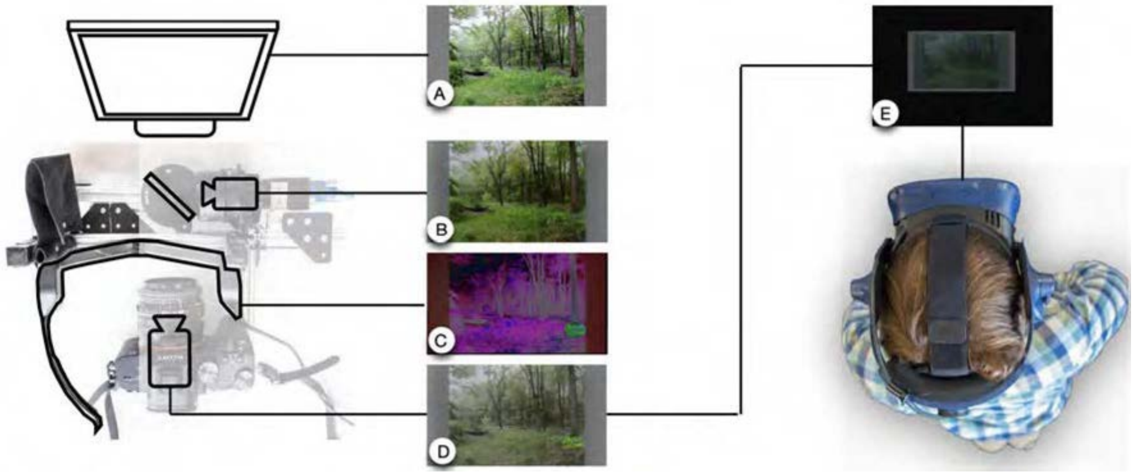


Figure 5.9: Study apparatus: (A) Scene displayed on a screen representing the real environment is captured by the scene camera (B). The computed saliency modulation (C) is then displayed on the Computational Glasses. The combined view (D) captured by the user-perspective camera (representing the human eye) is finally displayed in the virtual environment (E).

prototype in VR (See Figure 5.9). This enables us to utilise eye-tracking whilst providing a more controlled environment where we can overcome the confounding variables such as the eye display calibration quality.

For the VR environment, we created an unlit virtual room with black walls into which the user was placed. They then had a virtual screen placed in front of them that covered a 40^o angle. The screen was always placed directly in front of the user and maintained its visual position throughout the study described later. On that screen, we showed the camera feed as captured through the bench prototype. Thus, this system combined the visual results from an OSTHMD with the quality of dedicated eye tracking from an off-the-shelf VR system commonly used in research (See Figure 5.9). The VR environment provides a completely controlled setting were we can ensure all participants are exposed to the same conditions increasing internal validity.

Task: The task for this study was to simply look at a series of images. To encourage viewing of the images, and garner further insights we also asked the participants to rate the images on our metrics.

Participants: We recruited 20 participants (10 female, 10 male, age ranging from 19 to 47, $\overline{x} = 25.2$) from students at the university. All recruited participants completed the study according to the procedure above. All participants had normal vision or corrected to normal via contact lenses. Eye tracking was verified to an $\overline{x} = 0.8^o$ and $\theta = 0.36^o$.

Figure 5.10: An image from our dataset in both primary conditions; unmodulated (A), and Saliency modulation via Computational Glasses (B), as well as the secondary condition of circles (C). Overlaid is gaze data from a participant time-coded from red (start) to blue (end).

Procedure: Given the context of a global pandemic, we had to take extra precautions regarding health. Safety procedures for the study were following institutional and governmental guidelines for COVID-19 safety. As such a distance of >2m was maintained between participants and operator, participants were screened for symptoms, and sterilisation of equipment was used. Before entering the study, each participant completed the screening/contact tracing form, read the supplied information sheet, and signed a consent form. We also asked them to provide their information in a demographic survey collecting information on age, gender, vision impairments, and previous experience with VR (e.g., issues with simulator or motion sickness).

Once completed, we introduced the participants to the use of the HMD, provided an overview of the study procedure and questionnaires, and provide definitions for naturalness and obtrusion. After the participants put on the HMD, we calibrated the integrated eye tracker using the supplied Vive SRanipal calibration. We verified the calibration to be at least within 1.5^o average angular accuracy but often saw values $< 1^o$ across the measured area. This calibration was repeated after every 10 images throughout the study in case participants invalidated the calibration (e.g., by moving the HMD).

During the actual study, each participant was shown 3 different views (one for each condition) of each of 10 images for 5 seconds a piece. The images were shown in a randomised order and conditions in a semi-randomised order, as detailed in the design. No participant saw the same image in two conditions consecutively. Before showing a new image, we displayed a black screen with a white cross in the centre. Participants were instructed to look at the cross when it appeared. This was done to centre their gaze in the screen for each image. After each image we asked the participants to rate it on the Likert-like scale for naturalness, obtrusion, and quality. Answers were captured by the study conductor. We also asked for general feedback. Participants were gifted vouchers worth approximately \$13(USD) for their time.

Results: To determine whether participants focused on a target area, we implemented the IV-T fixation detection algorithm as described by Anneli Olsen (Olsen,

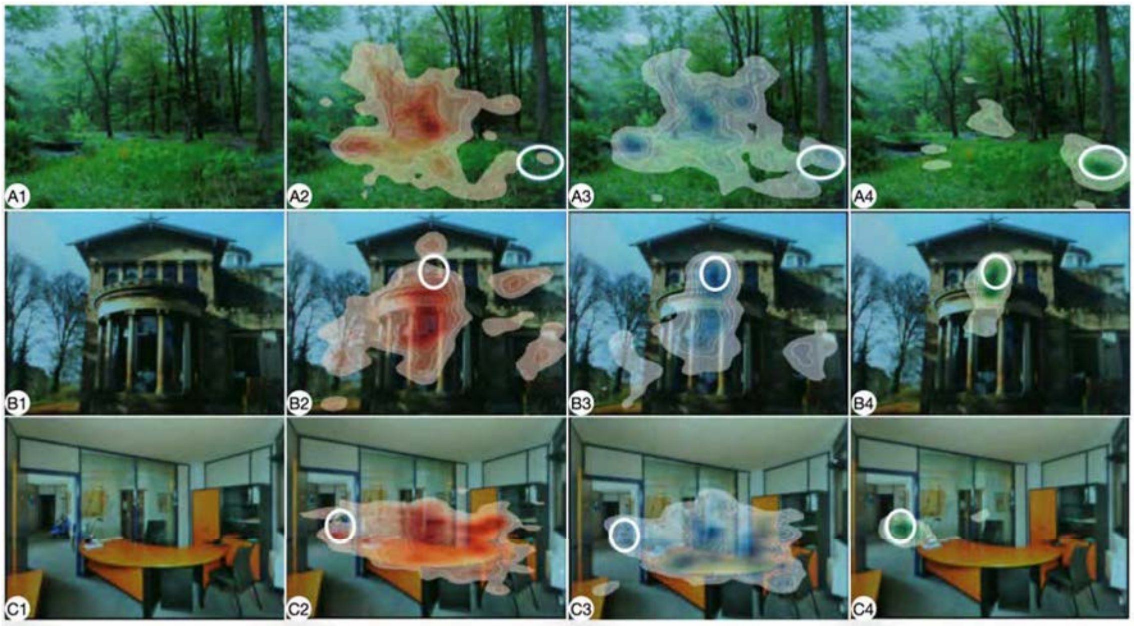


Figure 5.11: Resulting gaze heatmaps for 3 images. (A1, B1, C1) are the original images, (A2, B2, C2) are the heatmaps in the unmodified condition, (A3, B3, C3) are the heatmaps in the modified condition, (A4,B4,C4) are the heatmaps from the circle condition. Circles indicate the target area and have been overlaid only for the readers benefit.

2012). We identified that a participant fixated at an area of interest when at least one gaze-point associated with a fixation lay within the target area. An example of a participant's gaze data on an unmodified image is shown in Fig. 5.10(A), compared to the gaze data when modulated via the Computational Glasses (See Fig. 5.10(B)). We visually checked all recorded gaze patterns to detect possible errors in the recorded data. We exclude the gaze data of one participant as it exhibited large inconsistencies (e.g., consistent jumps between continuous gaze points). We checked normality of the collected data with the Shapiro-Wilk test and assumed significance at a $\alpha < 0.05$ level. We analysed normally distributed data with a paired one-sided t-test and used a Wilcoxon signed-rank test otherwise.

The results for the remaining 19 participants showed a significantly higher *number of detected fixations*(F) for the saliency modulated condition than the unmodulated condition using a McNemar test ($\chi^2 = 25.565$, $p < 0.001$) (See Fig. 5.12(a)). This is supported by our finding that we successfully attracted a higher F for all but one image, where the number of fixations went down from 4 to 2. This can also be observed in the heatmaps shown in Figure 5.11.

We also investigated whether the modulation prompted participants to look at a target faster if they did look at the area of interest by looking at the *time to first fixation* (TtFF). In our first analysis, we grouped by participants. When considering only image observations where participants had an actual fixation (*Cleaned* in Figure 5.12), participants fixated at the target area significantly faster in the saliency modulated condition than the unmodulated condition ($t(18) = 3.96, p <$

$0.001, d = 1.09$; CI 0.33-1.078). To compensate the effect of missing fixations in some image observations, Veas et al. (2011) assigned the maximum display time (in our cases 5s) for each image when not fixating(*All* in Figure 5.12). Applying this analysis shows the same result of a significantly faster TtFF when applying saliency modulated ($t(18) = 8.42, p < 0.001, d = 2.3$; CI 0.87-1.45) (See Fig. 5.12(b)).

In our second analysis, we assumed that TtFF is dominated by the observed image and consequently grouped by image. This evaluation is similar to that of Veas et al. (2011). The results violated the normality assumptions. We found that participants fixated onto the target area significantly faster in the modulated condition considering only observations with detected fixations (*Cleaned*) ($t(9) = 3.03, p = 1.4e{-2}, d = 0.81$; CI 0.17-1.19). Again, this effect was supported when assigning the maximum time for no detected fixations (*All*) ($t(9) = 4.36, p = 1.8e{-3}, d = 0.86$; CI 0.38-1.35)(See Figure 5.12(c)).

To investigate if participants perceived a difference in the images based on the Likert-like scales, we evaluated the difference using a paired Wilcoxon signed-rank test on the average scores per user under both the saliency modulated and unmodulated conditions. We found significant differences in all of our metrics; naturalness was significantly reduced ($Z = -3.530$, $p < 0.001$, $r = 0.75$) from a mean of 5.335 ($\sigma = 0.979$) to 4.305 ($\sigma = 0.750$), obtrusion was reduced ($Z = -3.723$, $p < 0.001$, $r = 0.79$) from a mean of 5.78 ($\sigma = 1.008$) to 4.58 ($\sigma = 0.797$), and quality was reduced ($Z = -3.530$, $p < 0.001$, $r = 0.75$) from a mean of 5.15 ($\sigma = 0.754$) to 4.48 ($\sigma = 0.556$) (See Figure 5.12(d)). We subsequently evaluated the unmodulated and modulated conditions for each image individually, using paired Wilcoxon tests. Here we found significant differences between the two conditions for the naturalness of $\frac{5}{10}$ images, obtrusion for $\frac{8}{10}$ of the images and quality for $\frac{3}{10}$.

To determine how much saliency modulation affects the participant's exploration of the images (*area of exploration* (AoE)) when compared against circle overlays, we compared the area covered by the heatmaps generated from the participant gaze data (See Fig. 5.11). Once again, we compare the AoE by averaging the data for each participant and each image. When averaging the generated heatmaps for each participant. Participants explored a much larger portion of the image in the saliency modulated condition (M=0.18, SD=0.02) than the circle condition (M=0.12, SD=0.02) ($t(9) = 6.484, p < 0.001, d = 0.61$; CI 0.036-0.078). We found differences also remained when we average the generated heatmaps for each image ($Z = -3.035, p = 0.002, r = 0.95$) (See Fig. 5.12(e)).

Discussion: For H5.1, our results show that saliency modulation via our algorithm on Computational Glasses not only increases the likelihood that participants look at a target area (F) but also do so faster (TtFF). We therefore can accept hypothesis H5.1.

However, we should point out that similar to existing works using saliency modulation outside AR (e.g., (Veas et al., 2011)) saliency modulation does not ensure fixations. Thus, not all participants fixated on the target. This could be due to the inherent saliency of different scenes, the limited time given to users to explore each

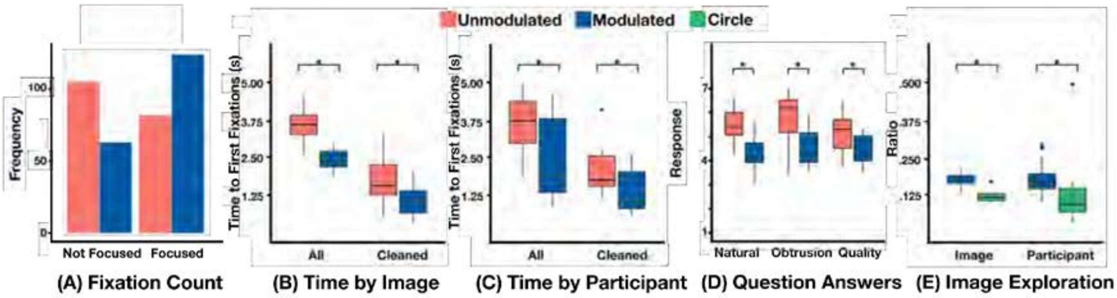


Figure 5.12: Quantitative results of our study for the Unmodulated (red), Modulated (blue), and the Circle (green) conditions. The F at the target area (a), TtFF averaged by image (b) and participant (c), Likert-like scale answers (d), and AoE averaged by image and participant (e). *Cleaned* includes only observations where participants fixated at the target area and *All* sets the fixation time for the remaining participants to the maximum observation time (5 seconds).

scene, and top-down influences while exploring an environment. For example, we observed that some participants tended to focus more on the centre of the scene, while others tried to explore as much as possible. For some scenes, F remained low, with only 5-6 participants focusing on it, while for others it was as high as 15 and improved by as little as 2 and as much as 10. In only one case, two participants less fixated onto the target area. We should also state again that images were shown exactly 5s and we cannot predict if or when participants would have gazed at the target after that. However, we would argue if participants gaze at the target after 5s the relevance for mentioned practical applications is relatively low.

For H5.2, based on the answers in our questionnaires, we must reject H5.2 as scores were significantly different. While this was the case over all images, it was not the case for each image which would indicate that there is potential here for further improvements. In particular, we saw a high number of images rated as significantly more obtrusive, whilst only three images were considered to have significantly reduced quality. Naturalness was split evenly. Although we do have a significant difference for all our metrics, all three show a mean reduction of only about one step on our scale and whilst our reduction is significant, we do not step past the neutral point. This implies that the differences created, whilst making the images less natural, more obtrusive, and lower quality, we do not expressly make them unnatural, obtrusive, or low quality.

We also should point out that we considered three images as challenging because the masked area was small, or the images were very salient before modulation. However, the results are mixed, and we would argue for more research also considering extreme images, but we see a trend that the amount of saliency modulation required to detract from an already very salient image can lead to artefacts that are considered more obtrusive.

From the feedback of the participants, we noted that they considered our saliency modulation to be part of the image. In contrast, they considered circles as an overlay or separate to the image, which is an interesting aspect generally supporting the concept of saliency modulation. Participants also noted that whilst the circles readily drew their attention and showed the areas of the image, which they liked for a short viewing of an image, they would not like to have this done constantly.

For H5.3, our aim for saliency modulation was also to enable a more natural exploration of the scene while directing the user's gaze to the area of interest. Thus, we compared how our modulation affects the gaze behaviour compared to a traditional circle overlay. We found that when presented with our modulation participants had a much larger AoE of the image, thus supporting our hypothesis H5.3. The presented circle overlay created a very strong anchoring effect, almost gluing the user's gaze to it. As such, although participants were given the instruction to explore the image for all conditions, the overlaid circle significantly hindered their exploration of the scene. These findings support our hypothesis H5.3. As our modulation did not create a similar anchoring effect, this could also explain why we did not detect fixations on the target areas from some participants. Overall, our saliency modulation using Computational Glasses seems preferable when the goal is to attract the user's attention while allowing natural scene exploration, and an AR overlay is preferable when the user's attention needs to be guided to a critical area.

5.6 Study 2: Space Exploration

Having demonstrated that Computational Glasses could provide subtle visual guidance, and having a saliency method that could work in OSTHMD and on our Computational Glasses, we looked to further explore our design space. We looked to understand how the various areas of it might apply to creating further visual guidance techniques for Computational Glasses, including expanding beyond spectral adjustments to those temporally based. We wanted to gain some insights into how techniques may be applied and where they might be best suited, as well as informing future research directions.

Hypotheses: As this was an explorative study into the relative effectiveness of each style we did not have overall hypothesis as to which techniques were best.

- RQ5.1: How might further visual guidance techniques for Computational Glasses be developed from within the design space?

Design: We designed a within-subject study to investigate and compare the effectiveness of guidance techniques at various levels of overtness.

For the study we evaluated the effectiveness of techniques using a set of real images which participants were asked to look at whilst their gaze was recorded. Our independent variables were the method of guidance provided (*None, Geometric, Temporal, Saliency*), and the degree to which the parameter is adjusted towards overtness (*25%, 50%, 75%, 100%*). We measured time to first fixation, the presence of fixations, and the area of image explored as our dependent variables to show effectiveness. We also collected subjective feedback as to the notability of any image modifications, and the overtness of the modifications on a 7-point semantically anchored scale in the case that they were noted.

Apparatus: For this study we again faced the challenge of recording accurate gaze data so showed the view through our bench prototype in VR once more. We kept the same VR environment we used in our initial efficacy study.

Images. In order to conduct our exploration of the design space we create a dataset of 80 images. These images were collected from around the campus and selected to ensure a diverse range of real-world scenarios in which visual guidance may be implemented. To this end we included images with and without high numbers of dynamic objects (cars and/or people), as well as both natural scenes and man-made structures.

Using a saliency estimation predictor (Cornia et al., 2018), we split the images into three levels of saliency (High, Medium, and Low). We then assigned each image to a desired level of modulation from 1 (minimal modulation) to 4 (maximum modulation). This resulted in a dataset of images comprised of 80 images divided into sets of 20 with an even distribution of inherent saliency and spread across a variety of real world scenarios.

Based on gaze maps of each image generated we also selected one object or area to be modulated. These objects were selected as places that were expected to see little attendance by viewers, but we would expect to see some.

Whilst datasets for exploring saliency and estimating user gaze already exist (Borji and Itti, 2015; Bylinskii et al., 2012), we choose to create a new dataset. This is due to the prior datasets not allowing for us to create a consistent dataset that meet our requirements for a large dataset of high-resolution natural images with an even spread of scenarios that we could allocate targets to various levels of saliency. We choose not to use videos in our dataset due to the introduction of confounding factors due to motion in videos, and a general inability to apply consistent constraints.

Methods: In order to explore the design space, we looked to explore the effect on techniques from various locations in the space. To this end we selected techniques from the previous literature that were; close to only one axis, therefore primarily designed to work in that axis, capable of being implemented on our Computational Glasses prototypes. As we wanted to explore various points regards *Goal*, rather than select and implement several techniques (of which some do not yet exist) we choose methods that we add a parameter to tune in *Goal*. Examples of each techniques subtly and overtly applied can be seen in Figure 5.13. We selected:

Geometric. For a geometric technique we chose to utilise the halo/circle technique (Waldner et al., 2014). This provides a readily implemented method that has been demonstrated to work in OSTHMD and is similar to several other methods such as tunnel effects for in view objects. It also avoids the placement issues present with arrows, and reduces the occlusions caused arrows and dots. We utilised a white circle which encompasses the target area.

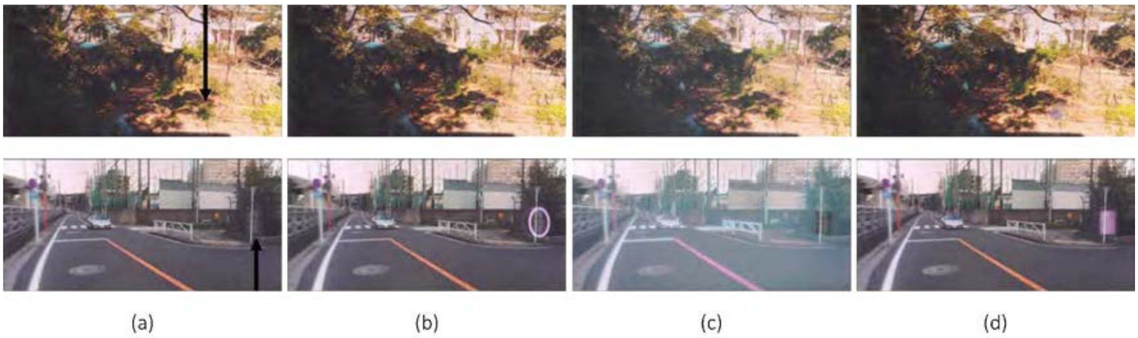


Figure 5.13: Examples of images used in our study. The top row shows an image modulated in the lowest condition for each technique. The bottom row shows an image modulated in the highest condition for each technique. Column (a) shows the original image, (b) the flicker condition at the midpoint of the largest intensity shift, (c) saliency modulation condition, and (d) the image in the geometric condition. Note that an arrow has been added to (a) to indicate the target area and was not visible to the participants.

Parameter: To vary the overtness of the geometry we adjusted the opacity of the circle and adjusted it between white and black.

Temporal. For a temporal technique we chose to utilise the flicker technique (Waldner et al., 2014). This represents a readily implemented temporal technique that can be implemented in OSTHMD. The advantage of using this technique is that, similar to the other techniques being tested, it does not require eye tracking input used by other temporal techniques (Bailey et al., 2009; McNamara et al., 2008). It also does not rely on achieving high frequency flicker at critical flicker frequency which is not able to be produced in the test apparatus, nor is consistent for all users and viewing environments. It works by briefly showing a high frequency flicker with high intensity shift before transitioning to a low frequency low intensity flicker. We modified the technique by adjusting the shape of the luminance adjustment to match the shape of the target, removing the introduction of a circular geometry.

Parameter: Based on the original presentation of the work we adjusted the time spent at the various flicker frequencies.

Inherent Saliency. For an inherent saliency technique we used the technique previously created for this purpose. The technique works additively adjusting contrast and saturation whilst also introducing additive sharpening and blur to increase saliency in a target area, and also reducing it everywhere else. We adjusted the technique by using direct outlines of the target areas rather than a blurred circle to reduce the geometrical impact of the technique.

Parameter: Originally, we utilised several components set to varying levels based on human responses to the effect of varying each component separately, looking for a minimally overt but effective level. The components were then combined into one parameter based on this. We chose not to apply an optimisation on these components as we were not looking to optimise any of the other techniques. Instead, we simply adjusted the levels of each component uniformly to adjust the overall overtness.

Task: We did not provide the participants with a formal task for the first part of the study as we wanted to evaluate methods in a neutral scenario. We informed participants that we wanted to record gaze patterns of people viewing a set of images some of which had been modified and encouraged them to explore the images.

Participants: We recruited 28 participants from around the campus (7 female, 21 male, mean age: 24.5, sd: 6.2). All 28 participants were able to calibrate the eye-tracker sufficiently and were recorded. All 28 participants took part in the interview.

Procedure: After signing a consent form and completing a demographic survey (age, gender, visual impairments) the participants put on the VR head set and were placed in the virtual room. The system was then calibrated for the participant using the provided calibration routine. The results of the calibration were verified and if the system reported errors greater than 1^o the calibration routine was rerun. This was repeated for a max of 5 times. If the system could not be calibrated properly for the participant, we did not record eye tracking data from them, however still showed them all images to ensure they had equal exposure to the techniques, before completing the interview questions. Once the calibration was verified the participant was shown a white cross at the centre of the virtual screen and instructed to focus on it whenever it was shown to them. After 3 seconds the cross was taken away and the participant was shown an image for 5 seconds. The participant was then shown a black screen with a question asking how obtrusive any modulations to the image were. They were shown a seven point semantically anchored scale with labels of *1: Very Subtle* and *7: Very Overt*. They were also given the option *0: No modulation*. After they answered the question, the image was then taken away and the cross shown again. This was repeated for all images in the dataset. Each image was shown modulated by one of the techniques at a given level. The level of modulation being applied to the images was set using a double latin squares to reduce ordering effects. The image and technique order within each level was randomised. After viewing all images the participant was given a break and a chance to remove the headset before being placed back in VR. The participant was then shown one of three images unmodified (either low, medium, or high saliency distribution) and all modified versions of the image at one level of modulation and asked for any further comments regarding the techniques. This was then repeated for all levels of modulation.

Results: When considering our results, all statistical tests were run as comparisons between every technique at a single level and the unmodulated condition. This allows us to evaluate the pertinent information to our investigation whilst reducing the number of tests being run. We calculated fixations using the IV-T fixation detection algorithm (Olsen, 2012). To determine if a fixation was within the target area we then tested if any fixation points lay within the area denoted by the geometrical primitive, allowing for an error of 1 degree. With this we determined if a user *fixated within the target area for each image* (F), and the *time to first fixation* (TtFF) if they did. In the case that the participant did not look at the target area we set TtFF to the maximum time (3 sec). Doing so assumes that in the best case for all images and techniques the participant would have looked at the target area immediately after the time shown. Whilst this is false this best case is equally applied across all conditions, and allows us to run statistical tests on a complete data set. We would expect actual values to show greater variance than this and that actual p-values to be less than those found. In order to investigate the *area of exploration* (AoE) we generated heatmaps based on the recorded gaze data. We also decide to look at the *duration of time* (D) spent fixated on the target area.

To compare F we calculated the percentage of times each participant fixated in the target area compared to the percentage of times they did not. We then completed a Friedmans test and Wilcoxons paired test as our post-hoc using a Holm-Bonferonni correction.

To compare TtFF we averaged the user's time to fixation across all images for each level and condition pairing and again completed a Friedmans test and a Wilcoxons post-hoc with Holm-Bonferonni correction.

To compare AoE we compared the ratio of area explored again using Friedmans and Wilcoxons tests.

When evaluating the overtness of the modulation as reported on the semantically anchored scale we first set any non-zero value to 1 and tested for noticeable *modulation* (M) by taking the percentage of times the participant rated the modulation as noticeable. We then tested using Friedmans and Wilcoxons.

We then compared how overt modulations were considered on the seven point semantically anchored scale (M7) again using the Friedmans and Wilcoxons.

Results at 25%: The results for conditions at 25% can be seen in Figure 5.14. Full sets of p-values from our statistical test are included in the appendix D.1.

When considering F at 25% we can see that only the Geometric guidance (mean = 0.54, sd = 0.22) was able to significantly increase the number of times participants looked at the target area compared to the None condition (p = 7.2e-05, mean = 0.16, sd = 0.16). Geometric guidance was also significantly better than Saliency (p = 0.00018, mean = 0.24, sd = 0.17) and Temporal (p = 0.00018, mean = 0.21, sd = 0.15). There were no significant effects between the other techniques or the baseline condition.

Looking at TtFF, Geometric (mean = 1.91, sd = 0.43) is again the only technique to have a statistically significant effect compared to None (p = 7.5e-08, mean = 2.74, sd = 0.3), Temporal (p = 3.0e-07, mean = 2.66, sd = 0.29) and Saliency (p = 4.5e-08, mean = 2.67, sd = 0.27).

The duration in which participants viewed the target area, D, only saw significant differences when comparing Geometric to the other conditions (None: p = 4.5e-08, Temporal: p = 4.5e-08, Saliency: p = 6.0e-08). Geometric held the gaze

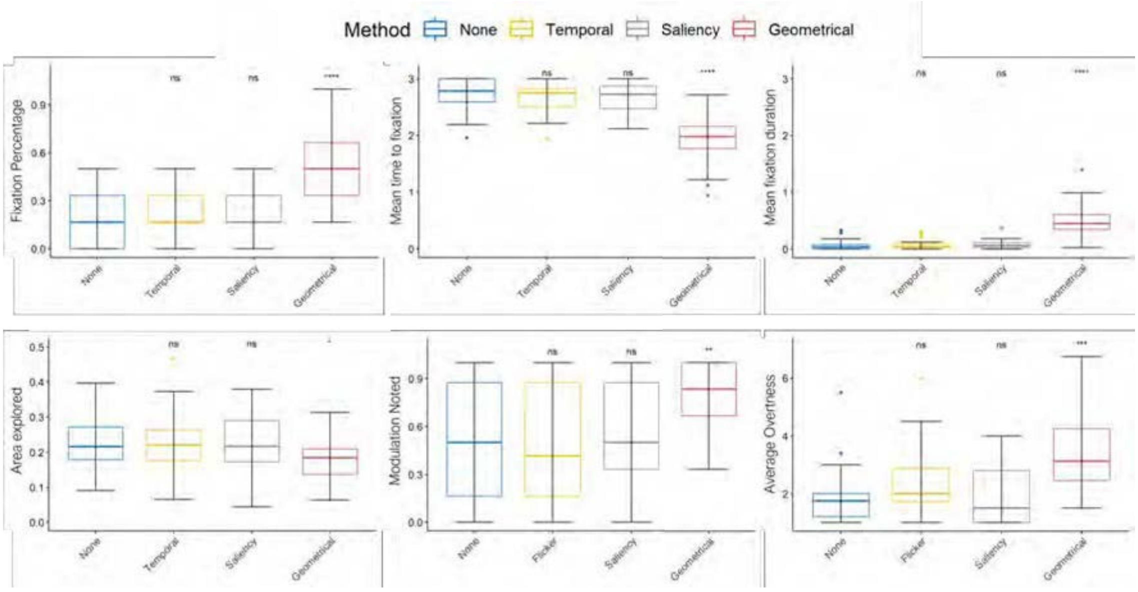


Figure 5.14: Results from the conditions where the modulation techniques were at 25% compared against the None condition. Significant differences are noted against the None condition.

for longer than the other conditions (mean = 0.51, sd = 0.31), with Saliency (mean = 0.08, sd = 0.08), Temporal (mean = 0.07, sd = 0.08) and None (mean = 0.06, sd = 0.08) barely holding it at all.

Comparing the AoE, the Geometric conditions caused the user to explore less of the image compared the other conditions (None: p = 0.0176, Temporal: p = 0.0048, Saliency: p = 0.0048). Where under the None (mean = 0.23, sd = 0.08), Temporal (mean = 0.23, sd = 0.09), and Saliency (mean = 0.22 sd, 0.8) over 20% of the image was explored, this was dropped to just under 20% when viewed in the Geometric condition (mean = 0.18, sd = 0.07).

Only Geometric had a significant effect when participants were asked if the modulation was noticed, M (None: p = 0.00196, Temporal: p = 0.00053, Saliency: p = 0.00067), and responses, M7 (None: p = 0.0023, Temporal: p = 0.0203, Saliency: p = 0.0048). The Geometric condition was noted as modified to some degree the most regularly (mean = 0.8, sd = 0.21), whilst the other conditions, None (mean = 0.52, sd = 0.39), Temporal (mean = 0.51, sd = 0.37), and Saliency (mean = 0.56, sd = 0.34) were only considered to have some degree of modification approximately 50% of the time. The semantically anchored ratings for Geometric (mean = 3.45, sd = 1.45) indicated that it was more overt than the other conditions, None (mean = 1.96, sd = 1.06), Saliency (mean = 1.99, sd = 1.05), and Temporal (mean = 2.38, sd = 1.19).

Results at 50%: The results for conditions at 50% can be seen in Figure 5.15. Full sets of p-values from our statistical test are included in the appendix D.2.

At 50%, when looking at F we see significant differences between all conditions

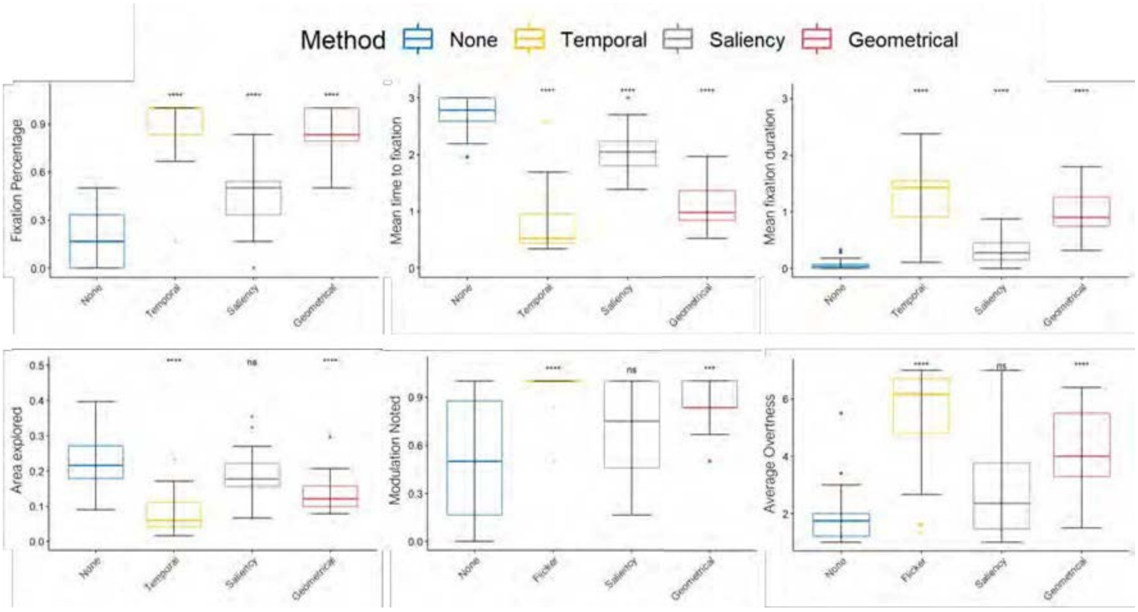


Figure 5.15: Results from the conditions where the modulation techniques were at 50% compared against the None condition. Significant differences are noted against the None condition.

in their ability to have participants fixate on the target except Temporal and Geometric($p = 0.083$) with mean fixations (mean = 0.91, sd = 0.18 and mean = 0.84, sd = 0.14 respectively), Saliency (mean = 0.46, sd = 0.2) and None (mean = 0.17, sd = 0.16).

Looking at the TtFF there is a significant difference in the times for all conditions. Here Temporal has the fastest mean guidance (mean = 0.74, sd = 0.49), followed by Geometric (mean = 1.08, sd = 0.4), then Saliency (mean = 2.09, sd = 0.41) and finally None (mean = 2.74, sd = 0.3).

The significant difference in all conditions also holds true for D with Temporal holding it the most (mean = 1.3, sd = 0.51), then Geometric (mean = 0.99, sd =0.39), followed by Saliency (mean = 0.31, sd = 0.22), and None (mean = 0.06, sd = 0.08) .

Likewise, AoE shows Temporal guidance having the smallest area explored (mean = 0.08, sd = 0.6), followed by Geometric (mean = 0.14, sd = 0.06), Saliency (mean = 0.19, 0.06), and then None (mean = 0.23, sd = 0.08), all with statistically significant differences.

Looking at the subjective responses, both M and M7 show significant differences between all conditions. For the notability of modulations, M, the None condition is the least noted (mean = 0.52, sd = 0.39), then the Saliency (mean = 0.68, sd = 0.3), followed by Geometric (mean = 0.88, 0.13), and finally the most noted condition for having some degree of modulation was Temporal (mean = 0.98, sd = 0.1). For the semantically anchored scale responses, M7, None had the lowest average rating (mean = 1.96, sd = 1.06), then the Saliency condition (mean = 2.7, sd = 1.48), Geometric (mean = 4.17, sd = 1.51), and the Temporal (mean = 5.49, sd = 1.68).

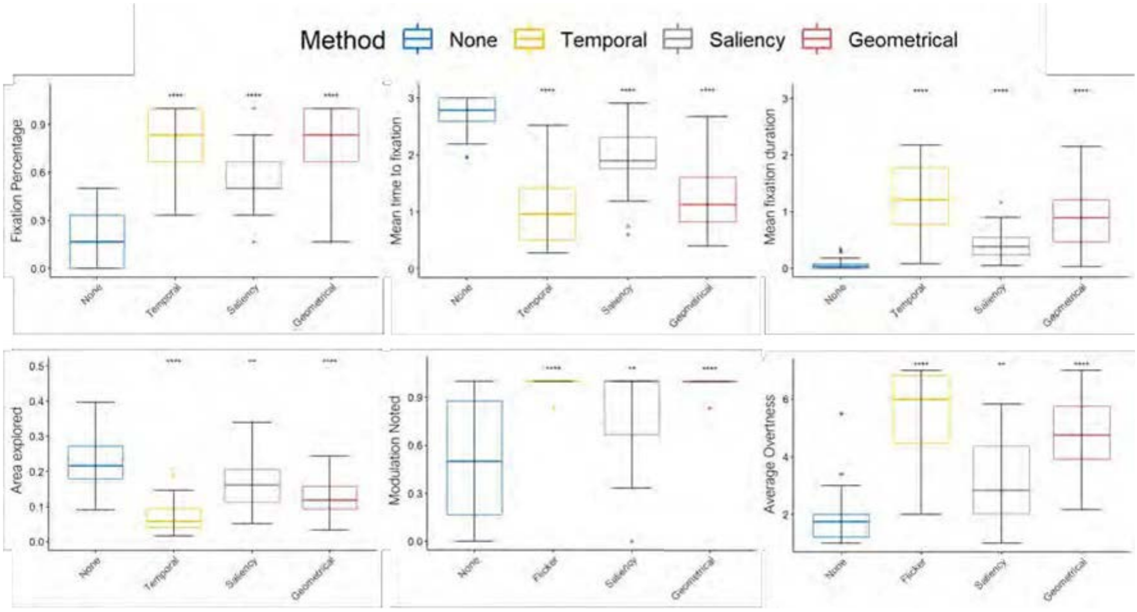


Figure 5.16: Results from the conditions where the modulation techniques were at 75% compared against the None condition. Significant differences are noted against the None condition.

Results at 75%: The results for conditions at 75% can be seen in Figure 5.16. Full sets of p-values from our statistical test are included in the appendix D.3.

Looking at the times participants fixated on the target, F, we see significant differences between all the techniques except Geometric–Temporal ($p = 0.59098$). We see that Temporal and Geometric have the highest mean fixations (mean = 0.82, sd = 0.2 and mean = 0.77, sd = 0.24 respectively), followed by Saliency (mean = 0.55, sd = 0.2) and then the None condition (mean = 0.17, sd = 0.16).

Regarding the TtFF, we do see significant difference in times between all conditions. Once again Temporal is the fastest form of guidance (mean = 1.02, sd = 0.59), then Geometric (mean = 1.22, sd = 0.59), Saliency (mean = 1.92, sd = 0.53), and None (mean = 2.74, sd = 0.3).

Similarly, we see significant differences in D for all conditions. Temporal was again longest (mean = 1.24, sd = 0.61), followed by Geometric (mean = 0.89, sd = 0.53), Saliency (mean = 0.43, sd = 0.26), and finally None (mean = 0.06, sd =0.08).

The area of the images explored, AoE, is also affected differently by all conditions with statistical significance. On average Temporal guidance reduced this to the smallest amount (mean = 0.07, sd = 0.4) compared to the None condition (mean = 0.23, sd = 0.08), whereas Saliency reduced in the least (mean = 0.16, 0.07). Geometric guidance sits in between the other guidance techniques (mean = 0.13, sd = 0.06).

Looking at M, all conditions performed significantly different from each other except Temporal–Flicker ($p = 0.07260$). Temporal (mean = 0.99, sd = 0.03), and Geometric (mean = 0.96, sd = 0.07) were noted the most often, with Saliency being noted approximately 80% of the time on average (mean = 0.82, sd = 0.27), whilst

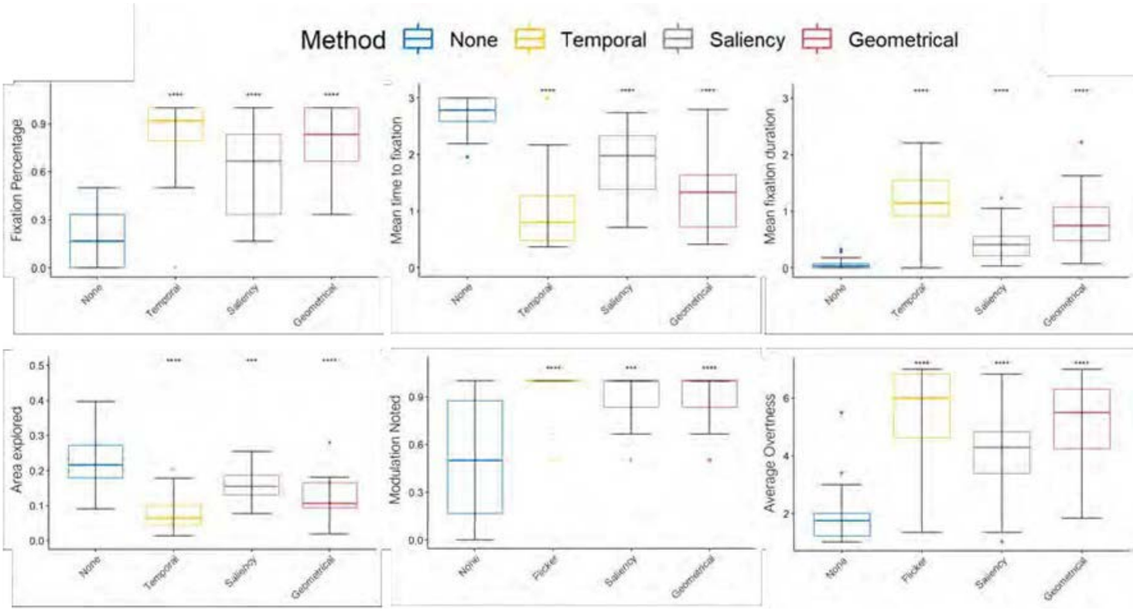


Figure 5.17: Results from the conditions where the modulation techniques were at 100% compared against the None condition. Significant differences are noted against the None condition.

None was noted approximately 50% of the time (mean = 0.52, sd = 0.39).

Different to the modulations being noted, M, the semantically anchored overtness ratings are significantly different between all conditions. We see that Temporal is again the most overt on average (mean = 5.46, sd = 1.64), followed by Geometric (mean= 4.71, sd = 1.33), Saliency (mean = 3.2, sd = 1.49), and None (mean = 1.96, sd = 1.06).

Results at 100%: The results for conditions at 100% can be seen in Figure 5.17. Full sets of p-values from our statistical test are included in the appendix D.4.

At 100% we do not see significant differences between Temporal and Geometric guidance (p = 0.07260), however we see significant differences between all other condition pairings. Temporal and Geometric have the highest mean fixations (mean = 0.83, sd = 0.24 and mean = 0.77, sd = 0.21 respectively), followed by Saliency (mean = 0.6, sd = 0.27) and then the None condition (mean = 0.17, sd = 0.16).

Differing to F, when looking at TtFF we do see significant differences between Temporal and Geometric alongside the rest of the condition pairings. The Temporal condition was fastest (mean = 1, sd = 0.66), followed by Geometric (mean = 1.27, sd = 0.61), Saliency (mean = 1.81, sd = 0.64) and slowest was None (mean = 2.74, sd = 0.3).

The durations, D, were again significantly different for all conditions. Temporal held the gaze for the longest (mean = 1.2, sd = 0.54), and None the shortest (mean = 0.06, sd = 0.08), with Geometric (mean = 0.82, sd = 0.46) and Saliency (mean = 0.44, sd = 0.31) in between them.

Again at 100% we see significant differences between all conditions when comparing AoE. The None condition was explored the most on average (mean = 0.23, sd = 0.08), followed by the Saliency condition (mean = 0.16, sd = 0.05), then Geometric (mean = 0.12, sd = 0.05), and finally Temporal (mean = 0.08, sd = 0.05).

Looking at the notability of the modulations at this level we see no significant differences between the guidance techniques (p-values: Temporal–Saliency = 0.25717, Temporal–Geometric = 0.25717, Saliency–Geometric = 0.32096), but significant differences to the None condition (p-values: Temporal = 0.00062, Saliency = 0.00055, Geometric = 0.00062). Temporal (mean = 0.97, sd = 0.11), Saliency (mean = 0.89, sd = 0.18), and Geometric (mean = 0.92, sd = 0.13) are all noted close to, or greater than, 90% of the time on average which is significantly higher than Flicker (mean = 0.52, sd = 0.39)

Looking at M7, we see significant difference between all conditions except for between Temporal and Geometric (p = 0.12407). We see Temporal (mean = 5.47, sd = 1.69) and Geometric (mean = 5.18, sd = 1.52) have the highest average rating followed by Saliency (mean = 4.18, sd = 1.44), and then None (mean = 1.96, sd = 1.06).

Discussion: One immediate thing to note from our results was that at the 25% level only the Geometric effect had any perceivable effect by the participants and was the only technique to affect their gaze patterns. More precise parameter setting of the other techniques is required to be able to be effective at lower levels.

Also, the noticeability of modulations in the None condition was rated at 50% on average. However, some participants never used a value of zero in their rating a few who were asked noted that they had ignored or forgotten it as an answer. This would explain this abnormality when considered alongside the low average score for the semantically anchored scale. Alternatively, the nature of the display, alongside the impact caused simply by placing the OSTHMD on the optical path when gathering images may have caused participants to perceive the images as modulated.

When comparing the various levels of guidance strength, we can see that all levels where there is an initial fixation component to the temporal flicker effect (levels 50%, 75% and 100%), it provides the fastest *TtFF* and greatest *F*, on par with traditional geometric techniques. It is notable that the geometric technique was the only technique to be effective across all levels of modulation when compared to the unmodulated condition. We also see that flicker, whilst being effective includes the same issues of attention tunnelling seen with geometric techniques. To avoid this saliency appears to be the best option, allowing for reduced time fixated on the target, beyond that necessary to identify it, and directing attention faster and more consistently than the unmodulated condition. However, it did not achieve the speed or guarantees of fixation that the other two techniques can achieve.

When looking at saliency we can again see that precise calibration is needed. Without using calibrations of the parameters, at lower levels of modulation the changes were not significant enough to have any effect, whilst at higher levels the lump sum changes were not appreciated by the participants, who found the sig-

Figure 5.18: Examples of images where modulations need adjustment without context awareness. The white circles show the target areas. Left: an image where modulations struggle due to the bright scene, particularly surrounding the small dark objects being highlighted. Right: an image where the modulations are quickly very apparent and overt due to modulating a dark image.

nificant washing out of colours undesirable. Based on the participants comments tuning saliency techniques towards directly increasing the saliency of the target, with limited reductions in the surrounding environment may be preferable.

We also found that as soon as a drawing time was introduced into the temporal flicker effect it was found to be very overt and participants thought it would become annoying. Turning off the effect once gaze has been drawn, as demonstrated in other works, (McNamara et al., 2008; Grogorick et al., 2018), is a potential means to alleviate this. Based on participant comments the transition to a lower level of flicker was also unwarranted and that maintaining a constant flicker at a set level may actually be preferable to shifting to a lower flicker frequency.

Notably, we can see that our levels selected are not equivalent based on our internal measure of obtrusion. This was a probable weakness of a first study into the space, however, enables a basis for further studies selecting more precisely equivalent levels of modulation for the various techniques. Doing so would require extensive level setting such as that demonstrated by (Veas et al., 2011) to be completed for each technique. In a widely unexplored area, this study provides initial understanding of the space and grounds in which such level setting can be based on.

One aspect to consider from our results is the need for tailoring of algorithms to both user and context. Whilst there is also the confounding variable of interpretation of the question, we can see that geometric guidance was often preferred in the subjective interview however, this was not always the case and for some participants was not the optimal means of guidance. We believe that tailoring modulations to the needs of an individual user will be an important step forward in the development of further methods for visual guidance. Furthermore, the need for context based modulations is self-apparent. We can clearly see in some images where the generic application of modulations can have little to no effect, for example those where the target is a light area surrounded by further light areas, and those where the effect is quickly evident, for example targeting a dark area surrounded by further dark areas causing modulations to create a quick transition (Figure 5.18).

Overall, we see that flicker and outlining techniques both provide effective means to quickly draw attention to a target area and hold their attention there. Notably, flicker appears to be the most effective at this. Saliency was still able to effectively draw attention when compared to no adjustment, however, was slower and less attention holding than the alternatives. This would indicate that saliency techniques may indeed be best left to utilisation in their currently indicated application areas of less obtrusive and more scene preserving subtle methods of visual guidance.

Regarding visual guidance in Computational Glasses, it is evident that there is potential for such techniques to provide effective forms of assistance by directly modulating the view of the user, rather than introducing virtual objects to it. Modulation via Computational Glasses can perform on par with and may even perform better than AR techniques. Furthermore, less obtrusive and attention demanding guidance is possible.

The results from this study also indicate the potential of the temporal-based techniques to provide effective forms of visual assistance and open up a yet unexplored area of research.

5.7 Discussion

In this work, we investigated real-world visual guidance using Computational Glasses and OSTHMDs. To the best of our knowledge this is the first detailing of a design space for visual guidance in AR, exploration of practical saliency modulation using Computational Glasses including the first presenting an actual prototypical implementation and a user study. Whilst prior comparisons have been made between guidance techniques, to the best of our knowledge this is the first to cover all cues of our space in isolation, or in OSTHMDs.

We presented a design space that allows for exploration of current and potential future methods of achieving visual guidance in AR. From our space we were clearly able to see the lack of dedicated techniques that were tailored to the needs of OSTHMDs. We also see a limited number of techniques that actively integrate multiple cueing styles. Cueing styles are also being heavily biased to obtrusive application.

We developed several Computational Glasses prototypes that at their core are based on commercially available OSTHMDs but extended them by integrating scene cameras via beamsplitters that capture the world as seen by the user. This is needed to accurately capture the user's perspective of the world, to compute the saliency modulation overlay, and its correct placement. The prototypes included a functional stereoscopic prototype and a bench prototype where users see through the prototype via a camera at the position of the user's eye (user perspective camera). The latter being needed for conducting the eye tracking study in a controlled study environment reducing confounding variables. In addition, we developed a working mobile prototype that we did not further address. The main reason is that there is perceivable latency (300-500ms), and maintaining a fixed calibrated position on the user's head is challenging. Finally, because the used OSTHMD in the mobile prototype is colour sequential, it is not well suited for capturing results.

We developed a custom natural saliency modulation technique for OSTHMDs. This method utilised various styles of saliency modulation demonstrated in prior techniques, however, did so in an additive only manner. This technique filled an aforementioned gap in the design space.

We explored our approach using images from a widely used image database also providing saliency data. The results show that we can guide the user's gaze towards relevant areas with significant effects in time for first fixation and number of actual fixations. These results are provided in a controlled lab environment with our prototype minimising confounding variables such as eye-display calibration quality. Whilst previous research on saliency modulation states that they are able to "imperceptibly" modify image material to change the saliency (Veas et al., 2011), we did not observe this effect when modulating the saliency via our Computational Glasses prototype as questionnaires showed significant differences in perceived image characteristics (e.g., naturalness and unobtrusiveness). After revisiting earlier studies, we would, in general, be careful with targeting the objective of imperceptibility as prior work showed only feeble evidence and critical image material seemed to be not considered. In particular, results suggest that for scenes that are already visually salient, strong saliency modulation is required which is often perceptible. We essentially argue that the line between imperceptible but effective saliency modulation is so thin (much thinner than previously indicated) that it is hard to generalise and, if possible, this imperceptible saliency modulation demands specific knowledge about the scene and parameters tuned for that scene (context-aware AR (Grubert et al., 2017)) which might be hard to realise in interactive non-controlled environments.

We subsequently explored our wider design space using a more generic set of images. We looked to compare the different styles of cues in isolation at various levels of overtness and found that across all levels of overtness geometric cues in OSTHMDs proved effective as anticipated based on the prior literature. However, we did also see that using temporal cues in Computational Glasses could perform on par with this more traditional cue above the lowest level. In fact, our results show that all methods of guidance were effective above an initial threshold, with variances in how quickly they guided attention (TtFF) and how much they held it (AoE). We did however also see that our chosen levels of overtness did not produce consistent levels when evaluated by participants. This indicates the need for more finely tuned parameters. Our subjective results indicated a preference for geometric techniques due to excessive washing created by saliency modulation, and annoyance caused by temporal flickering. This would indicate the need for careful consideration on how and where to employ visual guidance in Computational Glasses. It also points to the potential need for further temporal components in techniques, or situational awareness to best apply them.

Applications. Overall, we believe in the potential of real-world visual guidance via Computational Glasses. Not only because we can show a different gaze behaviour, but also because different cues caused users either to look at the targeted area whilst not being overly distracted by it or could be directed to quick find and focus on the target. This was indicated by the gaze analysis and is something that is relevant when modulating the real world.

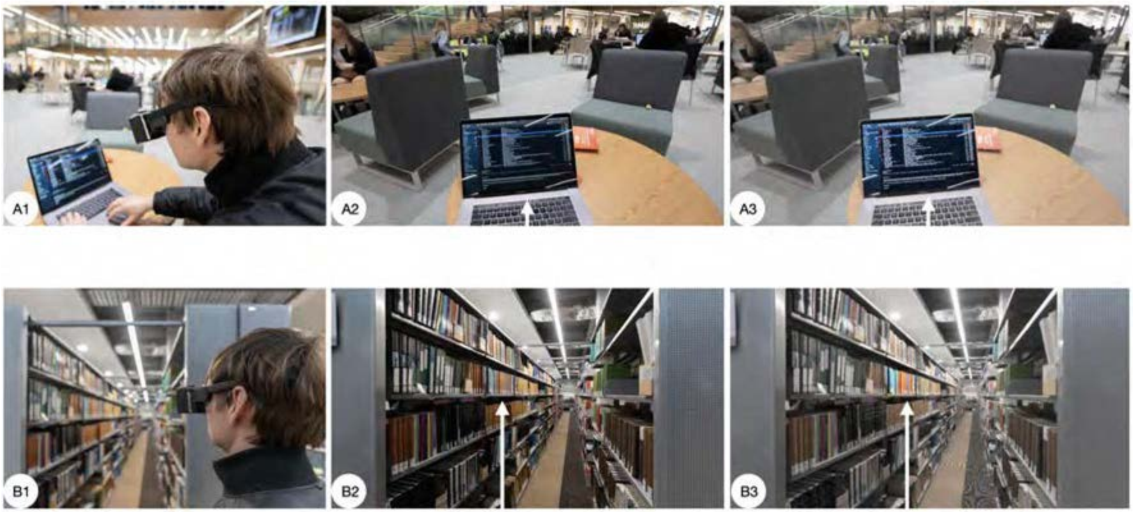


Figure 5.19: Two conceptual scenarios showing how the concept could be used for focusing on a main task (A) or finding objects (B). Our approach captures the real-world (A2 and B2) and modulates the saliency (A3 and B3) to guide the user's gaze to modulated scene elements (here simulated output of the computer and books on the shelf subtly modulated with our algorithm, as intended for an unobtrusive guide). The white arrow pointing out the emphasised area is for illustration only.

Generally speaking, we see most applications in guiding or highlighting information in the user's context. This was also the main direction of our work. This includes guidance during surgery, where occluding areas can be critical. Similarly, we see applications in guiding and navigation scenarios where introducing additional visual cues might occlude information or introduce visual clutter (see 5.19 B). We also think the concept of visual noise cancellation is a strong and interesting concept with applications in many directions (including medical) (see 5.19 A). However, it was not necessarily a focus within this work.

Ethics. A successful visual guidance with optical see-through AR also brings professional and ethical responsibilities for designers, developers, and researchers, amongst others. If our techniques turn out to be highly effective and are used in pervasive AR settings (Grubert et al., 2017), i.e., omnipresent, environmentally adaptive, and an everyday reality augmentation, then careful consideration should be taken in particular regarding health and safety, privacy issues, and produced illusion and belief. We should design our visual modulations in such a way that naturalness and unobtrusiveness can be controlled for the given task, user, and environment. The degree to which we control the perceived difference between reality and virtual reality can lead to unwanted side effects but can also lead to new and meaningful experiences in a host of applications. Visual guidance with saliency modulation can play a major role here.

5.7.1 Limitations

As the first venture into practical, real-world visual guidance using saliency modulation via Computational Glasses, there are several limitations.

Eyetracking. The first limitation is that the studies have not used the stereoscopic prototype, which as pointed out earlier would have affected the validity because of less stable gaze tracking. As stated earlier, using OSTHMDs with integrated eye tracking (e.g., MS HoloLens 2) adds other issues that are even more problematic. However, our approach of using bench prototypes is relatively common within the discipline as it allows for the exclusion of several confounding variables (e.g., quality of the individual eye-display calibration).

Focal Planes. Nearly all current OSTHMDs have a fixed focal plane (or two in the case of the Magic Leap One). Thus, our modulation mask is "sharp" when focusing on a fixed plane. The displayed images and our focal plane were not at the same distance but also not too far off (and the camera was set to an aperture closely resembling the human eye). There are research prototypes that support multiple focal planes as discussed. Our approach would benefit from them, but they are far away from being commercially available and whilst the commercial QD Laser Retissa is accommodation free, it is monocular and has a small eye-box. Similarly, there is research on OSTHMDs that allows environment modulation by subtracting light intensities (Itoh et al., 2019a) but if and when we will see this in commercial products is unclear.

Static Images. We have also limited ourselves to using an existing image dataset with saliency data. While we chose a wide range of scenes, including some we consider as challenging, the dataset is limited in comparison to the real world. Similarly, we decided to only change the saliency of static images via the Computational Glasses. This is similar to previous work (e.g., (Bailey et al., 2009; Grogorick et al., 2017)) and not a limitation of our algorithm. In fact, our approach works in real-time and it should be possible to implement it with similar performance on the latest mobile devices. However, using moving scenes as the material would have even further complicated the analysis of the results as finding good scenes with ground truth data (e.g., gaze) is challenging while other cues (dynamic cues such as moving objects) might introduce additional challenges.

Parameter setting. Concerning the algorithm, we tried to linearly combine parameters into one parameter, which is used for all image material. However, one could improve the results by optimising the combination and level parameters based on context. To our best knowledge, we are not aware of such an approach, but it seems possible. We also used the masking style from related work (e.g., (Veas et al., 2011)) but noticed the masks have a very short ramp between the emphasised area and the surroundings and future work could optimise it with likely improved results, particularly in the noticeability of the modulation.

Terminology. As there was no standardised questionnaires we were aware of that could be incorporated into our study to garner the insights we were looking for, we choose to provide the users with a semantically-anchored scale around chosen terms. We relied on users' understanding of the meaning of the terms in our given context and to be somewhat consistent. Whilst we provided definitions in an attempt of assure this, we have no way to verify this was the case.

Indirect Use of the Prototypes. Finally, in our main study the users were not directly seeing through the OSTHMD but indirectly using a user-perspective camera. We emphasize that this was an intentional decision to increase internal validity and we considered and tried several other options. All participants viewed the same content and this allowed us to circumvent potentially confounding variables, like incorrect alignment of the virtual content with the scene, different colour aberrations for each eye, and refocusing between the virtual content and the scene. We essentially traded some external validity of our study for increased internal validity which is a common approach in this situation. While this could be seen as a limitation it is actually a strength. Our prior studies using a similar approach could confirm transferability of findings from a modulated scene with a user-perspective camera to direct viewing (Langlotz et al., 2018, 2016; Sutton et al., 2022a). We believe this transferability also applies to our work. Whilst the human eye outperforms the cameras used to collect images in our study, the modulation of the view is done optically on the real environment light. This is in contrast to modulating in camera space as done in a VSTHMD system, and the output of the display can be fine-tuned for each participant to ensure that the degree of modulation is close to that observed by the camera. We visually verified we could produce a similar view when using the stereo prototype to that seen in the camera image.

5.7.2 Future Work

There are many interesting future research directions leading on from our design space and the results of our studies, both for visual guidance in general AR, and for its use in Computational Glasses.

A design space for guidance in AR provides a means to explore and develop new techniques for visual guidance. Whilst we provide some initial groundwork in exploring and providing insights into the space, it is evident that this is only explorative research. Further studies building on the outcomes of ours will further inform the development of techniques, a component of the design space that we did not utilise. In particular, developing techniques that look to better integrate various cues and span the axes is a promising future research direction (Hein et al., 2020).

Efficient parameter optimisation. One area of development that there is a great need for in this research is a consistent, established, and efficient manner of optimising parameters for visual guidance techniques. We simplified the steps used by Veas et al. (2011) who showed a complicated means of setting parameters for techniques via a lengthy series of user studies. However, this was of limited effect and time and user heavy. More efficient, established methods for consistently setting parameters across various techniques would greatly enhance and accelerate future research in this direction.

Standardised and Tested User Metrics. As mentioned a limitation to our studies was the user metrics. Better developed and understood user metrics that are more universally applied would greatly help in understanding the design space and relative value of various techniques and cues to application scenarios.

The successful ability to adjust the user's gaze patterns to apply attention to a target area using Computational Glasses shows promise for these devices to be used to augment the HVS. It also showed several areas where further research is possible or needed.

Adept Algorithms. As mentioned, modulations compatible with Computational Glasses were found to either be under-tuned and have no significant impact on gaze or were applied liberally and users were concerned about their application, in particular high levels of saliency modulation. This means that for subtler modulations careful fine tuning and application of modulation parameters is needed, as shown in our first efficacy study. It also means that if more overt modulations are to be employed research on how to better integrate them into the user's view, either by parameter tuning, further temporal adjustments such as varied overtness (Hata et al., 2016), limited areas of application or new algorithms are needed.

As part of the tuning algorithms comes the need to make them context aware and adjust to the input image. Whilst this has been explored in visual guidance before (Mendez et al., 2010) we did not implement it here, and even when implemented it required careful tuning (Veas et al., 2011). Properly integrated and tuned context awareness alongside further research into how best to apply techniques is needed.

Geometric Cues. One area of our design space that is still lacking is geometric cues for Computational Glasses. Whilst these exist for OSTHMDs, effects that can truly modify the geometry or textures of the real world in a way that can draw attention are lacking. Having a basis in spectral modulation and a large amount of groundwork to draw upon we created a custom saliency modulation to close the gap in the space, however did not approach geometric cues due to the amount of ground work needed first to be able to readily implement such a technique.

Temporal Modulations. Finally, the ability of temporal cues in Computational Glasses to perform on par with traditional geometric cueing indicates the potential use of such effects. Further research should consider the integration of this, particularly given the intention for Computational Glasses to be worn and used in non-static scenarios.

Overall, our work provides a design space to aid the development of real-world visual guidance techniques and shows the potential and the practical issues for them in AR and Computational Glasses. We have demonstrated the ability of Computational Glasses to provide vision augmentation to the HVS and this has implications for the use of Computational Glasses as a general means to provide vision assistance. The demonstration of a saliency-based modulation is an important achievement given that prior work has raised the conceptual idea but never actually explored the implementation of saliency modulation via Computational Glasses or OSTHMDs. The study of various styles of guidance cues addresses practical issues such as how to capture the environment and modulate it via Computational Glasses or the limited range for modulating images by only adding light via the Computational Glasses were not mentioned. Our research findings are thus of relevance for the *Human-Computer Interaction (HCI)* and AR communities with potential for future work in designing, developing, and comparing novel visual guidance techniques.

Chapter 6

Discussion and Conclusion

Assisting the *human visual system (HVS)* to better enable us to complete tasks, support it when it is impaired, or to generally improve aspects of life is commonly accepted in society today. This is achieved in many ways, from glasses for aiding refraction errors, to goggles for seeing at night, and utilised in many parts of society, from occupation specific uses like loupes to everyday uses likes glasses.

Traditional glasses provide a convenient way to assist the HVS, being commonplace in today's society and are one of the most ubiquitous solutions to aid it. The static optic lenses have developed greatly in quality over the centuries of their development, however remain constrained in a static state, unable to adjust to a user's needs and limited to the same applications for which they were originally developed. This holds true for most forms of vision assistance.

More recently, new devices that can affect the HVS in the form of *optical see-through head-mounted displays (OSTHMDs)* have become commonplace as a means to introduce virtual content into the real world, creating *augmented reality (AR)*.

There are various ways in which the HVS can be assisted with many forms of enhancement, and still other ways in which it can be impaired that remain unaided. As such, various methods for creating AR have been employed in an attempt to create new, more versatile means to assist the HVS. OSTHMDs have seen used in various ways, such as aiding with peripheral vision loss (Peli et al., 2009) and enhancing light adaptation abilities (Hiroi et al., 2017). However, all of these methods have come with their shortcomings when they are deployed as ubiquitous means to assist the HVS, with OSTHMDs not being able to directly modulate the HVS, but rather introducing new information or misaligned overlays.

Whilst some works with focus-tuneable lenses enable non-static aids for focus issues (Chakravarthula et al., 2018; Padmanaban et al., 2019a), and modifying the world light (Itoh et al., 2019a,b), until now no research has demonstrated how to create effective Computational Glasses utilising OSTHMDs. We looked to investigate how to create versatile Computational Glasses from OSTHMDs that are capable of assisting the HVS by directly modulating users' vision in a pixel-precise manner. This requires being able to analyse the world as the user sees it, creating precise overlays that can interact with the incoming light world to adjust it as needed, and accurately presenting these overlays to the user enabling the modulation of the world view.

6.1 Contributions

This thesis represents several contributions to the fields of assistive aids, augmenting humans, *Human-Computer Interaction (HCI)*, OSTHMDs, and AR.

Firstly, we present overarching research on Computational Glasses as forms of vision assistance, both to aid and enhance the HVS. We demonstrated the ability to create these devices from commercial OSTHMDs and their efficacy in multiple domains, looking to answer **H1**.

Secondly, we present our research into the ability of these Computational Glasses to provide new means to aid the HVS in ways that cannot be achieved by current glasses. Namely, we demonstrate their use as aids for *colour vision deficiency (CVD)* and answer **H2**.

Thirdly, we present our research into using these glasses to support the unimpaired HVS. We look to enhance its capabilities to find areas of interest which might otherwise go unnoticed via visual guidance, and answer **H3**.

6.1.1 Computational Glasses for Visual Assistance

When considering our first overarching hypothesis ***H1:****"Utilising AR OSTHMDs we will be able to create effective Computational Glasses that can modulate the users view with pixel precision, providing visual assistance"*, we can see that we are successfully able to prove this hypothesis.

We investigated the efficacy of Computational Glasses in two separate use cases, both as a means to aid the HVS and as a means to enhance it. In both cases we were able to create modulations that altered the users perception of the world in the desired manner. Both of these use cases were developed on the same platform of an OSTHMD modified with user perspective scene cameras that are placed on-axis with the user's view via beam-splitters. This demonstrates the potential versatility of such devices to be applied to various application areas and to provide a more holistic form of vision assistance.

The development of various prototypes all based on the same platform allowed us to test and verify the efficacy of Computational Glasses, whilst controlling for various confounding variables and backing the findings of each study.

We also saw that whilst techniques for providing assistance can draw heavily from traditional AR, OSTHMD, and general computing techniques, the unique nature of Computational Glasses necessitates the creation of new and tailored methods for these devices.

6.1.2 Computational Glasses for Visual Aid

When we look at our second overarching hypothesis ***H2:****"Utilising these Computational Glasses we will be able to adjust the visual spectra seen to provide aid for visual impairments. In particular, provide effective compensations for CVD."* we were also able to verify this by demonstrating their ability to assist those with CVD.

We developed a pipeline in order to analyse the world images for areas of difference for those with CVD and then created overlays that modulated the colours perceived by the user and compensated for their issues. We modified several techniques from the related literature on compensation in general computing to suit our system. The series of studies we ran on the system testing efficacy proves **H2** and we subsequently looked into the potential future research directions for aiding CVD.

In this manner we were able to show that spectral modulations could effectively compensate for a user's inability to differentiate colours, particularly in extreme cases of confusion. We also saw that such an aid could perform on par with alternative existing aids when considering usability and task load, even when said aids were advantaged.

We also found that working with those who would be aided by the Computational Glasses was complementary to creating effective aids and testing their efficacy.

6.1.3 Computational Glasses for Visual Augmentation

When considering our third overarching hypothesis ***H3:*** *"Computational Glasses will be able to not only assist impaired vision by providing visual aids, they will also be able to assist the unimpaired vision. In particular, they will able to effectively guide user's attention by modulating the saliency of a scene."*, we found that we could accept this hypothesis as we demonstrated their ability to guide vision.

To investigate this hypothesis we first looked to investigate methods to provide visual guidance, starting with those utilised in AR. We created a design space to cover these techniques. We then created a method of saliency modulation that was applicable to the OSTHMDs, building on our prior work in spectral modulations to create a technique to modulate the saliency of the real world via Computational Glasses. We verified that we could effectively adjust the gaze of users as intended using our technique. We then looked to explore a greater portion of our space looking at our new technique at various levels, alongside traditional AR geometric and Computational Glasses compatible temporal techniques. Further exploration of the space demonstrated the potential for Computational Glasses to be effective at providing visual guidance at various levels of overtness and to varied degrees based on need.

Continuing our research on spectral modulations, we were able to show that Computational Glasses could provide an alternative means to help users notice target areas of interest. It could do so in a manner that looked to only modulate the real world and produced a smaller anchoring effect than that of traditional geometric AR techniques. We also looked to investigate temporal effects for guiding attention in Computational Glasses and saw that by doing so, Computational Glasses could even outperform traditional geometric techniques and spectral saliency modulations.

We also identified the difficulties faced to finely tune techniques that look to modulate the world to a desired degree. This was evident in the need to finely tune the saliency algorithm to achieve the intended effect. It was also evident in the need for further tuning of parameters for all guidance techniques we tested.

6.1.4 Vision Assistance as a Discipline

When reflecting on this thesis one could argue that another contribution is the strengthening of Vision Assistance or Computational Vision aids as an independent research discipline. When starting this thesis, we understood our work as an application of AR. However, the more we progressed into our research on Computational Glasses it became more apparent that we share and use some techniques from AR but our research also has some fundamental differences. Firstly, the traditional concept of AR highlights the need for precise tracking in 3D. We showed in our work that traditional tracking was of much less use to us and for Vision Assistance in general. Instead, we emphasise that a per-pixel modulation is crucial and demonstrated how this can be achieved without following the traditional tracking approach. Similarly, traditional AR focuses very much on the idea of integrating digital information into the physical world. Instead, we tried to only modulate the physical world by changing its appearance, trying to keep the changes to the bare minimum needed to achieve the effect. Whilst we acknowledge that AR is interdisciplinary in its core, we also see that research on Vision Assistance and Computational Glasses would benefit from the involvement of specific disciplines such as cognitive science, ophthalmology, academics from AR, augmented human, optics, and also researchers from HCI and here in particular those that specialise in assistive technologies. We believe that the differences in the application and the emerging differences in the requirements and the manner in which they are applied may indicate the start of a separate discipline to be studied.

6.2 Considerations

Beyond our contributions, limitations, and directions for future research, there are several things that arise from this research and the overarching field of Computational Glasses that it inhabits that bare discussion.

6.2.1 Social Acceptability

Throughout this thesis the goal for the form factor and style of Computational Glasses is to become like that of traditional glasses in order to be socially acceptable and common place. However, one must also consider the social acceptability of glasses and how they too were not always accepted. Whilst developments in form factor have help them, understanding of their value and increased uptake were also instrumental. Improving the form factor of Computational Glasses will assist in uptake, however one should also consider the need for social integration. For example, many people reported being unwilling to accept current aids for *low vision* due to social stigma and impact on employment (Sivakumar et al., 2020). Furthermore, there is room for consideration that allowances may be able to be made for the form of such glasses if their value becomes sufficient and understood.

Another component to the social acceptability of Computational Glasses is the difference in aiding versus enhancing our vision and the notion of modulating our unimpaired vision. The reception for new forms of aids for the impaired is normally warm, and they are often lauded. However, anecdotally, the reception for forms of assistance and augmentation for betterment, without leading cause can be less so.

6.2.2 Ethics

As touched on when discussing the use of our saliency modulation, the use of Computational Glasses comes with some ethical considerations that must be considered by the responsible researchers. This is particularly important if Computational Glasses are pervasively used.

The utilisation of omnipresent cameras analysing the world around the user raises various privacy concerns. This issue is not one unique to Computational Glasses, as OSTHMDs already face this issue (Wassom, 2014; Slater et al., 2020).

Another issue facing OSTHMDs that is transferred to Computational Glasses is that of modifying a user's world in a manner they find believable (Regenbrecht et al., 2022). This issue becomes even greater with Computational Glasses when we consider that the intention is to directly modulate the user's perception of the world.

One must also consider the impact on communication and understanding presented by a believed modulation of the environment. For example, colours are referred to by name, and changing these impacts users' ability to understand and communicate what they see, both to another party, or when using and not using the glasses.

6.2.3 Hardware

We built all of our prototypes using simpler OSTHMDs and basic beamsplitters as these provided us with many advantages when it came to prototyping, running user studies and documenting our work.

For OSTHMDs we heavily utilised the Epson Moverio series and specifically the BT-300. This is due to the relative simplicity of the display architecture used. A simple waveguide with an OLED allows for a large colour range, with precise colours, and a bright screen. Also, the display is not time-sequential (images are displayed as a whole, rather than in parts), making capturing images through the device readily achievable. These devices are more readily available than other alternatives on the market, and are relatively inexpensive. The final advantage to using the BT-300 is their form factor. Being small devices without bulky housings and straps, they are easily modified and integrated into our system. The BT-300 is however limited in *Field of View (FoV)*. We also use the Lumus DK-52 for our mobile study prototype. This display provided a much brighter screen with a high colour range, direct HDMI input for reduced latency and increased FoV.

We chose to utilise the Lumus and Espon displays over some of the more popular devices in recent research trends, such as the HoloLens V1 or V2 and the MagicLeap One for various reasons including; display quality, customisation, display control and photograph-ability. One of the big advantages of these devices is their inbuilt tracking capabilities. However due to the nature of Computational Glasses we do not need to utilise this. Another advantage to using these devices is the integrated computing power. This has advantages for creating a mobile device as an additional computational unit does not need to be created and tethered. The final advantage to using these devices is the integrated eye tracking in the HoloLens 2 and MagicLeap One. This would allow for studies to be conducted utilising these, as well as modulations to integrate user gaze as part of their operations. One large problem with the aforementioned devices is their form factor. Due to the integration of world tracking and the complex displays used, integrating these devices into prototypes is complex due to the increased bulk of the designs. Integrating the required hardware for on-axis scene views from the user's perspective would also require dismantling and modification of expensive, intricate, pieces of hardware. Furthermore, using these displays would introduce perceptual issues. Even without the introduction of the beamsplitters current used for user perspective cameras, such devices greatly reduce the visual light perceived by the wearer, acting like sunglasses. This comes alongside various optical aberrations.

For our beamsplitters we used 50/50 half-silvered mirrors throughout our prototypes. These provide a ready amount of light to both the camera and users eyes, however they do reduce the amount of light being perceived by the user. As the human response is logarithmic this is not a significant reduction, however is still noticeable. Alternative plates such as 80/20 could be employed with more sensitive cameras to mitigate this. Furthermore, whilst simple mirror beamsplitters are easy to work with when creating various prototypes they do occupy are large amount of space. Alternative forms of beamsplitters, such as utilising waveguides used in OS-THMDs, *holographic optical elements (HOEs)*, or alternative camera setups, could be employed to create smaller form factors and better light passage through, however come with trade-offs of cost, complexity and reduced flexibility. Practical use of these alternatives also still requires further research and development. As such they were not used in our prototypes.

6.3 Limitations

Whilst we covered the limitations of our individual use cases, there are some limitations to the research conducted in this thesis, predominantly about the studies we ran and the devices used that apply to our overall conclusions around the capabilities of Computational Glasses.

Technology. There are a number of ways in which the technology we used in our studies is still a long way from being mature enough to be used in the intended use case of a ubiquitous form of vision assistance.

The setups we used were generally relatively fragile with the bench prototypes providing the most stable of our prototypes. This is an effect of early products produced by various components bound together in a temporary manner. As of yet none of our prototypes are capable of being fully detethered from a computer, and all would suffer quickly when used outside the carefully constrained lab conditions and without constant maintenance.

Something that may have been a confounding factor on the effect of our glasses when our users were looking directly through prototypes, and had an impact on our mobile study, is the latency of the system. As covered, there is a notable amount of latency in the system, with a lot occurring as display latency. Whilst we could mitigate this effect when using the bench prototypes, in particular during our visual guidance studies, and were still able to demonstrate efficacy, it would have had an impact our results. Furthermore, the currently latency in the system makes them unusable for highly dynamic scenarios, such as general everyday life.

As can be seen from our prototypes, the vast majority of them are still large bulky systems, most being on par with or larger than *video see-through head-mounted displays (VSTHMDs)*. This places them a long way from being used as ubiquitous aids. We were able to demonstrate how we could miniaturise these prototypes to a more acceptable size, however we still utilised simple optics that introduced bulk. To be able to further miniaturise prototypes and develop more generalisable prototypes additional research is needed.

At odds with the size of prototypes is the FoV that they can cover. Increasing the size of the FoV covered generally requires an increase in size (particularly with simple optics). The limitations of the FoV of OSTHMDs strictly limits the FoV which can be aided and modulated. Whilst vignetting can mitigate the transition between modulated and unmodulated vision for Computational Glasses to be able to provide vision assistance in a manner similar to traditional glasses, they need to be able to modulate all relevant areas of vision. This is something that we were unable to achieve in our work.

Internal Validity. Another limitation on our work is the internal validity of our results. There are a number of confounding variables that impact our results such as those imposed by the aforementioned technology. Other confounding variables include the quality of calibrations, limitations of cameras and displays, and the interaction with study apparatus by participants. We were however able to mitigate this limitation somewhat by the various replicated studies under which we were able to control for different confounding variables. We replicated both specific results, such as in our CVD use case, as well as broader result trends in both our use cases.

External Validity. As well as impacts on our internal validity there are number of ways in which the external validity of our results also has limitations applied to it. Where we could mitigate some of the internal limitations via replication under various conditions, the same is not possible for external validity.

The first impact on our external validity was the environments in which our studies were conducted. As our work is the first to develop Computational Glasses with OSTHMD we constrained our studies to highly controlled lab studies. Whilst this allows for the control of confounding variables, it limits the situations in which the efficacy of the glasses has been evaluated.

The second impact to note is the media we used for our studies. Throughout our studies we relied on static images to test efficacy. This enabled us to control for the effects of motion of vision in general as well as the impact of changing images on the output of overlay calculations. It also allowed us to minimise the impact of display latency. However, in particular situations these effects will have an impact and so need to be taken into consideration.

The third impact is that, as stated several times now, the state of OSTHMDs means they are unable to achieve complete control over the light a user perceives, or the ways in which modulations can be applied whilst still allowing the user to perceive a coherent world. This meant that for the purposes of our research, the modulations we choose to apply and the manner in which we applied them was limited and we would hypothesise that further potential will be found with further developments of the underlying technology.

Finally, our parameter selections for our techniques in both use cases were very limited. Whilst we made attempts to set our parameters carefully when testing for effective subtle saliency-based guidance, the rest of the studies used parameters more generically set based on prior works. As we noted in both use cases, an important finding from our results was the potential impact of user and context-based tuning. Therefore, if improved methods of parameter setting were utilised that reflected this, we would expect that results could be very different, and significant improvements may be seen.

6.4 Future Work

The conclusion of our research opens up a number of potential and interesting future directions for research to go.

6.4.1 Extended Studies

As mentioned, due to the explorative nature of our research, the studies conducted in our work were limited and leave a large area of future exploration.

Duration. One interesting component to be further explored in using Computational Glasses is extended and repeated use. In our studies, users were only exposed to the Computational Glasses for up to 30 minutes. This leaves a lot to be investigated about what repeat uses, or extended use and more time to learn how modulation works, would impact people's opinion of them and their efficacy.

Freedom. Currently we have run our studies in very constrained conditions where users' heads are always placed facing an image and unable to turn away. The only exception to this the limited moveability provided in our explorative mobile study on CVD. This creates an artificial viewing scenario and as such more research is needed where users are given increased freedom to move about and rotate their heads.

Location. Associated with enabling more user freedom and movement is running studies outside heavily constrained lab environments. This would provide interesting research into the use of Computational Glasses. Extensions beyond the lab would enable better understanding of the requirements of Computational Glasses to be fully deployed, the scenarios in which they would be used, and how they could best provide assistance.

Content. The content we explored in our studies was static images from limited data sets. As noted with our work on CVD, extended image sets that cover more general scenarios are needed and the development of content for testing Computational Glasses in general scenarios is an interesting research area. Furthermore, as mentioned in our limitations, we have only tested Computational Glasses using static images. Using videos, or non-static views in studies will extend the realism of the studies and whilst this would not have a large impact on the output overlays from our techniques, it would impact how users perceive the modulated world.

6.4.2 Hardware

In order for Computational Glasses to develop further towards pervasive tools to assist vision they need to become less bench bound. Mobile computing, reduced form factor and weight, and increased capability to match the HVS, particularly if aspects such as FoV are needed. Moving away from bench prototypes reduces the desire to use simple optics and allows for more complex designs to be used. However, as discussed, a large consideration is still the optical properties of the perceived vision. Optical aberrations such as those demonstrated by the HoloLens 2 would cause issues with correctly modifying vision and could cause confusions.

Extensions to studies requires many of these further developments in the hardware.

Eye-Tracking. A crucial component to be added to Computational Glasses for many applications, both in research studies (e.g., further exploration of visual guidance) and practical applications (e.g., context adjustments), is to add physiological measures. In particular, practical integration of eye-tracking into prototypes and devices is interesting future work that would open up further opportunities for computational glasses.

Miniaturisation. A particularly interesting component of further research into the underlying hardware of Computational Glasses is the miniaturisation of prototypes. Whilst we showed some smaller prototypes and the potential for further miniaturisation of our work, developing further towards a form factor similar to that of glasses is an interesting and challenging research area.

6.4.3 Further Aids

Having demonstrated the potential for Computational Glasses to assist the HVS, there are a myriad of ways in which they could potential enhance the HVS and visual impairments that remain unaided that could be investigated.

6.4.3.1 Customised Techniques

As part of developing further aids to be deployed on Computational Glasses customised techniques will need to be investigate and created. Whilst prior methods for achieving assistance on computer screens and in AR can provide a basis for the development of methods, as demonstrated by our research, we also underline the need for new and customised techniques to achieve assistance.

Further methods for saliency modulation using intelligent adaptation based on the saliency of the input scene, and tailored CVD compensation techniques are also areas of future research.

Situational Variance Throughout our research it became clear that future modulation techniques would need to consider their current context of use at any given time. In the case of compensating for CVD it was evident that techniques would need to adjust based on the confusions being seen. Furthermore, from participant feedback, techniques should adjust to the wants of the user or be adjustable by the user to desired degrees of modulation. Investigations into integrating such context awareness and user control are interesting research topics in both AR and HCI, and also have application here. We similarly saw a need for context awareness and adjustment to the scene in our investigations into visual guidance. There was a need for techniques to respond differently based on the input scene. We also saw that the application of techniques could be adjusted based on the current needs, and that potential issues with techniques raised by participants may be alleviated by simple actions such as techniques disabling themselves once they have achieved their goal, something demonstrated in the literature.

Given that this was a notably finding from both our investigations into Computational Glasses for vision assistance, we believe it is an important future research direction to investigate the integration of context awareness into techniques. This also extends to Computational Glasses being applied in multiple scenarios and the potential to switch between various combinations of modulations based on context.

6.4.4 Non-Static Scenarios and the Temporal Domain

Our research to date has focused on removing the confounding variable of time from our studies and other than in our final study utilising flicker to guide attention, we have not considered the temporal domain. These timed components provide potential and necessary areas of research.

Non-Static Scenarios. Introducing motion to the test scenarios introduces problems with display latency and calibration. Further developments would include the need to update calibrations at run-time using techniques such as *interaction-free display calibration (INDICA)* (Itoh and Klinker, 2014a). However, Computational

Glasses are designed to be worn and provide assistance throughout life, therefore overcoming these issues is necessary. There is a need to test the efficacy of modulations in non-static scenarios, and to develop techniques that can account for them.

Temporal Domain. Our research focused predominately on the ability of Computational Glasses to modulate the users view by applying constant spectral shifts. However, as seen by the efficacy of temporal flicker as an effective means to guide attention, the non-static nature of Computational Glasses lends them to also applying temporally varying modulations, as well as creating temporal effects. This opens up a new, interesting avenue of approach for using Computational Glasses.

6.5 Conclusion

With this research we have provided prototypes and initial evidence to show that Computational Glasses can be created with OSTHMDs to modulate the vision of wearers with pixel precision. Doing so enables the ability to affect the HVS with positive effects in assisting it in various scenarios. Whilst OSTHMDs still face challenges to enable the creation of the ideal Computational Glasses, research trends show that these challenges are currently being solved. These first steps indicate that upon the creation of Computational Glasses in the desired form factor, there is evidence that they can provide a new style of glasses capable of aiding currently unaided visual impairments and provide general assistance in life.

Bibliography

Ahn, E., Lee, S., and Kim, G. J. (2018). Real-time adjustment of contrast saliency for improved information visibility in mobile augmented reality. *Virtual Reality*, 22(3):245–262.

Akşit, K., Kautz, J., and Luebke, D. (2015). Slim near-eye display using pinhole aperture arrays. *Appl. Opt.*, 54(11):3422–3427.

Akeley, K., Watt, S. J., Girshick, A. R., and Banks, M. S. (2004). A stereo display prototype with multiple focal distances. *ACM TOG*, 23(3):804–813.

Akşit, K., Chakravarthula, P., Rathinavel, K., Jeong, Y., Albert, R., Fuchs, H., and Luebke, D. (2019). Manufacturing application-driven foveated near-eye displays. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1928–1939.

Amano, T. and Kato, H. (2010). Appearance control by projector camera feedback for visually impaired. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pages 57–63, San Francisco, CA, USA. IEEE.

An, J. and Park, J. (2014). A Contrast-based Color Conversion Method for the Maintenance of Sense of the People with Color Vision Deficiency. *Journal of Digital Contents Society*, 15(6):751–761.

Anagnostopoulos, C. and Anagnostopoulos, I. (2007). Intelligent modification for the daltonization process of digitized paintings. In *Proceedings of the 5th International Conference on Computer Vision Systems*, San Francisco, CA, USA. IEEE.

Ananto, B. S., Sari, R. F., and Harwahyu, R. (2011). Color transformation for color blind compensation on augmented reality system. In *Proceedings - 2011 International Conference on User Science and Engineering, i-USEr 2011*, pages 129–134, Selangor, Malaysia. IEEE.

Angelopoulos, A. N., Ameri, H., Mitra, D., and Humayun, M. (2019). Enhanced depth navigation through augmented reality depth mapping in patients with low vision. *Scientific reports*, 9(1):1–10.

Arntz, A., Keßler, D., Borgert, N., Zengeler, N., Jansen, M., Handmann, U., and Eimler, S. C. (2020). Navigating a heavy industry environment using augmented reality - a comparison of two indoor navigation designs. In Chen, J. Y. C. and Fragomeni, G., editors, *Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications*, pages 3–18, Cham. Springer International Publishing.

Ates, H. C., Fiannaca, A., and Folmer, E. (2015). Immersive simulation of visual impairments using a wearable see-through display. In *Proceedings of the ninth international conference on tangible, embedded, and embodied interaction*, pages 225–228.

Azuma, K. and Koike, H. (2018). A study on gaze guidance using artificial color shifts. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, New York, USA. Association for Computing Machinery.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385.

Bailey, R., McNamara, A., Sudarsanam, N., and Grimm, C. (2009). Subtle gaze direction. *ACM Transactions on Graphics*, 28(4):1–14.

Bao, J., Wang, Y., Ma, Y., and Gu, X. (2008). Re-coloring images for dichromats based on an improved adaptive mapping algorithm. In *ICALIP 2008 - 2008 International Conference on Audio, Language and Image Processing, Proceedings*, pages 152–156, Shanghai, China. IEEE.

Bao, S., Tanaka, G., Tamukoh, H., and Suetake, N. (2015). Improvement of Lightness Modification Method Based on Craik-O ' Brien Effect for Dichromats. In *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) November*, pages 75–78, Nusa Dua Bali, Indonesia. IEEE.

Benko, H., Ofek, E., Zheng, F., and Wilson, A. D. (2015). Fovear: Combining an optically see-through near-eye display with projector-based spatial augmented reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 129–135. ACM.

Besic, I., Omanovic, S., and Boskovic, D. (2019). Time-domain color mapping for color vision deficiency assistive technology. In *2019 2nd International Conference on Signal Processing and Information Security (ICSPIS)*, pages 1–4, Dubai, United Arab Emirates. IEEE.

Bimber, O., Coriand, F., Kleppe, A., Bruns, E., Zollmann, S., and Langlotz, T. (2005a). Superimposing pictorial artwork with projected imagery. In *ACM SIGGRAPH 2005 Courses*, SIGGRAPH '05, page 6, New York, USA. Association for Computing Machinery.

Bimber, O., Emmerling, A., and Klemmer, T. (2005b). Embedded entertainment with smart projectors. *Computer*, 38(1):48–55.

Bimber, O., Iwai, D., Wetzstein, G., and Grundhöfer, A. (2008). The visual computing of projector-camera systems. In *ACM SIGGRAPH 2008 Classes*, SIGGRAPH '08, pages 84:1–84:25, New York, USA. ACM.

Bimber, O. and Raskar, R. (2005). *Spatial augmented reality: merging real and virtual worlds*. CRC Press.

Biocca, F., Owen, C., Tang, A., and Bohil, C. (2007). Attention Issues in Spatial Information Systems: Directing Mobile Users' Visual Attention Using Augmented Reality. *Journal of Management Information Systems*, 23(4):163–184.

Biocca, F., Tang, A., Owen, C., and Xiao, F. (2006). Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, page 1115.

Booth, T., Sridharan, S., McNamara, A., Grimm, C., and Bailey, R. (2013). Guiding attention in controlled real-world environments. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '13, page 75–82, New York, USA. Association for Computing Machinery.

Borji, A. and Itti, L. (2015). Cat2000: A large scale fixation dataset for boosting saliency research. *CVPR 2015 workshop on "Future of Datasets"*. arXiv preprint arXiv:1505.03581.

Bowers, A. R., Luo, G., Rensing, N. M., and Peli, E. (2004). Evaluation of a prototype Minified Augmented-View device for patients with impaired night vision. *Ophthalmic and physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 24(4):296–312.

Bozzelli, G., De Nino, M., Pero, C., and Ricciardi, S. (2020). Ar based user adaptive compensation of metamorphopsia. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 1–5.

Bräunl, T., McCane, B., Rivera, M., and Yu, X. (2016). Color Conversion for Color Blindness Employing Multilayer Neural Network with Perceptual Model. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9431:8–19.

Brettel, H., Viénot, F., and Mollon, J. D. (1997). Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 14(10):2647–2655.

Browne, D., Totterdell, P., and Norman, M., editors (1990). *Adaptive User Interfaces*. Academic Press Ltd., London, UK.

Buker, T. J., Vincenzi, D. A., and Deaton, J. E. (2012). The effect of apparent latency on simulator sickness while using a see-through helmet-mounted display: reducing apparent latency with predictive compensation. *Human factors*, 54(2):235–249.

Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A., and Torralba, A. (2012). Mit saliency benchmark. http://saliency.mit.edu/.

Cakmakci, O., Ha, Y., and Rolland, J. P. (2004). A compact optical see-through head-worn display with occlusion support. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2004*, pages 16–25.

Cakmakci, O. and Rolland, J. (2006). Head-worn displays: a review. *Journal of Display Technology*, 2(3):199–216.

Carmack, J. (2013 (accessed October, 2019)). Latency mitigation strategies. https://danluu.com/latency-mitigation/.

Cem, A., Hedili, M. K., Ulusoy, E., and Urey, H. (2020). Foveated near-eye display using computational holography. *Scientific reports*, 10(1):1–9.

Chakravarthula, P., Dunn, D., Akşit, K., and Fuchs, H. (2018). Focusar: Autofocus augmented reality eyeglasses for both real world and virtual imagery. *IEEE TVCG*, 24(11):2906–2916.

Chakravarthula, P., Peng, Y., Kollin, J., Fuchs, H., and Heide, F. (2019). Wirtinger holography for near-eye displays. *ACM Transactions on Graphics (TOG)*, 38(6):1–13.

Chakravarthula, P., Tseng, E., Srivastava, T., Fuchs, H., and Heide, F. (2020). Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. *ACM Transactions on Graphics (TOG)*, 39(6):186.

Chan, X. B. V., Goh, S. M. S., and Tan, N. C. (2014). Subjects with colour vision deficiency in the community: what do primary care physicians need to know? *Asia Pacific Family Medicine*, 13(1):10.

Cheadle, S. W., Parton, A., Müller, H. J., and Usher, M. (2011). Subliminal gamma flicker draws attention even in the absence of transition-flash cues. *Journal of Neurophysiology*, 105(2):827–833.

Chen, L., Ghilardi, M., Busfield, J. J., and Carpi, F. (2021). Electrically tunable lenses: a review. *Frontiers in Robotics and AI*, page 166.

Chen, Y. C. and Liao, T. S. (2011). Hardware digital color enhancement for color vision deficiencies. *ETRI Journal*, 33(1):71–77.

Cheng, D., Wang, Y., Hua, H., and Sasian, J. (2011). Design of a wide-angle, lightweight head-mounted display using free-form optics tiling. *Opt. Lett.*, 36(11):2098–2100.

Chia, A., Gazzard, G., Tong, L., Zhang, X., Sim, E.-L., Fong, A., and Saw, S. M. (2008). Red-green colour blindness in singaporean children. *Clinical & experimental ophthalmology*, 36(5):464–467.

Childs, B. (1989). Human biology: An introduction to human evolution, variation, growth, and adaptability. *American Journal of Human Genetics*, 44(4):593.

Ching, S. L. and Sabudin, M. (2010). Website image colour transformation for the colour blind. In *ICCTD 2010 - 2010 2nd International Conference on Computer Technology and Development, Proceedings*, pages 255–259, Cairo, Egypt. IEEE.

Choi, J., Lee, J., Moon, H., Yoo, S. J., and Han, D. (2019). Optimal color correction based on image analysis for color vision deficiency. *IEEE Access*, 7:154466–154479.

Chua, S. H., Zhang, H., Hammad, M., Hau Chua, S., Zhang, H., Hammad, M., Zhao, S., Goyal, S., and Singh, K. (2015). ColorBless: Augmenting Visual Information for Colorblind People with Binocular Luster Effect. *ACM Transactions on Computer-Human Interaction*, 21(6):1–20.

Cieza, A., Keel, S., Kocur, I., Mccoy, M., and Mariotti, S. P. (2019). World Report on Vision. Technical report, World Health Organization.

Clark, J., Hasselfeld, K., Bigsby, K., and Divine, J. (2017). Colored glasses to mitigate photophobia symptoms posttraumatic brain injury. *Journal of Athletic Training*, 52(8):725–729.

Cobb, S. V., Nichols, S., Ramsey, A., and Wilson, J. R. (1999). Virtual reality-induced symptoms and effects (vrise). *Presence: teleoperators and virtual environments*, 8(2):169–186.

Cornia, M., Baraldi, L., Serra, G., and Cucchiara, R. (2018). Predicting human eye fixations via an lstm-based saliency attentive model. *IEEE Transactions on Image Processing*, 27:5142–5154.

Cornia, M., Baraldi, L., Serra, G., and Cucchiara, R. (2018). Sam: Pushing the limits of saliency prediction models. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1971–19712.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods*, 29(1):27–36.

Dalton, J. (1798). Extraordinary facts relating to the vision of colours. :: Color and Optics. *Memoirs of the literary and philosophical society of Manchester*, 5(1):28–45.

Danieau, F., Guillo, A., and Doré, R. (2017). Attention guidance for immersive video content in head-mounted displays. In *2017 IEEE Virtual Reality (VR)*, pages 205–206.

David Hincapie-Ramos, J., Ivanchuk, L., Sridharan, S. K., and Irani, P. (2014). Smartcolor: Real-time color correction and contrast for optical see-through head-mounted displays. In *ISMAR*, pages 187–194.

Davis, J., Hsieh, Y.-H., and Lee, H.-C. (2015). Humans perceive flicker artifacts at 500 hz. *Scientific reports*, 5:7861.

Deeb, S. (2005). The molecular basis of variation in human color vision. *Clinical genetics*, 67(5):369–377.

Deng, Y., Wang, Y., Ma, Y., Bao, J., and Gu, X. (2007). A Fixed Transformation of Color Images for Dichromats Based on Similarity Matrices. In *Internation Conference on Intelligent Computing*, pages 1018–1028, Lectur Notes in Computer Science. SpringerLink.

Dillman, K. R., Mok, T. T. H., Tang, A., Oehlberg, L., and Mitchell, A. (2018). A visual interaction cue framework from video game environments for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–12, New York, USA. Association for Computing Machinery.

Do, H., Kim, Y. M., and Min, S.-W. (2019). Focus-free head-mounted display based on maxwellian view using retroreflector film. *Applied optics*, 58(11):2882–2889.

Doliotis, P., Tsekouras, G., Anagnostopoulos, C. N., and Athitsos, V. (2009). Intelligent modification of colors in digitized paintings for enhancing the visual perception of Color-blind viewers. In *IFIP International Federation for Information Processing*, volume 296, pages 293–301, Boston, MA. Springer US.

Dorr, M., Martinetz, T., Gegenfurtner, K., and Barth, E. (2004). Guidance of eye movements on a gaze-contingent display. In *Dynamic Perception Workshop of the GI Section "Computer Vision"*, pages 89–94, Tübingen, Germany.

Dunn, D. (2019). Required accuracy of gaze tracking for varifocal displays. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1838–1842.

Dunn, D., Chakravarthula, P., Dong, Q., Akşit, K., and Fuchs, H. (2018). 10-1: Towards varifocal augmented reality displays using deformable beamsplitter membranes. *SID Symposium Digest of Technical Papers*, 49(1):92–95.

Dunn, D., Tippets, C., Torell, K., Kellnhofer, P., Aksit, K., Didyk, P., Myszkowski, K., Luebke, D., and Fuchs, H. (2017). Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. *IEEE TVCG*.

Eckert, M., Volmerg, J. S., and Friedrich, C. M. (2019). Augmented reality in medicine: Systematic and bibliographic review. *JMIR mHealth and uHealth*, 7(4).

Fernandez, A., Fernandez, P., López, G., Calderón, M., and Guerrero, L. A. (2015). Troyoculus: an augmented reality system to improve reading capabilities of night-blind people. In *International Work-Conference on Ambient Assisted Living*, pages 16–28. Springer.

Ferwerda, J. A., Pattanaik, S. N., Shirley, P., and Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 249–258, New York, USA. ACM.

Findlater, L., Stearns, L., Du, R., Oh, U., Ross, D., Chellappa, R., and Froehlich, J. (2015). Supporting Everyday Activities for Persons with Visual Impairments Through Computer Vision-Augmented Touch. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '15*, pages 383–384.

Flatla, D. and Gutwin, C. (2012). Ssmrecolor: Improving recoloring tools with situation-specific models of color differentiation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2297–2306, New York, USA. Association for Computing Machinery.

Flatla, D., Reinecke, K., Gutwin, C., and Gajos, K. (2013). SPRWeb: preserving subjective responses to website colour schemes through automatic recolouring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2013*, pages 2069–2078, New York, USA. Association for Computing Machinery.

Flatla, D. R., Andrade, A. R., Teviotdale, R. D., Knowles, D. L., and Stewart, C. (2015). Colourid: Improving colour identification for people with impaired colour vision. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, page 3543–3552, New York, USA. Association for Computing Machinery.

Fukiage, T., Oishi, T., and Ikeuchi, K. (2014). Visibility-based blending for real-time applications. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 63–72.

Furness III, T. A. and Kollin, J. S. (1995). Virtual retinal display. US Patent 5,467,104.

Gabbard, J. L., Zedlitz, J., Swan, J. E., and Winchester, W. W. (2000). More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces. *Proceedings - IEEE Virtual Reality*, pages 79–86.

Gao, C., Lin, Y., and Hua, H. (2012). Occlusion capable optical see-through head-mounted display using freeform optics. In *11th IEEE ISMAR*, pages 281–282.

Gao, C., Lin, Y., and Hua, H. (2013). Optical see-through head-mounted display with occlusion capability. In Marasco, P. L. and Havig, P. R., editors, *Head- and Helmet-Mounted Displays XVIII: Design and Applications*, volume 8735, pages 107 – 115. SPIE.

Gao, Q., Liu, J., Duan, X., Zhao, T., Li, X., and Liu, P. (2017). Compact see-through 3d head-mounted display based on wavefront modulation with holographic grating filter. *Optics express*, 25(7):8412–8424.

Gao, Q., Liu, J., Han, J., and Li, X. (2016). Monocular 3d see-through head-mounted display via complex amplitude modulation. *Optics express*, 24(15):17372–17383.

Ghafoor, M. J., Junaid, M., and Ali, A. (2019). Nav-cane a smart cane for visually impaired people. In *2019 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, volume 4, pages 1–4, New York, USA. IEEE.

Goel, V. and Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, 16:395–429.

Goldstein, E. and Brockmole, J. (2016). *Sensation and Perception*. Cengage Learning.

Gómez-Robledo, L., Valero, E. M., Huertas, R., Martínez-Domingo, M. A., and Hernández-Andrés, J. (2018). Do enchroma glasses improve color vision for colorblind subjects? *Opt. Express*, 26(22):28693–28703.

Gonçalves, P., Orlosky, J., and Machulla, T.-K. (2020). An augmented reality assistant to support button selection for patients with age-related macular degeneration. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 730–731. IEEE.

Grasset, R., Langlotz, T., Kalkofen, D., Tatzgern, M., and Schmalstieg, D. (2012). Image-driven view management for augmented reality browsers. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, pages 177–186.

Green, A., Cohen-Zion, M., Haim, A., and Dagan, Y. (2017). Evening light exposure to computer screens disrupts human sleep, biological rhythms, and attention abilities. *Chronobiology International*, 34(7):855–865.

Grogorick, S., Albuquerque, G., Tauscher, J.-P., and Magnor, M. (2018). Comparison of unobtrusive visual guidance methods in an immersive dome environment. *ACM Transactions on Applied Perception*, 15(4).

Grogorick, S., Stengel, M., Eisemann, E., and Magnor, M. (2017). Subtle gaze guidance for immersive environments. *SAP 2017*.

Grogorick, S., Tauscher, J.-P., Albuquerque, G., Kassubeck, M., and Magnor, M. (2019). Towards vr attention guidance: Environment-dependent perceptual threshold for stereo inverse brightness modulation. In *ACM Symposium on Applied Perception 2019*, SAP '19, New York, USA. Association for Computing Machinery.

Grubert, J., Itoh, Y., Moser, K., and Swan, J. E. (2018). A survey of calibration methods for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 24(9):2649–2662.

Grubert, J., Langlotz, T., Zollmann, S., and Regenbrecht, H. (2016). Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1.

Grubert, J., Langlotz, T., Zollmann, S., and Regenbrecht, H. (2017). Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1706–1724.

Guenter, B., Finch, M., Drucker, S., Tan, D., and Snyder, J. (2012). Foveated 3d graphics. *ACM Trans. Graph.*, 31(6):164:1–164:10.

Hagiwara, A., Sugimoto, A., and Kawamoto, K. (2011). Saliency-Based Image Processing for Guiding Visual Attention. In *PETMEI*, pages 1–8.

Hamasaki, T. and Itoh, Y. (2019). Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1961–1969.

Hasana, S., Fujimoto, Y., Plopski, A., Kanbara, M., and Kato, H. (2019). Improving color discrimination for color vision deficiency (cvd) with temporal-domain modulation. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 243–244, New York, USA. IEEE.

Hata, H., Koike, H., and Sato, Y. (2016). Visual Guidance with Unnoticed Blur Effect. *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '16*, pages 28–35.

He, Z., Sui, X., Jin, G., and Cao, L. (2019). Progress in virtual reality and augmented reality based on holographic display. *Applied optics*, 58(5):A74–A81.

Hein, P., Bernhagen, M., and Bullinger-Hoffmann, A. (2020). Two is Better than One. Improved Attention Guiding in AR by Combining Techniques. *IEEE Computer Graphics and Applications*, 40(5):57–66.

Herbst, M. and Brinkman, B. (2014). Color-via-pattern: distinguishing colors of confusion without affecting perceived brightness. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*, pages 245–246, New York, USA. Association for Computing Machinery.

Hicks, S. L., Wilson, I., Muhammed, L., Worsfold, J., Downes, S. M., and Kennard, C. (2013). A depth-based head-mounted visual display to aid navigation in partially sighted individuals. *PloS one*, 8(7).

Hiroi, Y., Itoh, Y., Hamasaki, T., and Sugimoto, M. (2017). Adaptivisor: Assisting eye adaptation via occlusive optical see-through head-mounted displays. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, New York, USA. Association for Computing Machinery.

Hiroi, Y., Kaminokado, T., Mori, A., and Itoh, Y. (2020). Dehazeglasses: Optical dehazing with an occlusion capable see-through display. In *Proceedings of the Augmented Humans International Conference*, AHs '20, New York, USA. Association for Computing Machinery.

Hirzle, T., Gugenheimer, J., Geiselhart, F., Bulling, A., and Rukzio, E. (2019). A design space for gaze interaction on head-mounted displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, USA. Association for Computing Machinery.

Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33.

Hou, X. and Zhang, L. (2007). Saliency detection: A spectral residual approach. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8.

Houston, K. E., Bowers, A. R., Peli, E., and Woods, R. L. (2018). Peripheral Prisms Improve Obstacle Detection during Simulated Walking for Patients with Left Hemispatial Neglect and Hemianopia. *Optometry and Vision Science*, 95(9):795–804.

Howard, I. P. and Rogers, B. J. (1996). *Binocular Vision and Stereopsis.* Oxford University Press.

Howlett, E. M. (1992). High-resolution inserts in wide-angle head-mounted stereoscopic displays. In *Proceedings of the SPIE*, volume 1669.

Hu, C., Zhai, G., and Li, D. (2015). An augmented-reality night vision enhancement application for see-through glasses. In *2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, pages 1–6.

Hu, X., Liu, X., Mao, X., and Wong, T. (2019a). Colorblind-shareable videos. In *2019 International Conference on Cyberworlds (CW)*, pages 356–359, New York, USA. IEEE.

Hu, X., Zhang, Y., Isoyama, N., Sakata, N., and Kiyokawa, K. (2021). Design and prototyping of computational sunglasses for autism spectrum disorders. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 581–582.

Hu, X., Zhang, Z., Liu, X., and Wong, T. (2019b). Deep visual sharing with colorblind. *IEEE Transactions on Computational Imaging*, 5(4):649–659.

Hua, H. (2017). Enabling focus cues in head-mounted displays. *Proceedings of the IEEE*, 105(5):805–824.

Hua, H. and Javidi, B. (2014). A 3d integral imaging optical see-through head-mounted display. *Opt. Express*, 22(11):13484–13491.

Huang, C. R., Chiu, K. C., and Chen, C. S. (2011). Temporal color consistency-based video reproduction for dichromats. *IEEE Transactions on Multimedia*, 13(5):950–960.

Huang, F.-C., Chen, K., and Wetzstein, G. (2015). The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM TOG*, 34(4):60.

Huang, J., Kinateder, M., Dunn, M. J., Jarosz, W., Yang, X.-D., and Cooper, E. A. (2019). An augmented reality sign-reading assistant for users with reduced vision. *PloS one*, 14(1).

Huang, J.-B., Chen, C. S., Jen, T.-C., and Wang, S.-J. (2009). Image Recolorization for The Colorblind. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1161–1164, New York, USA. IEEE.

Huang, J. B., Tseng, Y. C., Wu, S. I., and Wang, S. J. (2007). Information preserving color transformation for protanopia and deuteranopia. *IEEE Signal Processing Letters*, 14(10):711–714.

Hung, P.-c. and Hiramatsu, N. (2013). A Colour Conversion Method Which Allows Colourblind and Normal-Vision People Share Documents with Colour Content. Technical report, Konica Minolta.

Hwang, A. D. and Peli, E. (2014). An augmented-reality edge enhancement application for google glass. *Optometry and vision science: official publication of the American Academy of Optometry*, 91(8):1021.

Iaccarino, G. and Malandrino, D. (2006). Efficient edge-services for colorblind users. In *Proceedings of the 15th international conference on World Wide Web*, pages 919–920, New York, USA. Association for Computing Machinery.

Ichikawa, M., Tanaka, K., Kondo, S., Hiroshima, K., Ichikawa, K., Tanabe, S., and Fukami, K. (2003). Web-Page Color Modification for Barrier-Free Color Vision with Genetic Algorithm. In *Genetic and Evolutionary Computation — GECCO 2003*, volume 2724, pages 2134–2146, Germany. Springer.

Ichikawa, M., Tanaka, K., Kondo, S., Hiroshima, K., Ichikawa, K., Tanabe, S., and Fukami, K. (2004). Preliminary study on color modification for still images to realize barrier-free color vision. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pages 36–41, New York, USA. IEEE.

Itoh, Y., Amano, T., Iwai, D., and Klinker, G. (2016a). Gaussian light field: Estimation of viewpoint-dependent blur for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2368–2376.

Itoh, Y., Dzitsiuk, M., Amano, T., and Klinker, G. (2015). Semi-Parametric Color Reproduction Method for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visusalization and Computer Graphics (TVCG)*, 21(11):1269–1278. Proceedings of ISMAR 2015: long papers.

Itoh, Y., Hamasaki, T., and Sugimoto, M. (2017). Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator. *IEEE transactions on visualization and computer graphics*, 23(11):2463–2473.

Itoh, Y. and Klinker, G. (2014a). Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In *2014 IEEE symposium on 3d user interfaces (3dui)*, pages 75–82, New York, USA. IEEE.

Itoh, Y. and Klinker, G. (2014b). Performance and sensitivity analysis of INDICA: Interaction-free display calibration for optical see-through head-mounted displays. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2014*, pages 171–176.

Itoh, Y. and Klinker, G. (2015a). Light-field correction for spatial calibration of optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)*, 21(4):471–480.

Itoh, Y. and Klinker, G. (2015b). Simultaneous direct and augmented view distortion calibration of optical see-through head-mounted displays. In *2015 IEEE International Symposium on Mixed and Augmented Reality*, pages 43–48, New York, USA. IEEE.

Itoh, Y. and Klinker, G. (2015c). Vision enhancement: Defocus correction via optical see-through head-mounted displays. In *Proceedings of the 6th Augmented Human International Conference*, AH '15, page 1–8, New York, USA. Association for Computing Machinery.

Itoh, Y., Langlotz, T., Iwai, D., Kiyokawa, K., and Amano, T. (2019a). Light attenuation display: Subtractive see-through near-eye display via spatial color filtering. *IEEE transactions on visualization and computer graphics*, 25(5):1951–1960.

Itoh, Y., Langlotz, T., Sutton, J., and Plopski, A. (2021). Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Comuting Surveys*, X(X):1–35.

Itoh, Y., Langlotz, T., Zollmann, S., Iwai, D., Kiyokawa, K., and Amano, T. (2019b). Computational phase-modulated eyeglasses. *IEEE transactions on visualization and computer graphics*.

Itoh, Y., Langlotz, T., Zollmann, S., Iwai, D., Kiyokawa, K., and Amano, T. (2021). Computational phase-modulated eyeglasses. *IEEE Transactions on Visualization & Computer Graphics*, 27(3):1–1.

Itoh, Y., Orlosky, J., Huber, M., Kiyokawa, K., and Klinker, G. (2016b). Ost rift: Temporally consistent augmented reality with a consumer optical see-through head-mounted display. In *2016 IEEE Virtual Reality (VR)*, pages 189–190.

Itoh, Y., Orlosky, J., Kiyokawa, K., and Klinker, G. (2016c). Laplacian vision: Augmenting motion prediction via optical see-through head-mounted displays. In *Proceedings of the 7th Augmented Human International Conference 2016*, pages 1–8.

Itoh, Y., Yamamoto, K., and Ochiai, Y. (2018). Retinal hdr: Hdr image projection method onto retina. In *SIGGRAPH Asia 2018 Posters*, SA '18, New York, USA. Association for Computing Machinery.

Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:1254–1259.

Jacobs, G. H. (2009). Evolution of colour vision in mammals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1531):2957–2967.

Jang, C., Bang, K., Moon, S., Kim, J., Lee, S., and Lee, B. (2017). Retinal 3d: Augmented reality near-eye display via pupil-tracked light field projection on retina. *ACM Trans. Graph.*, 36(6):190:1–190:13.

Jefferson, L. and Harvey, R. (2006). Accommodating color blind computer users. In *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility Assets 06*, pages 40–47, New York, USA. Association for Computing Machinery.

Jefferson, L. and Harvey, R. (2007). An interface to support color blind computer users. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems - CHI '07*, pages 1535–1538, New York, USA. Association for Computing Machinery.

Jeffri, N. F. S. and Rambli, D. R. A. (2020). Guidelines for the interface design of ar systems for manual assembly. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*, ICVARS 2020, page 70–77, New York, USA. Association for Computing Machinery.

Jenny, B. and Kelso, N. V. (2007a). Color design for the color vision impaired. *Cartographic Perspectives*, 4(58):61–67.

Jenny, B. and Kelso, N. V. (2007b). Designing maps for the colour-vision impaired. *Bulletin of the Society of Cartographers*, 41(1-2):9–12.

Jeong, J. Y., Kim, H. J., Kim, Y. H., Wang, T. S., and Ko, S. J. (2012a). Enhanced re-coloring method with an information preserving property for color-blind person. In *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, pages 600–601, New York, USA. IEEE.

Jeong, J. Y., Kim, H. J., Wang, T. S., Jung, S. W., and Ko, S. J. (2012b). Real-time video re-coloring algorithm considering the temporal color consistency for the color-blind. *IEEE Transactions on Consumer Electronics*, 58(2):721–729.

Jeong, J. Y., Kim, H. J., Wang, T. S., Yoon, Y. J., and Ko, S. J. (2011). An efficient re-coloring method with information preserving for the color-blind. *IEEE Transactions on Consumer Electronics*, 57(4):1953–1960.

Jerald, J. and Whitton, M. (2009). Relating scene-motion thresholds to latency thresholds for head-mounted displays. In *Proceedings of the 2009 IEEE Virtual Reality Conference*, VR '09, pages 211–218, Washington, DC, USA. IEEE Computer Society.

Jia-Bin, H., Sih-Ying, W., and Chu-Song, C. (2008). Enhancing color representation for the color vision impaired. In *Workshop on Computer Vision Applications for the Visually Impaired*, Marseille, France. HAL.

Jones, P. R., Somoskeöy, T., Chow-Wing-Bom, H., and Crabb, D. P. (2020). Seeing other perspectives: evaluating the use of virtual and augmented reality to simulate visual impairments (openvissim). *NPJ digital medicine*, 3(1):1–9.

Jota, R., Ng, A., Dietz, P., and Wigdor, D. (2013). How fast is fast enough?: A study of the effects of latency in direct-touch pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2291–2300, New York, USA. ACM.

Jung, J.-H. and Peli, E. (2018). Field Expansion for Acquired Monocular Vision Using a Multiplexing Prism. *Optometry and Vision Science*, 95(9):814–828.

Kalloniatis, M. and Luu, C. (2007). The perception of color. *Webvision: The Organization of the Retina and Visual System [Internet]*.

Kaminokado, T., Hiroi, Y., and Itoh, Y. (2020). Stainedview: Variable-intensity light-attenuation display with cascaded spatial color filtering for improved color fidelity. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3576–3586.

Kato, H. and Billinghurst, M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85–94. IEEE.

Khan, A., Matejka, J., Fitzmaurice, G., and Kurtenbach, G. (2005). Spotlight: Directing users' attention on large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, page 791–798.

Kim, J., Jeong, Y., Stengel, M., Akşit, K., Albert, R., Boudaoud, B., Greer, T., Kim, J., Lopes, W., Majercik, Z., et al. (2019a). Foveated ar: dynamically-foveated augmented reality display. *ACM Transactions on Graphics (TOG)*, 38(4):99.

Kim, K., Heo, D., and Hahn, J. (2019b). Occlusion-capable head-mounted display. In *Proc. of the 7th PHOTOPTICS 2019, Prague, Czech Republic, February 25-27, 2019*, pages 299–302. SciTePress.

Kim, K., Heo, D., and Hahn, J. (2019c). Occlusion-capable head-mounted display. In Raposo, M., Ribeiro, P. A., and Andrews, D., editors, *Proceedings of the*

7th International Conference on Photonics, Optics and Laser Technology, PHOTOPTICS 2019, Prague, Czech Republic, February 25-27, 2019, pages 299–302. SciTePress.

Kinateder, M., Gualtieri, J., Dunn, M. J., Jarosz, W., Yang, X.-D., and Cooper, E. A. (2018). Using an augmented reality device as a distance-based vision aid—promise and limitations. *Optometry and Vision Science*, 95(9):727.

Kiyokawa, K. (2007). A wide field-of-view head mounted projective display using hyperbolic half-silvered mirrors. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '07, pages 1–4, Washington, DC, USA. IEEE Computer Society.

Kiyokawa, K., Billinghurst, M., Campbell, B., and Woods, E. (2003). An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *2nd IEEE/ACM ISMAR*, page 133.

Kiyokawa, K., Kurata, Y., and Ohno, H. (2000). An optical see-through display for mutual occlusion of real and virtual environments. In *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 60–67.

Klein, G. and Murray, D. (2007). Parallel tracking and mapping for small ar workspaces. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 225–234.

Klein, G. and Murray, D. (2008). Compositing for small cameras. In *Proc. Seventh IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'08)*, Cambridge.

Klemm, M., Seebacher, F., and Hoppe, H. (2016). Non-parametric camera-based calibration of optical see-through glasses for ar applications. In *2016 International Conference on Cyberworlds (CW)*, pages 33–40.

Koch, C. and Ullman, S. (1987). Shifts in selective visual attention: towards the underlying neural circuitry. In *Matters of intelligence*, pages 115–141. Springer.

Kokui, T., Takimoto, H., Mitsukura, Y., Kishihara, M., and Okubo, K. (2013). Color image modification based on visual saliency for guiding visual attention. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pages 467–472.

Konrad, R., Padmanaban, N., Molner, K., Cooper, E. A., and Wetzstein, G. (2017). Accommodation-invariant computational near-eye displays. *ACM Trans. Graph.*, 36(4):88:1–88:12.

Koshi, M., Sakata, N., and Kiyokawa, K. (2019). Augmented concentration: Concentration improvement by visual noise reduction with a video see-through hmd. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1030–1031.

Kotera, H. (2012). Optimal Daltonization by Spectral Shift for Dichromatic Vision. In *Color and Imaging Conference*, pages 302–308, Springfield, USA. Society for Imaging Science and Technology.

Krajancich, B., Padmanaban, N., and Wetzstein, G. (2020). Factored occlusion: Single spatial light modulator occlusion-capable optical see-through augmented reality display. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1871–1879.

Kramida, G. (2016). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics*, 22(7):1912–1931.

Kress, B. and Starner, T. (2013). A review of head-mounted displays (hmd) technologies and applications for consumer electronics. *Photonic Applications for Aerospace, Commercial, and Harsh Environments IV*, 8720:62–74.

Kress, B. C. and Chatterjee, I. (2021). Waveguide combiners for mixed reality headsets: a nanophotonics design perspective. *Nanophotonics*, 10(1):41–74.

Kroner, A., Senden, M., Driessens, K., and Goebel, R. (2019). Contextual encoder-decoder network for visual saliency prediction. *CoRR*, abs/1902.06634.

Kuhn, G. R., Oliveira, M. M., and Fernandes, L. A. F. (2008). An efficient naturalness-preserving image-recoloring method for dichromats. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1747–1754.

Kunkel, T. and Reinhard, E. (2010). A reassessment of the simultaneous dynamic range of the human visual system. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, pages 17–24, New York, USA. ACM.

Kuo, G., Waller, L., Ng, R., and Maimone, A. (2020). High resolution étendue expansion for holographic displays. *ACM Transactions on Graphics (TOG)*, 39(4):66–1.

Lai, C. L. and Chang, S. W. (2009). An image processing based visual compensation system for vision defects. *IIH-MSP 2009 - 2009 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 5(1):559–562.

Lai, C. L., Chang, S. W., and Sheen, J. (2009). An integrated portable vision assistant agency for the visual impaired people. In *2009 IEEE International Conference on Control and Automation, ICCA 2009*, pages 2311–2316, New York, USA. IEEE.

Lang, F. and Machulla, T. (2021). Pressing a button you cannot see: Evaluating visual designs to assist persons with low vision through augmented reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10.

Lang, F., Schmidt, A., and Machulla, T. (2020). Augmented reality for people with low vision: symbolic and alphanumeric representation of information. In *International Conference on Computers Helping People with Special Needs*, pages 146–156. Springer.

Langlotz, T., Cook, M., and Regenbrecht, H. (2016). Real-Time Radiometric Compensation for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization & Computer Graphics*, 22(11):2385–2394.

Langlotz, T., Sutton, J., Zollmann, S., Itoh, Y., and Regenbrecht, H. (2018). Chromaglasses: Computational glasses for compensating colour blindness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 1–12, New York, USA. Association for Computing Machinery.

Lanman, D. and Luebke, D. (2013). Near-eye light field displays. *ACM TOG*, 32(6):220.

Lau, C., Perdu, N., Rodríguez-Pardo, C. E., Süsstrunk, S., and Sharma, G. (2015). An interactive app for color deficient viewers. In *Proc. of SPIE-IS&T Electronic Imaging*, page 939512, Bellingham, USA. SPIE.

Lausegger, G., Spitzer, M., and Ebner, M. (2017). Omnicolor–a smart glasses app to support colorblind people. *International Journal Of Interactive Mobile Technologies*, 11(5).

Lee, B., Hong, J.-Y., Jang, C., Jeong, J., and Lee, C.-K. (2017). Holographic and light-field imaging for augmented reality. In *Proc. SPIE*, volume 10125, pages 101251A–1.

Lee, C.-K., Moon, S., Lee, S., Yoo, D., Hong, J.-Y., and Lee, B. (2016). Compact three-dimensional head-mounted display system with savart plate. *Optics Express*, 24(17):19531–19544.

Lee, G.-Y., Hong, J.-Y., Hwang, S., Moon, S., Kang, H., Jeon, S., Kim, H., Jeong, J.-H., and Lee, B. (2018). Metasurface eyepiece for augmented reality. *Nature comm/s*, 9(1):1–10.

Lee, J. and Dos Santos, W. P. (2011). An adaptive fuzzy-based system to simulate, quantify and compensate color blindness. *Integrated Computer-Aided Engineering*, 18(1):29–40.

Lee, J. S., Kim, Y. K., Lee, M. Y., and Won, Y. H. (2019). Enhanced see-through near-eye display using time-division multiplexing of a maxwellian-view and holographic display. *Opt. Express*, 27(2):689–701.

Lee, S., Cho, J., Lee, B., Jo, Y., Jang, C., Kim, D., and Lee, B. (2018). Foveated retinal optimization for see-through near-eye multi-layer displays. *IEEE Access*, 6:2170–2180.

Lee, S., Wang, M., Li, G., Lu, L., Sulai, Y., Jang, C., and Silverstein, B. (2020). Foveated near-eye display for mixed reality using liquid crystal photonics. *Scientific Reports*, 10(1):1–11.

Lincoln, P., Blate, A., Singh, M., State, A., Whitton, M. C., Whitted, T., and Fuchs, H. (2017). Scene-adaptive high dynamic range display for low latency augmented reality. In *Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, I3D '17, pages 15:1–15:7, New York, USA. ACM.

Lincoln, P., Blate, A., Singh, M., Whitted, T., State, A., Lastra, A., and Fuchs, H. (2016a). From Motion to Photons in 80 Microseconds: Towards Minimal Latency for Virtual and Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1367–1376.

Lincoln, P., Blate, A., Singh, M., Whitted, T., State, A., Lastra, A., and Fuchs, H. (2016b). From Motion to Photons in 80 Microseconds: Towards Minimal Latency for Virtual and Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1367–1376.

Lindsay, P. H. and Norman, D. A. (2013). *Human information processing: An introduction to psychology*. Academic press.

Liu, S., Cheng, D., and Hua, H. (2008). An optical see-through head mounted display with addressable focal planes. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2008*, pages 33–42.

Liu, S., Li, Y., Zhou, P., Li, X., Rong, N., Huang, S., Lu, W., and Su, Y. (2016). A multi-plane optical see-through head mounted display design for augmented reality applications. *Journal of the Society for Information Display*, 24(4):246–251.

Livingston, M. A. and Ai, Z. (2008). The effect of registration error on tracking distant augmented objects. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 77–86.

Lu, W., Duh, B. L. H., and Feiner, S. (2012). Subtle cueing for visual search in augmented reality. *ISMAR 2012 - 11th IEEE International Symposium on Mixed and Augmented Reality 2012, Science and Technology Papers*, pages 161–166.

Lu, W., Duh, H. B. L., Feiner, S., and Zhao, Q. (2014). Attributes of subtle cues for facilitating visual search in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 20(3):404–412.

Lu, W., Feng, D., Feiner, S., Zhao, Q., and Duh, H. B. L. (2013). Subtle cueing for visual search in head-tracked head worn displays. *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, pages 271–272.

Luo, G. and Peli, E. (2006). Use of an augmented-vision device for visual search by patients with tunnel vision. *Investigative Ophthalmology and Visual Science*, 47(9):4152–4159.

Ma, Y., Gu, X., and Wang, Y. (2009). Color discrimination enhancement for dichromats using self-organizing color transformation. *Information Sciences*, 179(6):830–843.

MacHado, G. M. and Oliveira, M. M. (2010). Real-time temporal-coherent color contrast enhancement for dichromats. *IEEE-VGTC Symposium on Visualization 2010*, 29(3):933–942.

Machado, G. M., Oliveira, M. M., and Fernandes, L. A. F. (2010). A Physiologically-Based Model for Simulation of Color Vision Deficiency (vol 15, pg 1291, 2009). *Ieee Transactions on Visualization and Computer Graphics*, 16(2):352.

Maimone, A. and Fuchs, H. (2013). Computational augmented reality eyeglasses. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, pages 29–38.

Maimone, A., Georgiou, A., and Kollin, J. S. (2017). Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.*, 36(4):85:1–85:16.

Maimone, A., Lanman, D., Rathinavel, K., Keller, K., Luebke, D., and Fuchs, H. (2014). Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources. In *ACM SIGGRAPH Emerging Technologies*, page 20. ACM.

Maimone, A. and Wang, J. (2020). Holographic optics for thin and lightweight virtual reality. *ACM Transactions on Graphics (TOG)*, 39(4):67–1.

Makibuchi, N., Kato, H., and Yoneyama, A. (2013). Vision-based robust calibration for optical see-through head-mounted displays. In *2013 IEEE International Conference on Image Processing*, pages 2177–2181, New York, USA. IEEE, IEEE.

Mann, S. (1994). Mediated reality.

Mann, S., Lo, R. C. H., Ovtcharov, K., Gu, S., Dai, D., Ngan, C., and Ai, T. (2012). Realtime hdr (high dynamic range) video for eyetap wearable computers, fpga-based seeing aids, and glasseyes (eyetaps). In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–6.

Masia, B., Wetzstein, G., Didyk, P., and Gutierrez, D. (2013). A survey on computational displays: Pushing the boundaries of optics, computation, and perception. *Computers & Graphics*, 37(8):1012–1038.

Mateescu, V. and Bajic, I. (2014). Attention Retargeting by Color Manipulation in Images. *Proceedings of the 1st International Workshop on Perception Inspired Video Processing*, 1.

Mateescu, V. A. and Bajić, I. V. (2014). Can Subliminal Flicker Guide Attention in Natural Images? *Proceedings of the 1st International Workshop on Perception Inspired Video Processing*, pages 33–34.

Matsuda, N., Fix, A., and Lanman, D. (2017). Focal surface displays. *ACM Transactions on Graphics (TOG)*, 36(4):1–14.

Mckendrick, A. and Johnson, C. (2003). *Temporal Properties of Vision*, pages 511–530. Mosby International, United States, 10 edition.

McNamara, A., Bailey, R., and Grimm, C. (2008). Improving search task performance using subtle gaze direction. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08. ACM.

Mechrez, R., Shechtman, E., and Zelnik-Manor, L. (2019). Saliency driven image manipulation. *Machine Vision and Applications*, 30(2):189–202.

Melillo, P., Riccio, D., Di Perna, L., Sanniti Di Baja, G., De Nino, M., Rossi, S., Testa, F., Simonelli, F., and Frucci, M. (2017). Wearable Improved Vision System for Color Vision Deficiency Correction. *IEEE Journal of Translational Engineering in Health and Medicine*, 5:1–7.

Mendez, E., Feiner, S., and Schmalstieg, D. (2010). Focus and context in mixed reality by modulating first order salient features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6133 LNCS:232–243.

Meng, M. and Tanaka, G. (2016). Proposal of Minimization Problem Based Lightness Modification for Protanopia and Deuteranopia. In *2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 1–6, New York, USA. IEEE.

Meng, M. and Tanaka, G. (2019). Proposal of minimization problem based lightness modification method considering visual characteristics of protanopia and deuteranopia. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 1417–1422, New York, USA. IEEE.

Meyer, G. W. and Greenberg, D. P. (1988). Color-Defective Vision and Computer Graphics Displays. *IEEE Computer Graphics and Applications*, 8(5):28–40.

Milić, N., Hoffmann, M., Tómács, T., Novaković, D., and Milosavljević, B. (2015). A Content-Dependent Naturalness-Preserving Daltonization Method for Dichromatic and Anomalous Trichromatic Color Vision Deficiencies. *Journal of Imaging Science and Technology*, 59(1):105041–1050410.

Min Htike, H., H. Margrain, T., Lai, Y.-K., and Eslambolchilar, P. (2021). Augmented Reality Glasses as an Orientation and Mobility Aid for People with Low Vision: a Feasibility Study of Experiences and Requirements. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15.

Mine, M. R. (1993). Characterization of end-to-end delays in head-mounted display systems. Technical report, University of North Carolina, Chapel Hill, NC, USA.

Miyamoto, J., Koike, H., and Amano, T. (2018). Gaze navigation in the real world by changing visual appearance of objects using projector-camera system. *VRST 2018*, pages 1–5.

Mochizuki, R., Nakamura, T., Chao, J., and Lenz, R. (2008). Color-weak correction by discrimination threshold matching. In *Conference on Colour in Graphics, Imaging, and Vision*, pages 208–213, Springfield, USA. Society for Imaging Science and Technology.

Mochizuki, R., Oshima, S., and Chao, J. (2011a). Fast color-weakness compensation with discrimination threshold matching. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 6626 LNCS, pages 176–187, Lectur Notes in Computer Science. SpringerLink.

Mochizuki, R., Oshima, S., Lenz, R., and Chao, J. (2011b). Exact compensation of color-weakness with discrimination threshold matching. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6768 LNCS(4):155–164.

Moon, E., Kim, M., Roh, J., Kim, H., and Hahn, J. (2014). Holographic head-mounted display with rgb light emitting diode light source. *Optics express*, 22(6):6526–6534.

Mori, S., Ikeda, S., Plopski, A., and Sandor, C. (2018). Brightview: Increasing perceived brightness of optical see-through head-mounted displays through unnoticeable incident light reduction. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 251–258.

Moser, K., Itoh, Y., Oshima, K., Swan, J. E., Klinker, G., and Sandor, C. (2015). Subjective evaluation of a semi-automatic optical see-through head-mounted display calibration technique. *IEEE transactions on visualization and computer graphics*, 21(4):491–500.

Mukawa, H., Akutsu, K., Matsumura, I., Nakano, S., Yoshida, T., Kuwahara, M., Aiki, K., and Ogawa, M. (2008). 8.4: Distinguished paper: A full color eyewear display using holographic planar waveguides. *SID Symposium Digest of Technical Papers*, 39(1):89–92.

Mulloni, A., Seichter, H., and Schmalstieg, D. (2011). Handheld augmented reality indoor navigation with activity-based instructions. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, page 211–220, New York, USA. Association for Computing Machinery.

Nagahara, H., Yagi, Y., and Yachida, M. (2006). Super wide field of view head mounted display using catadioptrical optics. *Presence: Teleoper. Virtual Environ.*, 15(5):588–598.

Narain, R., Albert, R. A., Bulbul, A., Ward, G. J., Banks, M. S., and O'Brien, J. F. (2015). Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Trans. Graph.*, 34(4).

Ng, A., Lepinski, J., Wigdor, D., Sanders, S., and Dietz, P. (2012). Designing for low-latency direct-touch input. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pages 453–464, New York, USA. ACM.

Nguyen, T. V., Ni, B., Liu, H., Xia, W., Luo, J., Kankanhalli, M., and Yan, S. (2013). Image re-attentionizing. *IEEE Transactions on Multimedia*, 15(8):1910–1919.

Nigam, P. K. and Bhattacharya, M. (2013). Colour vision deficiency correction in image processing. In *Proceedings - 2013 IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2013*, pages 1–8, New York, USA. IEEE.

Nooney Jr., T. W. (1972). An optical approach to aid cerebral hemiplegics. *MCV/Q, Medical College of Virginia Quarterly*, 8(4):274–277.

Noorlander, C., Koenderink, J. J., Olden, R. J. D., and Edens, B. W. (1983). Sensitivity to spatiotemporal colour contrast in the peripheral visual field. *Vision Research*, 23(1):1 – 11.

Nowak, A., Wozniak, M., Pieprzowski, M., and Romanowski, A. (2018). Towards amblyopia therapy using mixed reality technology. In *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 279–282. IEEE.

Oda, O., Elvezio, C., Sukan, M., Feiner, S., and Tversky, B. (2015). Virtual replicas for remote assistance in virtual and augmented reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, UIST '15, page 405–415, New York, USA. Association for Computing Machinery.

Ohkubo, T. and Kobayashi, K. (2008). A color compensation vision system for color-blind people. In *Proceedings of the SICE Annual Conference*, pages 1286–1289, New York, USA. IEEE.

Oka, M., Ozawa, N., Mori, H., and Sakuri, A. (2010). A Study on Color Conversion for the anomalous trichromat to Identify Color. In *Proceedings of SICE Annual Conference 2010*, pages 1441–1443, New York, USA. IEEE.

Olsen, A. (2012). The tobii i-vt fixation filter. *Tobii Technology*, pages 1–21.

Orii, H., Kawano, H., Maeda, H., and Kouda, T. (2014). Color conversion algorithm for color blindness using self-organizing map. In *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems, SCIS 2014 and 15th International Symposium on Advanced Intelligent Systems, ISIS 2014*, pages 910–913, New York, USA. IEEE.

Orlosky, J. (2014). Depth based interaction and field of view manipulation for augmented reality. *Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology - UIST'14 Adjunct*, pages 5–8.

Orlosky, J., Kim, P., Kiyokawa, K., Mashita, T., Ratsamee, P., Uranishi, Y., and Takemura, H. (2017). Vismerge: Light adaptive vision augmentation via spectral and temporal fusion of non-visible light. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 22–31.

Oshima, S., Mochizuki, R., Lenz, R., and Chao, J. (2016). Modeling, Measuring, and Compensating Color Weak Vision. *IEEE Transactions on Image Processing*, 25(6):2587–2600.

Ostia, C. F., Padilla, D. A., Reidj, F., Cruz, G., Galang, R. J. D., Josafat, A. S. M., and Victoria, E. F. (2019). Electronic vision system with personalized calibration of color compensation for people with partial color vision deficiency using raspberry pi digital image processing. In *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, pages 315–319, New York, USA. IEEE.

Otao, K., Itoh, Y., Osone, H., Takazawa, K., Kataoka, S., and Ochiai, Y. (2017). Light field blender: designing optics and rendering methods for see-through and aerial near-eye display. In *SIGGRAPH Asia 2017 Technical Briefs*, pages 1–4. ACM.

Owen, C. B., Zhou, J., Tang, A., and Xiao, F. (2004). Display-relative calibration for optical see-through head-mounted displays. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2004*, pages 70–78.

Padmanaban, N., Konrad, R., and Wetzstein, G. (2019a). Autofocals: Evaluating gaze-contingent eyeglasses for presbyopes. *Science Advances*, 5(6):eaav6187.

Padmanaban, N., Peng, Y., and Wetzstein, G. (2019b). Holographic near-eye displays based on overlap-add stereograms. *ACM Transactions on Graphics (TOG)*, 38(6):1–13.

Pal, R. and Roy, D. (2018). Enhancing Saliency of an Object Using Genetic Algorithm. *Proceedings - 2017 14th Conference on Computer and Robot Vision, CRV 2017*, 2018-January:337–344.

Park, J., Choi, J., and Han, D. (2011). Applying enhanced confusion line color transform using color segmentation for mobile applications. In *Proceedings - 1st ACIS/JNU International Conference on Computers, Networks, Systems, and Industrial Engineering, CNSI 2011*, pages 40–44, New York, USA. IEEE.

Parkhurst, D., Law, K., and Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1):107 – 123.

Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. (2016). Towards foveated rendering for gaze-tracked virtual reality. *ACM Trans. Graph.*, 35(6):179:1–179:12.

Peli, E. (2007). Vision multiplexing: an optical engineering concept for low-vision aids. In *Current Developments in Lens Design and Optical Engineering VIII*, volume 6667, page 66670C. International Society for Optics and Photonics.

Peli, E. and Jung, J. H. (2017). Multiplexing prisms for field expansion. *Optometry and Vision Science*, 94(8):817–829.

Peli, E., Luo, G., Bowers, A., and Rensing, N. (2009). Development and evaluation of vision multiplexing devies for vision impairments. *International journal of artificial intelligence tools : architectures, languages, algorithms*, 18(3):365–378.

Peng, Y., Choi, S., Padmanaban, N., and Wetzstein, G. (2020). Neural holography with camera-in-the-loop training. *ACM Transactions on Graphics*, 39(6).

Plopski, A., Itoh, Y., Nitschke, C., Kiyokawa, K., Klinker, G., and Takemura, H. (2015). Corneal-imaging calibration for optical see-through head-mounted displays. *IEEE transactions on visualization and computer graphics*, 21(4):481–490.

Poret, S., Dony, R. D., and Gregori, S. (2009). Image processing for colour blindness correction. In *TIC-STH'09: 2009 IEEE Toronto International Conference - Science and Technology for Humanity*, pages 539–544, New York, USA. IEEE.

Poupyrev, I., Billinghurst, M., Kato, H., and May, R. (2000). Integrating real and virtual worlds in shared space. In *Proceedings of the 5th International Symposium on Artificial Life and Robotics*, pages 22–25.

Pucihar, K. v. and Coulton, P. (2015). Exploring the evolution of mobile augmented reality for future entertainment systems. *Comput. Entertain.*, 11(2).

Pullin, G. (2009). Fashion meets discretion. In *Design Meets Disability*, pages 16–64. MIT Press, Cambridge.

Pundlik, S., Yi, H., Liu, R., Peli, E., and Luo, G. (2017). Magnifying smartphone screen using google glass for low-vision Users. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(1):49–58.

Ragan, E., Wilkes, C., Bowman, D. A., and Hollerer, T. (2009). Simulation of augmented reality systems in purely virtual environments. In *2009 IEEE VR*, pages 287–288.

Rasche, K., Geist, R., and Westall, J. (2005a). Detail preserving reproduction of color images for monochromats and dichromats. *IEEE Computer Graphics and Applications*, 25(3):22–30.

Rasche, K., Geist, R., and Westall, J. (2005b). Re-coloring images for gamuts of lower dimension. *Computer Graphics Forum*, 24(3):423–432.

Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., and Forlines, C. (2003). Ilamps: Geometrically aware and self-configuring projectors. *ACM Trans. Graph.*, 22(3):809–818.

Rathinavel, K., Wang, H., Blate, A., and Fuchs, H. (2018). An extended depth-at-field volumetric near-eye augmented reality display. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2857–2866.

Rathinavel, K., Wetzstein, G., and Fuchs, H. (2019). Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics. *IEEE transactions on visualization and computer graphics*, 25(11):3125–3134.

Regenbrecht, H., Zwanenburg, S., and Langlotz, T. (2022). Pervasive augmented reality—technology and ethics. *IEEE Pervasive Computing*, pages 1–8.

Reitmayr, G., Langlotz, T., Wagner, D., Mulloni, A., Schall, G., Schmalstieg, D., and Pan, Q. (2010). Simultaneous localization and mapping for augmented reality. In *Ubiquitous Virtual Reality (ISUVR), 2010 International Symposium on*, pages 5–8.

Renner, P. and Pfeiffer, T. (2017). Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*, pages 186–194.

Ribeiro, M. G. and Gomes, A. J. (2013). A skillet-based recoloring algorithm for dichromats. In *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services, Healthcom 2013*, pages 702–706, New York, USA. IEEE.

Robert, M. (2002). *Display Interfaces:Fundamentals and Standards.* John Wiley & Sons.

Roglic, G. (2016). WHO Global report on diabetes: A summary. *International Journal of Noncommunicable Diseases*, 1(1):3–8.

Rolland, J. P., Krueger, M. W., and Goon, A. (2000). Multifocal planes head-mounted displays. *Appl. Opt.*, 39(19):3209–3215.

Rolland, J. P., Yoshida, A., Davis, L. D., and Reif, J. H. (1998). High-resolution inset head-mounted display. *Appl. Opt.*, 37(19):4183–4193.

Rothe, S., Buschek, D., and Hußmann, H. (2019). Guidance in cinematic virtual reality-taxonomy, research status and challenges. *Multimodal Technologies and Interaction*, 3(1).

Rumiński, J., Wtorek, J., Rumińska, J., Kaczmarek, M., Bujnowski, A., Kocejko, T., and Poliński, A. (2010). Color transformation methods for dichromats. In *3rd International Conference on Human System Interaction*, pages 634–641, New York, USA. IEEE.

Sajadi, B., Majumder, A., Oliveira, M. M., Schneider, R. G., and Raskar, R. (2013). Using patterns to encode color information for dichromats. *IEEE Transactions on Visualization and Computer Graphics*, 19(1):118–129.

Santos, P., Gierlinger, T., Machui, O., and Stork, A. (2008). The daylight blocking optical stereo see-through hmd. In *Proceedings of the 2008 workshop on Immersive projection technologies/Emerging display technologiges*, page 4. ACM.

Sayed, A. M., Abdel-Mottaleb, M., Kashem, R., Roongpoovapatr, V., Elsawy, A., Abdel-Mottaleb, M., Parrish II, R. K., and Abou Shousha, M. (2020). Expansion of peripheral visual field with novel virtual reality digital spectacles. *American journal of ophthalmology*, 210:125–135.

Schmitt, S., Stein, S., Hampe, F., and Paulus, D. (2012). Mobile services supporting color vision deficiency. In *2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, pages 1413–1420, New York, USA. IEEE.

Schmitz, A., MacQuarrie, A., Julier, S., Binetti, N., and Steed, A. (2020). Directing versus attracting attention: Exploring the effectiveness of central and peripheral cues in panoramic videos. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 63–72.

Schwerdtfeger, B. and Klinker, G. (2008). Supporting order picking with augmented reality. *Proceedings - 7th IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008*, pages 91–94.

Schwerdtfeger, B., Reif, R., Gunthner, W. A., Klinker, G., Hamacher, D., Schega, L., Bockelmann, I., Doil, F., and Tumler, J. (2009). Pick-by-vision: A first stress test. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pages 115–124.

Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A., and Vorozcovs, A. (2004). High dynamic range display systems. *ACM Trans. Graph.*, 23(3):760–768.

Sharpe, L. T., Stockman, A., Jägle, H., and Nathans, J. (1999). *Opsin genes, cone photopigments, color vision, and color blindness*, pages 3–51. Cambridge U. Press, Cambridge, England.

Shen, W., Mao, X., Hu, X., and Wong, T.-T. (2016). Seamless visual sharing with color vision deficiencies. *ACM Transactions on Graphics*, 35(4):1–12.

Shi, L., Huang, F.-C., Lopes, W., Matusik, W., and Luebke, D. (2017). Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics. *ACM Trans. Graph.*, 36(6):236:1–236:17.

Shi, T. and Sugimoto, A. (2014). Video Saliency Modulation in the HSI Color Space for Drawing Gaze. *PSIVT*, 8333(July 2015).

Singh, R., Bailey, S., Chang, P., Olyaei, A., Hekmat, M., and Winoto, R. (2021). 34.2 a 21pj/frame/pixel imager and 34pj/frame/pixel image processor for a low-vision augmented-reality smart contact lens. In *2021 IEEE International Solid-State Circuits Conference (ISSCC)*, volume 64, pages 482–484.

Sivakumar, P., Vedachalam, R., Kannusamy, V., Odayappan, A., Venkatesh, R., Dhoble, P., Moutappa, F., and Narayana, S. (2020). Barriers in utilisation of low vision assistive products. *Eye*, 34(2):344–351.

Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., Watson, Z., Breen, G., Schwarz, R., Steptoe, W., et al. (2020). The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality*, 1:1.

Song, W., Wang, Y., Cheng, D., and Liu, Y. (2014). Light field head-mounted display with correct focus cue using micro structure array. *Chinese Optics Letters*, 12(6).

Spalding, J. (1999). Colour vision deficiency in the medical profession. *British journal of general practice*, 49(443):469–475.

Spjut, J., Boudaoud, B., Kim, J., Greer, T., Albert, R., Stengel, M., Akşit, K., and Luebke, D. (2020). Toward standardized classification of foveated displays. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2126–2134.

Sprague, R. (2010). Method and apparatus to process display and non-display information. WO Patent App. PCT/US2009/055,758.

Sridharan, S. K., Hincapié-Ramos, J. D., Flatla, D. R., and Irani, P. (2013). Color correction for optical see-through displays using display color profiles. In *VRST*, pages 231–240.

Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., Picard, R. W., and Pentland, A. (1997). Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments*, 6(4).

Stearns, L., DeSouza, V., Yin, J., Findlater, L., and Froehlich, J. E. (2017). Augmented reality magnification for low vision users with the microsoft hololens and a finger-worn camera. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '17, page 361–362, New York, USA. Association for Computing Machinery.

Stearns, L., Findlater, L., and Froehlich, J. E. (2018). Design of an Augmented Reality Magnification Aid for Low Vision Users. *Proc. ASSETS 2018 (To Appear)*, pages 28–39.

Steward, J. M. and Cole, B. L. (1989a). What do color vision defectives say about everyday tasks? *Optometry and Vision Science*, 66(5):288–295.

Steward, J. M. and Cole, B. L. (1989b). What do color vision defectives say about everyday tasks? *Optometry and Vision Science*, 66(5):288–295.

Strasburger, H. and Pöppel, E. (2002). Visual field.

Su, S. L., Durand, F., and Agrawala, M. (2005). De-emphasis of distracting image regions using texture power maps. *Proceedings of the 2nd symposium on Appied perception in graphics and visualization - APGV '05*, page 164.

Suetake, N., Tanaka, G., Hashii, H., and Uchino, E. (2012). Simple lightness modification for color vision impaired based on Craik-O'Brien effect. *Journal of the Franklin Institute*, 349(6):2093–2107.

Sutherland, I. E. (1968). A head-mounted three dimensional display. *Fall Joint Computer Conference*, pages 757–764.

Sutton, J., Langlotz, T., and Itoh, Y. (2019). Computational glasses: Vision augmentations using computational near-eye optics and displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 438–442, Beijing, China. IEEE.

Sutton, J., Langlotz, T., and Plopski, A. (2022a). Seeing colours: Addressing colour vision deficiency with vision augmentations using computational glasses. *ACM Transactions on Computer-Human Interaction*, 29(3).

Sutton, J., Langlotz, T., and Plopski, A. (TBD). A visual guidance design space for on-screen ar. *Completing Final Studies*, 0(0).

Sutton, J., Langlotz, T., Plopski, A., Zollmann, S., Itoh, Y., and Regenbrecht, H. (2022b). Look over there! investigating saliency modulation for visual guidance with augmented reality glasses. In *The 35th Annual ACM Symposium on User Interface Software and Technology*, New York, USA. Association for Computing Machinery.

Suyama, S., Date, M., and Takada, H. (2000). Three-dimensional display system with dual-frequency liquid-crystal varifocal lens. *Japanese Journal of Applied Physics*, 39(2R):480.

Suzuki, N. and Nakada, Y. (2018). Effects selection technique for improving visual attraction via visual saliency map. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings*, 2018-January:1–8.

Swarbrick, H. A., Nguyen, P., Nguyen, T., and Pham, P. (2001). The chromagen contact lens system: colour vision test results and subjective responses. *Ophthalmic and Physiological Optics*, 21(3):182–196.

Takimoto, H., Hitomi, S., Yamauchi, H., Kishihara, M., and Okubo, K. (2017). Image modification based on spatial frequency components for visual attention retargeting. *IEICE Transactions on Information and Systems*, 100(6):1339–1349.

Takimoto, H., Kokui, T., Yamauchi, H., Kishihara, M., and Okubo, K. (2015). Image modification based on a visual saliency map for guiding visual attention. *IEICE Transactions on Information and Systems*, 98(11):1967–1975.

Takimoto, H., Yamamoto, K., Kanagawa, A., Kishihara, M., and Okubo, K. (2019). Attention retargeting using saliency map and projector–camera system in real space. *IEEJ Transactions on Electrical and Electronic Engineering*, 14(6):853–861.

Talom, O., P, N., and A., A. (2016). Efficient Image Re-coloring Mechanism for Color Vision Deficient (CVD) Persons. *Elysium Journal of Engineering Research & Management*, 3(1):1–6.

Tanaka, G., Suetake, N., and Uchino, E. (2009). Visibility improvement of undiscriminatable colors by modification of yellow-blue components for protanopia and deuteranopia. In *ISPACS 2009 - 2009 International Symposium on Intelligent Signal Processing and Communication Systems, Proceedings*, pages 562–565, New York, USA. IEEE.

Tanaka, G., Suetake, Noriaki, and Uchino, E. (2010). Lightness Modification of Color Image for Protanopia and Deuteranopia. *Optical Review*, 17(1):14–23.

Tang, Y., Zhu, Z., Toyoura, M., Go, K., Kashiwagi, K., Fujishiro, I., and Mao, X. (2018). Arriving light control for color vision deficiency compensation using optical see-through head-mounted display. In *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '18, New York, USA. Association for Computing Machinery.

Tang, Y., Zhu, Z., Toyoura, M., Go, K., Kashiwagi, K., Fujishiro, I., and Mao, X. (2020). Alcc-glasses: Arriving light chroma controllable optical see-through head-mounted display system for color vision deficiency compensation. *Applied Sciences*, 10(7).

Tanuwidjaja, E., Huynh, D., Koa, K., Nguyen, C., Shao, C., Torbett, P., Emmenegger, C., and Weibel, N. (2014). Chroma: AWearable Augmented-Reality Solution for Color Blindness. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, pages 799–810, New York, USA. Association for Computing Machinery.

Tatzgern, M., Kalkofen, D., Grasset, R., and Schmalstieg, D. (2014). Hedgehog labeling: View management techniques for external labels in 3d space. In *2014 IEEE Virtual Reality (VR)*, pages 27–32.

Tatzgern, M., Orso, V., Kalkofen, D., Jacucci, G., Gamberini, L., and Schmalstieg, D. (2016). Adaptive information density for augmented reality displays. In *2016 IEEE Virtual Reality (VR)*, pages 83–92.

Torralba, A., Oliva, A., Castelhano, M., and Henderson, J. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Pyschol Rev*, pages 766–786.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1):97 – 136.

Troiano, L., Birtolo, C., and Miranda, M. (2008). Adapting palettes to color vision deficiencies by genetic algorithm. In *GECCO'08: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation 2008*, pages 1065–1072, New York, USA. Association for Computing Machinery.

Tuceryan, M., Genc, Y., and Navab, N. (2002). Single-point active alignment method (spaam) for optical see-through hmd calibration for augmented reality. *Presence: Teleoper. Virtual Environ.*, 11(3):259–276.

Uchida, T., Sato, K., and Inokuchi, S. (2002). An optical see-through mr display with digital micro-mirror device. *Transactions of the Virtual Reality Society of Japan*, 7(2).

Ueda, T., Iwai, D., Hiraki, T., and Sato, K. (2020). Illuminated focus: Vision augmentation using spatial defocusing via focal sweep eyeglasses and high-speed projector. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2051–2061.

Veas, E., Mendez, E., Feiner, S. K., and Schmalstieg, D. (2011). Directing attention and influencing memory with visual saliency modulation. *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 1471.

Verriest, G., Neubauer, O., Marre, M., and Uvijls, A. (1980). New investigations concerning the relationships between congenital colour vision defects and road traffic security. *International ophthalmology*, 2(2):87–99.

Viénot, F., Brettel, H., and Mollon, J. D. (1999). Digital video colourmaps for checking the legibility of displays by dichromats. *Color Research and Application*, 24(4):243–252.

Volmer, B., Baumeister, J., Von Itzstein, S., Bornkessel-Schlesewsky, I., Schlesewsky, M., Billinghurst, M., and Thomas, B. H. (2018). A comparison of predictive spatial augmented reality cues for procedural tasks. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2846–2856.

Wagner, A. and Rozenblit, J. W. (2017). Augmented reality visual guidance for spatial perception in the computer assisted surgical trainer. In *Proceedings of the Symposium on Modeling and Simulation in Medicine*, MSM '17. Society for Computer Simulation International.

Wagner, D., Langlotz, T., and Schmalstieg, D. (2008). Robust and unobtrusive marker tracking on mobile phones. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, pages 121–124, Washington, DC, USA. IEEE Computer Society.

Wakita, K. and Shimamura, K. (2005). SmartColor: disambiguation framework for the colorblind. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility - Assets '05*, pages 158–165, New York, USA. Association for Computing Machinery.

Waldin, N., Waldner, M., and Viola, I. (2017). Flicker Observer Effect: Guiding Attention Through High Frequency Flicker in Images. *Computer Graphics Forum*, 36(2):467–476.

Waldner, M., Le Muzic, M., Bernhard, M., Purgathofer, W., and Viola, I. (2014). Attractive flicker-Guiding attention in dynamic narrative visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2456–2465.

Wandell, B. (1995). *Foundations of Vision.* Sinauer Associates.

Wang, W., Shen, J., and Shao, L. (2018). Video salient object detection via fully convolutional networks. *IEEE Transactions on Image Processing*, 27(1):38–49.

Wang, X., Zhu, Z., Chen, X., Toyoura, M., and Mao, X. (2020). Evaluation of color vision compensation algorithms for people with varying degrees of color vision deficiency. In *2020 International Conference on Cyberworlds (CW)*, pages 149–152, Caen, France. IEEE.

Wassom, B. (2014). *Augmented reality law, privacy, and ethics: Law, society, and emerging AR technologies*. Syngress.

Weiland, C., Braun, A., and Heiden, W. (2009). Colorimetric and photometric compensation for optical see-through displays. *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, pages 603–612.

Wetzstein, G., Heidrich, W., and Luebke, D. (2010). Optical image processing using light modulation displays. *Computer Graphics Forum*, 29(6):1934–1944.

Wheatstone, C. (1838). Xviii. contributions to the physiology of vision.—part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical transactions of the Royal Society of London*, 1(128):371–394.

Wilson, A. and Hua, H. (2017). Design and prototype of an augmented reality display with per-pixel mutual occlusion capability. *Optics express*, 25(24):30539–30549.

Wilson, A. and Hua, H. (2019). Design and demonstration of a vari-focal optical see-through head-mounted display using freeform alvarez lenses. *Optics express*, 27(11):15627–15637.

Wolfe, J. M. and Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6):495–501.

Wong, L. K. and Low, K. L. (2011). Saliency retargeting: An approach to enhance image aesthetics. *2011 IEEE Workshop on Applications of Computer Vision, WACV 2011*, 2(c):73–80.

Wood, J. M., Lacherez, P., Black, A. A., Cole, M. H., Boon, M. Y., and Kerr, G. K. (2011). Risk of Falls, Injurious Falls, and Other Injuries Resulting from Visual Impairment among Older Adults with Age-Related Macular Degeneration. *Investigative Ophthalmology and Visual Science*, 52(8):5088–5092.

Woźniak, P., Knaving, K., Obaid, M., Carcedo, M. G., Ünlüer, A., and Fjeld, M. (2015). Chromaglove: A wearable haptic feedback device for colour recognition. In *Proceedings of the 6th Augmented Human International Conference*, AH '15, page 219–220, New York, USA. Association for Computing Machinery.

Xia, X., Guan, Y., State, A., Chakravarthula, P., Rathinavel, K., Cham, T.-J., and Fuchs, H. (2019). Towards a switchable ar/vr near-eye display with accommodation-vergence and eyeglass prescription support. *IEEE transactions on visualization and computer graphics*, 25(11):3114–3124.

Xiao, L., Kaplanyan, A., Fix, A., Chapman, M., and Lanman, D. (2018). Deepfocus: Learned image synthesis for computational displays. *ACM Trans. Graph.*, 37(6).

Xu, M. and Hua, H. (2017). High dynamic range head mounted display based on dual-layer spatial modulation. *Opt. Express*, 25(19):23320–23333.

Yamaguchi, Y. and Takaki, Y. (2016). See-through integral imaging display with background occlusion capability. *Applied optics*, 55(3):A144–A149.

Yang, S. and Ro, Y. M. (2003). Visual contents adaptation for color vision deficiency. In *ICIP 2003: Proceedings of International Conference on Image Processing*, volume 1, pages 453–456, New York, USA. IEEE.

Yang, S., Ro, Y. M., Nam, J., Hong, J., Choi, S. Y., and Lee, J. H. (2004). Improving visual accessibility for color vision deficiency based on MPEG-21. *ETRI Journal*, 26(3):195–202.

Yang, S., Ro, Y. M., Wong, E. K., and Lee, J.-H. (2005). Color compensation for anomalous trichromats based on error score of FM-100 Hue test. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, volume 27, pages 6571–6574, New York, USA. IEEE.

Yang, S., Ro, Y. M., Wong, E. K., and Lee, J. H. (2008). 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*, 2008:1–12.

Yano, Y., Orlosky, J., Kiyokawa, K., and Takemura, H. (2016). Dynamic view expansion for improving visual search in video see-through ar. In *Proceedings of the 26th International Conference on Artificial Reality and Telexistence and the 21st Eurographics Symposium on Virtual Environments*, ICAT-EGVE '16, page 57–60, Goslar, DEU. Eurographics Association.

Yanoff, M. and Duker, J. (2018). *Ophthalmology*. Elsevier Saunders, Edinburgh, 5th edition.

Yaraş, F., Kang, H., and Onural, L. (2010). State of the art in holographic displays: a survey. *Journal of display technology*, 6(10):443–454.

Yaraş, F., Kang, H., and Onural, L. (2011). Circular holographic video display system. *Optics Express*, 19(10):9147–9156.

Yoo, C., Bang, K., Jang, C., Kim, D., Lee, C.-K., Sung, G., Lee, H.-S., and Lee, B. (2019). Dual-focal waveguide see-through near-eye display with polarization-dependent lenses. *Optics letters*, 44(8):1920–1923.

Yoshimura, A., Khokhar, A., and Borst, C. W. (2019). Visual cues to restore student attention based on eye gaze drift, and application to an offshore training system. In *Symposium on Spatial User Interaction*, SUI '19, New York, USA. Association for Computing Machinery.

Young, C. A. (1929). Homonymous hemianopsia during pregnancy aided by reflecting prism. *Archives of Ophthalmology*, 2(5):560–565.

Younis, O., Al-Nuaimy, W., Al-Taee, M. A., and Al-Ataby, A. (2017). Augmented and virtual reality approaches to help with peripheral vision loss. *2017 14th International Multi-Conference on Systems, Signals and Devices, SSD 2017*, 2017-Janua:303–307.

Yuen, S. C.-Y., Yaoyuneyong, G., and Johnson, E. (2011). Augmented reality: An overview and five directions for ar in education. *Journal of Educational Technology Development and Exchange (JETDE)*, 4(1):11.

Zhang, L., Xu, Q., Zhu, G., Song, J., Zhang, X., Shen, P., Wei, W., Shah, S. A. A., and Bennamoun, M. (2018). Improved colour-to-grey method using image segmentation and colour difference model for colour vision deficiency. *IET Image Processing*, 12(3):314–319.

Zhang, Y., Hu, X., Kiyokawa, K., Isoyama, N., Sakata, N., and Hua, H. (2021a). Optical see-through augmented reality displays with wide field of view and hard-edge occlusion by using paired conical reflectors. *Optics letters*, 46(17):4208–4211.

Zhang, Y., Hu, X., Kiyokawa, K., Isoyama, N., Uchiyama, H., and Hua, H. (2021b). Realizing mutual occlusion in a wide field-of-view for optical see-through augmented reality displays based on a paired-ellipsoidal-mirror structure. *Optics Express*, 29(26):42751–42761.

Zhao, X., Go, K., Kashiwagi, K., Toyoura, M., Mao, X., and Fujishiro, I. (2019a). Computational alleviation of homonymous visual field defect with ost-hmd: the effect of size and position of overlaid overview window. In *2019 International Conference on Cyberworlds (CW)*, pages 175–182. IEEE.

Zhao, Y., Matsuda, N., Wang, X., Zannoli, M., and Lanman, D. (2020). High dynamic range near-eye displays. In Kress, B. C. and Peroz, C., editors, *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR)*, volume 11310, pages 268 – 279. SPIE.

Zhao, Y., Szpiro, S., and Azenkot, S. (2015). ForeSee : A Customizable Head-Mounted Vision Enhancement System for People with Low Vision. *ASSETS '15 The 17th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 239–249.

Zhao, Y., Szpiro, S., Knighten, J., and Azenkot, S. (2016). CueSee. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*, pages 73–84.

Zhao, Y., Szpiro, S., Shi, L., and Azenkot, S. (2019b). Designing and evaluating a customizable head-mounted vision enhancement system for people with low vision. *ACM Trans. Access. Comput.*, 12(4).

Zheng, D. D., Christ, S. L., Lam, B. L., Arheart, K. L., Galor, A., and Lee, D. J. (2012). Increased mortality risk among the visually impaired: the roles of mental well-being and preventive care practices. *Investigative ophthalmology and visual science*, 53(6):2685–2692.

Zheng, F., Whitted, T., Lastra, A., Lincoln, P., State, A., Maimone, A., and Fuchs, H. (2014). Minimizing latency for augmented reality displays: Frames considered harmful. In *IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2014, Munich, Germany, September 10-12, 2014*, pages 195–200.

Acronyms

AR augmented reality. i, 3–8, 11, 13–15, 22, 23, 28, 34–36, 44, 45, 89, 97–109, 111, 113, 116, 120, 122, 133–135, 137, 139, 141–144, 150

CFF critcal flicker frequency. 18

cpd cycles per degree. 18, 32

CRT cathode-ray tube. 23

CVD colour vision deficiency. i, viii, 1, 2, 6–12, 21, 25–28, 47–57, 60–64, 69, 70, 72, 74–82, 84, 86–95, 98, 142, 147–150

CVR cinematic virtual reality. 103, 104

DMD digital micromirror device. 28, 35, 37, 38

FoV Field of View. 1, 2, 5, 18, 21, 24–29, 32, 36, 38–42, 44, 145, 147, 149

HCI Human-Computer Interaction. 139, 141, 144, 150

HDR high dynamic range. 35

HMD head-mounted display. 3, 22, 23, 26, 33–35, 39, 41, 42, 57

HOE holographic optical element. 24, 42, 146

HVS human visual system. i, 1, 2, 4–8, 11, 14–16, 18, 19, 21, 25, 26, 28–32, 37, 38, 42, 44, 49, 53, 98, 138, 139, 141, 142, 149, 151

INDICA interaction-free display calibration. 34, 150

IPD interpupillary distance. 19

IR infra-red. 26

LCD liquid-crystal display. 1, 25, 27, 37, 38, 40

LUT look-up table. 36

OST optical see-through. 37

OSTHMD optical see-through head-mounted display. i, 3–9, 11–14, 16, 18, 19, 22–25, 27–46, 48, 56, 58, 80, 89, 90, 94, 97, 99, 100, 102, 104–113, 115–117, 122–124, 131, 133, 134, 136–139, 141–143, 145–148, 151

PSF point spread function. 34

PSLM phased spatial-light modulator. 25, 27, 37

SAR spatial augmented reality. 3, 23, 28, 36, 41, 100–102, 104–107, 109

SGD Subtle Gaze Direction. 101, 102

SLAM simultaneous localisation and mapping. 33

SLM spatial-light modulator. 31, 37, 38, 40, 41

SPAAM single-point active alignment method. 34, 58, 59

UV ultra-violet. 1

VR virtual reality. 3, 35, 39–41, 100–105, 107, 113, 117, 118, 123, 125

VST video see-through. 44

VSTHMD video see-through head-mounted display. 3, 5, 23, 26, 28, 29, 31, 36, 44, 100–102, 104–107, 137, 147

WHO World Health Organization. i, 2

Glossary

accommodation–vergence conflict The accommodation–vergence conflict is a visual issue with single focal-plane displays that present the user stereoscopic content. When rendering the virtual content in the real world, depth can be achieved using stereoscopic vision, and causing the eyes to converge at the desired depth. However the light from the display is still being focused on a set plane, irregardless of the stereoscopic depth and the eyes must adjust their accommodation to this. When the depth on the virtual content does not match the depth of the focal plane there is a conflict between the accommodation and the convergence of the eyes. 25, 38, 39, 41, 43

eyebox The eyebox is the volume relative to a display in which an eye must be placed in order to properly view the presented image. Similar to a viewing angle of a monitor, only for most optical see through display designs these are small areas just in front of the display that the eye must be placed in to properly see the content rendered. 24

half-silvered mirror A half-silvered mirror is a piece of glass with a coating applied to it that partially reflects the light striking the glass whilst allowing the rest through. This enables the ability to reflect light onto or away from a path whilst still allowing light on the path already to continue. 24, 27, 39, 41

low vision Low vision is a general term for reduced sight or vision that in turn reduces ones ability to complete everyday tasks. Sight loss in low vision is considered to be untreatable and unassisted. Low vision can be caused by a number of separate or overlapping causes and impacts all parts of the lives of those who have to live with it. 26, 27, 144

Appendix A

CVD Image Set

Figure A.1: CVD test images pt.1

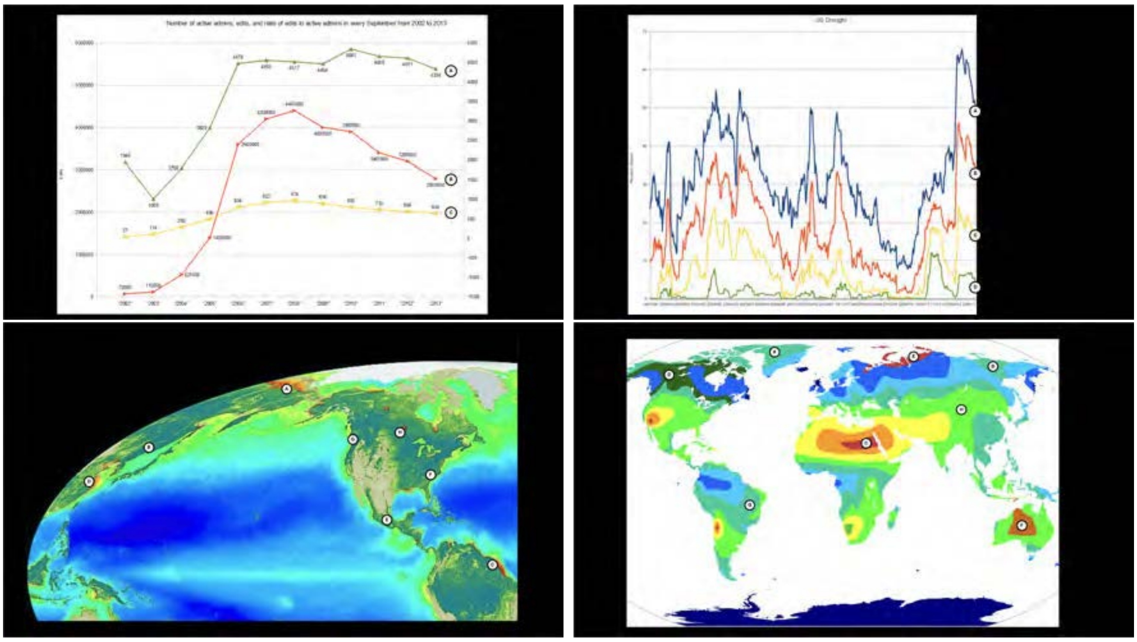


Figure A.2: CVD test images pt.2

Figure A.3: CVD test images pt.3

Appendix B

CVD Explorative Questionnaire

1. Have you been part of our previous study?
    - If yes:
        - What did you see when you looked at the poster?
        - What was your impression?
        - How did what you saw today compare to what you have seen the last time?
    - If no:
        - What did you see when you looked at the poster?
        - What was your impression?
        - What do you think we were trying to do?
2. Please take a look at the poster again. How does it compare to what you saw earlier?
3. If they say nothing about CG: Was it easier to distinguish and answer questions with our prototype? Please explain the reason for your answer.
4. If you could fix your CVD, would you consider doing so? What procedure/ device would you consider, e.g., surgery?
5. Do you wear vision correction?
    - If yes: Would you like your glasses to include functionality that assists your CVD similar to what you experienced today?
    - If no: If you could wear glasses that could help you distinguish between colors, would you consider wearing them?
6. What situations in daily like do you feel this technology could assist you with?
7. Who do you think could benefit from such technology the most?
8. What would you change to make this visualization better address your needs?
9. Any other feedback you could give us?

Appendix C

Visual Guidance Image Sets

C.1 Study 1

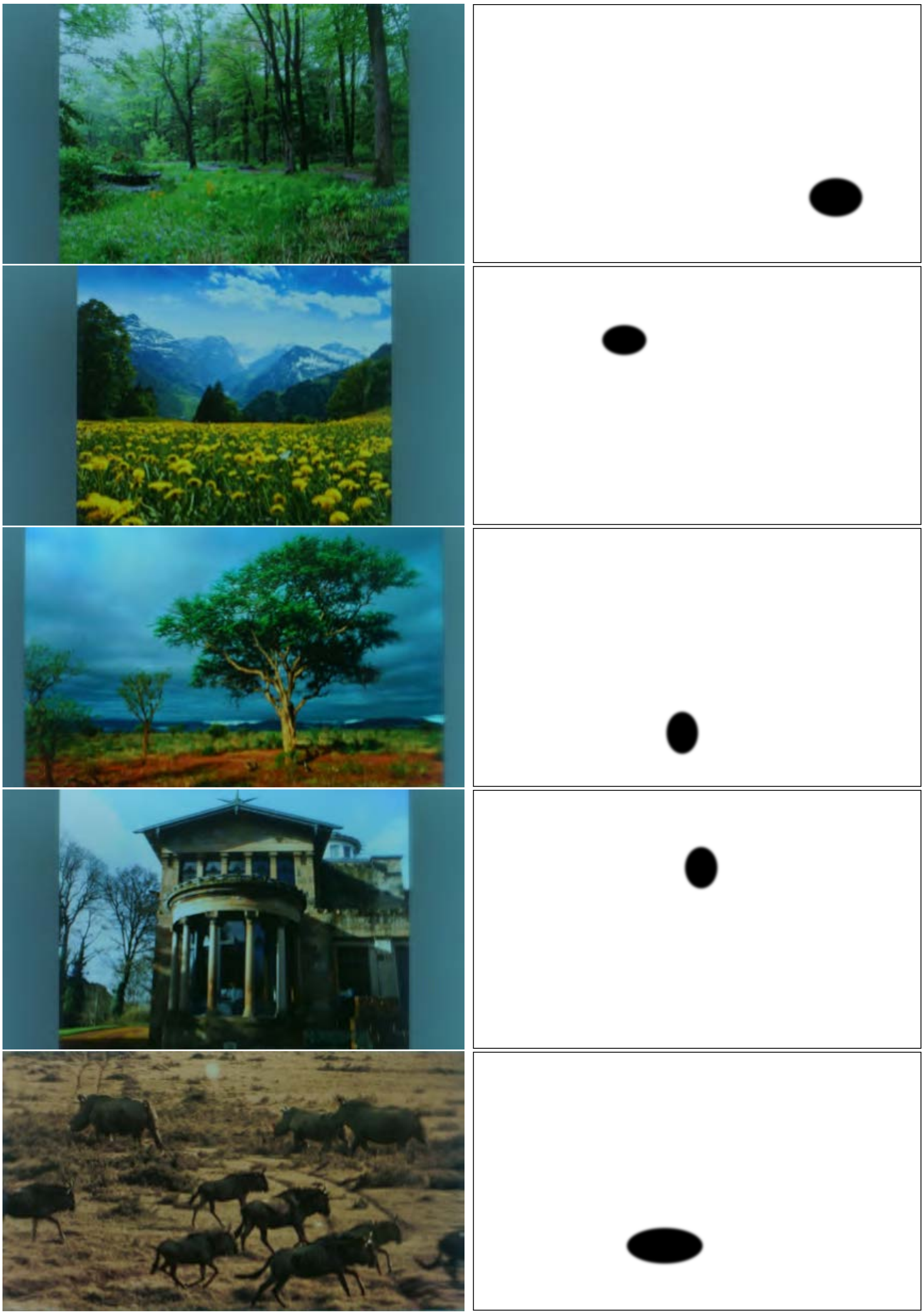


Figure C.1: Images from the CAT2000 used in our studies (Left) and the corresponding masks (Right).

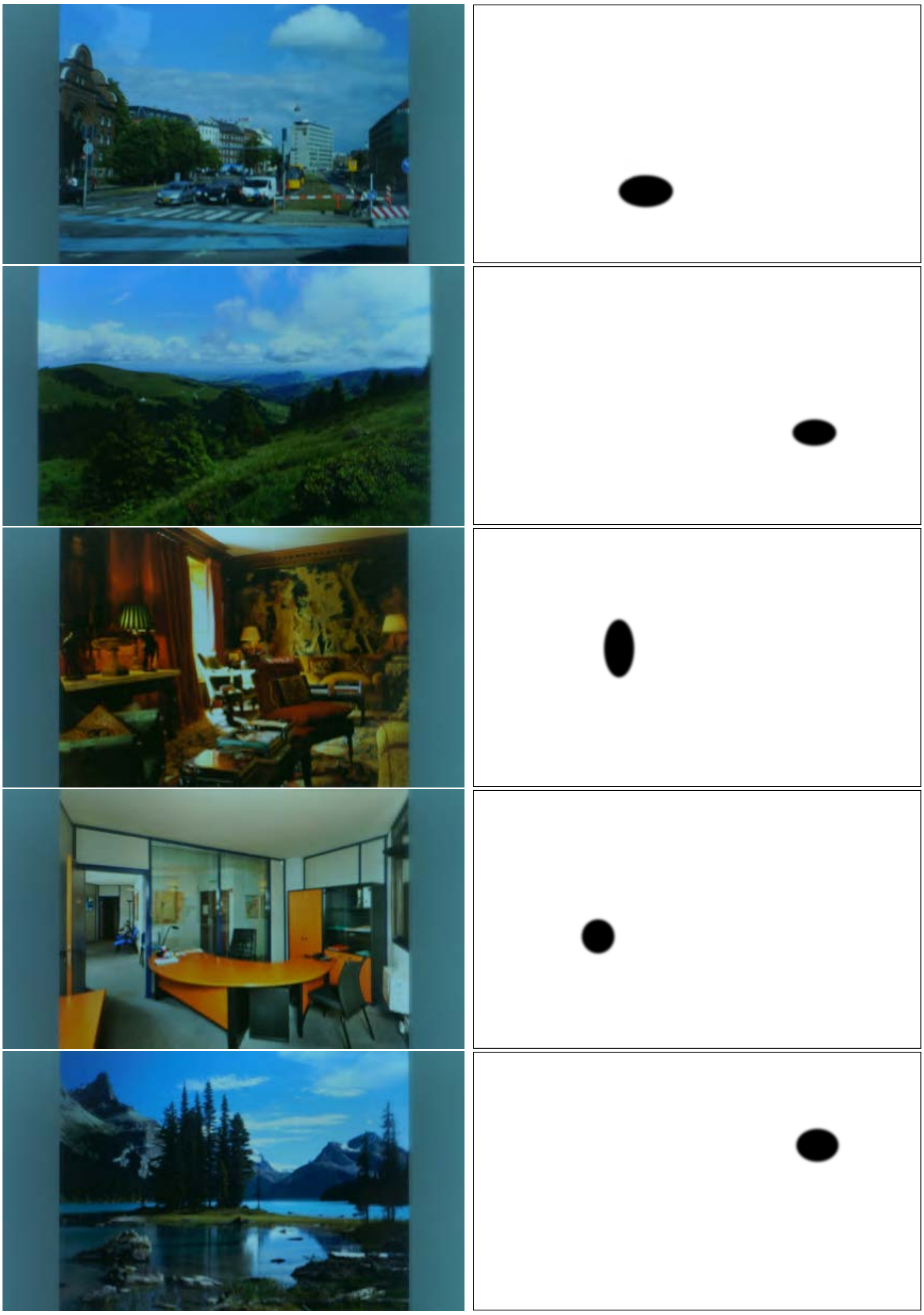


Figure C.2: Images from the CAT2000 used in our studies (Left) and the corresponding masks (Right).

C.2 Study 2

Due to the size of the data set we do not include a full set of images here, only a sample from each level.

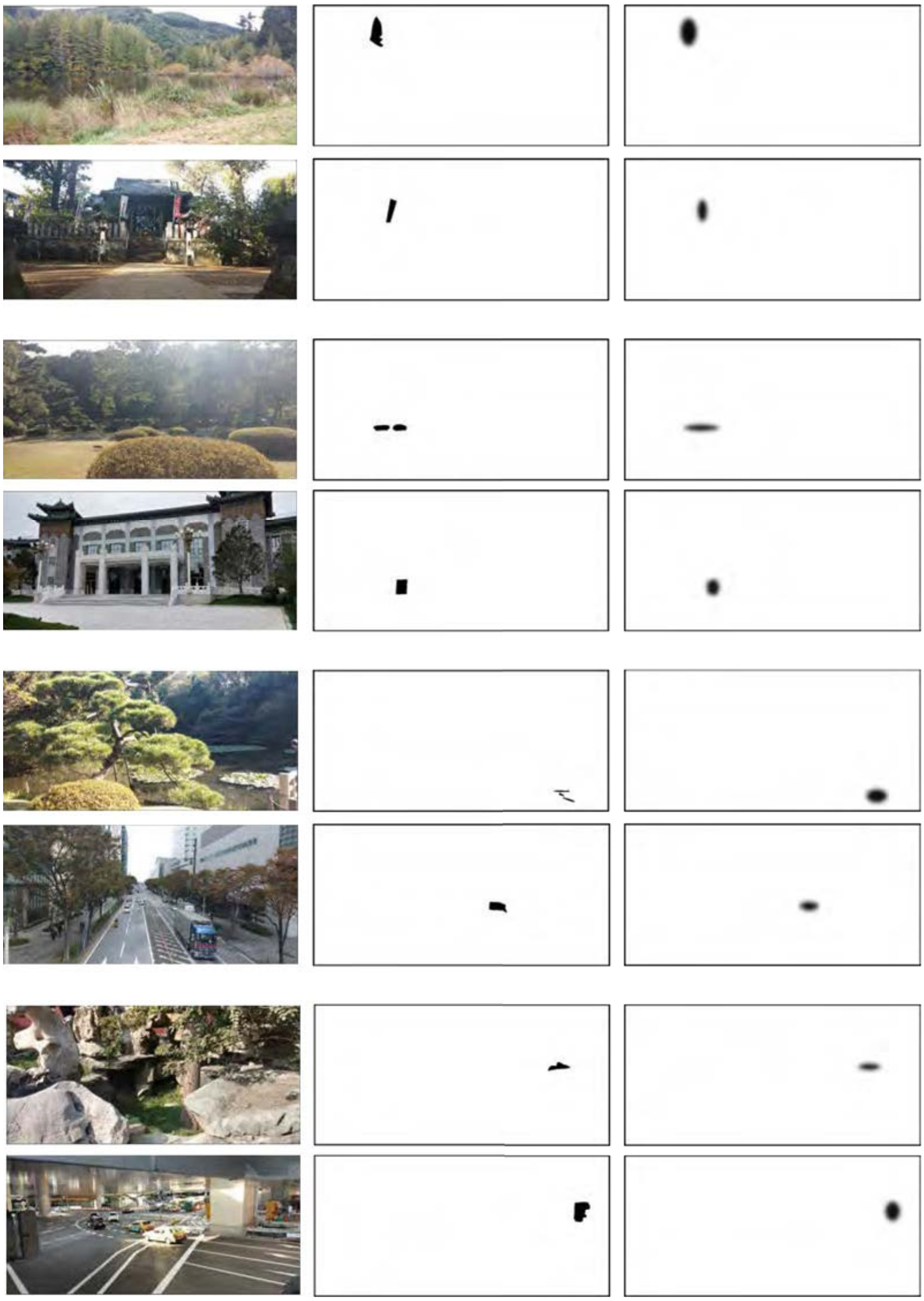


Figure C.3: Sample of images used for our exploration of our design space separated by level tested. Left: Original image. Centre: Mask used for saliency and temporal modulations. Right: Mask used for geometric outlines.

Appendix D

Visual Guidance Study: Results

The following tables summaries the statistical p-values from our test run on the data gathered during the exploration of our design space 5.6.

D.1 25%

	None	Flicker	Saliency
Flicker	0.81258	-	-
Saliency	0.35257	0.81258	-
Geometrical	7.2e-05	0.00018	0.00018

Table D.1: Fixation percentages at 25%. Freidman p-value = 2.71e-09

	None	Flicker	Saliency
Flicker	0.69	-	-
Saliency	0.40	0.94	-
Geometrical	7.5e-08	3.0e-07	4.5e-08

Table D.2: Average fixation time at 25%. Freidman p-value = 5.306e-10

	None	Flicker	Saliency
Flicker	0.89	-	-
Saliency	0.12	0.89	-
Geometrical	4.5e-08	4.5e-08	6.0e-08

Table D.3: Average fixation duration at 25%. Freidman p-value = 1.817e-11

	None	Flicker	Saliency
Flicker	1.0000	-	-
Saliency	1.0000	1.0000	-
Geometrical	0.0176	0.0048	0.0048

Table D.4: Average area explore at 25%. Freidman p-value = 0.005262

	None	Flicker	Saliency
Flicker	1.00000	-	-
Saliency	1.00000	0.48443	-
Geometrical	0.00196	0.00053	0.00067

Table D.5: Percentage of times modulation was noted at 25%. Freidman p-value = 8.039e-07

	None	Flicker	Saliency
Flicker	1.0000	-	-
Saliency	1.0000	1.0000	-
Geometrical	0.0023	0.0203	0.0048

Table D.6: Likert-like responses at 25%. Freidman p-value = 0.009654

D.2 50%

	None	Flicker	Saliency
Flicker	2.5e-05	-	-
Saliency	7.2e-05	3.0e-05	-
Geometrical	2.2e-05	0.083	3.0e-05

Table D.7: Fixation percentages at 50%. Freidman p-value = 3.326e-15

	None	Flicker	Saliency
Flicker	4.5e-08	-	-
Saliency	2.6e-05	4.5e-08	-
Geometrical	4.5e-08	0.0022	5.6e-07

Table D.8: Average fixation time at 50%. Freidman p-value = 2.42e-15

	None	Flicker	Saliency
Flicker	4.5e-08	-	-
Saliency	2.6e-05	4.5e-08	-
Geometrical	4.5e-08	0.00079	4.5e-08

Table D.9: Average fixation duration at 50%. Freidman p-value = 2.345e-16

	None	Flicker	Saliency
Flicker	6.3e-07	-	-
Saliency	0.0232	4.1e-06	-
Geometrical	4.4e-05	7.1e-05	0.0013

Table D.10: Average area explore at 50%. Freidman p-value = 8.873e-10

	None	Flicker	Saliency
Flicker	0.00052	-	-
Saliency	0.00549	0.00163	-
Geometrical	0.00101	0.00444	0.00267

Table D.11: Percentage of times modulation was noted at 50%. Freidman p-value = 1.606e-09

	None	Flicker	Saliency
Flicker	0.00044	-	-
Saliency	0.00760	0.00024	-
Geometrical	0.00065	0.00024	0.00636

Table D.12: Likert-like responses at 50%. Freidman p-value = 3.536e-06

D.3 75%

	None	Flicker	Saliency
Flicker	2.1e-05	-	-
Saliency	4.6e-05	0.00046	-
Geometrical	3.9e-05	0.59098	0.00055

Table D.13: Fixation percentages at 75%. Freidman p-value = 9.83e-13

	None	Flicker	Saliency
Flicker	4.5e-08	-	-
Saliency	2.1e-06	3.1e-06	-
Geometrical	4.5e-08	0.074	6.7e-06

Table D.14: Average fixation time at 75%. Freidman p-value = 2.978e-12

	None	Flicker	Saliency
Flicker	4.5e-08	-	-
Saliency	3.0e-07	3.1e-07	-
Geometrical	4.5e-08	0.00042	5.5e-05

Table D.15: Average fixation duration at 75%. Freidman p-value = 3.706e-14

	None	Flicker	Saliency
Flicker	1.3e-07	-	-
Saliency	3.5e-05	5.0e-06	-
Geometrical	2.0e-05	3.3e-06	0.019

Table D.16: Average area explore at 75%. Freidman p-value = 7.989e-10

	None	Flicker	Saliency
Flicker	0.00036	-	-
Saliency	0.00138	0.00617	-
Geometrical	0.00036	0.07260	0.00866

Table D.17: Percentage of times modulation was noted at 75%. Freidman p-value = 2.608e-10

	None	Flicker	Saliency
Flicker	0.00049	-	-
Saliency	0.00355	0.00075	-
Geometrical	0.00034	0.00355	0.00355

Table D.18: Likert-like responses at 75%. Freidman p-value = 3.409e-07

D.4 100%

	None	Flicker	Saliency
Flicker	2.3e-05	-	-
Saliency	5.0e-05	0.00049	-
Geometrical	2.3e-05	0.07625	0.00499

Table D.19: Fixation percentages at 100%. Freidman p-value = 8.175e-13

	None	Flicker	Saliency
Flicker	7.5e-08	-	-
Saliency	7.5e-08	5.7e-06	-
Geometrical	4.5e-08	0.00170	0.00068

Table D.20: Average fixation time at 100%. Freidman p-value = 4.661e-13

	None	Flicker	Saliency
Flicker	6.0e-08	-	-
Saliency	4.5e-08	2.2e-07	-
Geometrical	4.5e-08	4.1e-05	4.8e-05

Table D.21: Average fixation duration at 100%. Freidman p-value = 3.001e-14

	None	Flicker	Saliency
Flicker	6.3e-07	-	-
Saliency	0.0232	4.1e-06	-
Geometrical	4.4e-05	7.1e-05	0.0013

Table D.22: Average area explore at 100%. Freidman p-value = 8.873e-10

	None	Flicker	Saliency
Flicker	0.00062	-	-
Saliency	0.00055	0.25717	-
Geometrical	0.00062	0.25717	0.32096

Table D.23: Percentage of times modulation was noted at 100%. Freidman p-value = 1.362e-08

	None	Flicker	Saliency
Flicker	0.00056	-	-
Saliency	0.00056	0.03435	-
Geometrical	0.00034	0.12407	0.03435

Table D.24: Likert-like responses at 100%. Freidman p-value = 1.357e-07