

# Appendix

## Chroma Glasses: Computational Glasses for Compensating Colour Blindness

### INTRODUCTION

In this appendix, we will provide a few more details on implementation, parameters, sample images and data from the studies where there was no space in the paper. The main purpose is to support the replication of the implemented system and the performed studies.

### IMPLEMENTATION DETAILS

In the Software section in the paper we describe that our implementation of correcting CVD consists of two main tasks 1) identifying colours critical for CVD and 2) modulate the environment to compensate for CVD. The mathematical details for these computations are given in the following.

#### Adjustment Matrices for Calculating CVD in LMS

The first task is to identify the colours in the physical environment that are critical for CVD. For this purpose, we compute an image  $I_{CVD}$  that simulates CVD based on the input image  $I$ . The exact computation thereby depends on the type of colour blindness we are working with. As described in the paper this is based on the work of Viénot et al. [33]. For Protanopia we use the following matrix:

$$M_{adjust} = \begin{bmatrix} 0 & 2.02344 & -2.5281 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

For Deuteranopia, we use:

$$M_{adjust} = \begin{bmatrix} 1 & 0 & 0 \\ 0.494207 & 0 & 1.24827 \\ 0 & 0 & 1 \end{bmatrix}$$

As explained in the paper, these matrices are setup to work in LMS colour space. So in order to apply them, we have to convert the input image  $I$  from RGB colour space to LMS colour space. The overall equation to simulate CVD is then given by:

$$I_{CVD} = M_{LMS2RGB} \times M_{adjust} \times M_{RGB2LMS} \times I, \quad (1)$$

where  $M_{LMS2RGB}$  is the matrix to convert from LMS colour space to RGB colour space and  $M_{RGB2LMS}$  is the matrix to

convert from RGB colour space to LMS colour space. The CVD simulation image  $I_{CVD}$  is then in RGB colour space to directly display it on the ChromaGlasses.

For simplification and since these computations will not change, we include the matrix that combines the previous equation into one matrix for either Protanopia:

$$I_{CVD} = \begin{bmatrix} 0.112091 & 0.885306 & -0.00191592 \\ 0.112654 & 0.88974 & 0.000153488 \\ 0.00453387 & 0.000137308 & 1.00027 \end{bmatrix} \times I$$

and Deuteranopia:

$$I_{CVD} = \begin{bmatrix} 0.291994 & 0.70541 & -0.000193098 \\ 0.293446 & 0.708944 & 0.000149885 \\ -0.020975 & 0.0256464 & 1.00027 \end{bmatrix} \times I$$

Based on the simulated image  $I_{CVD}$ , we compute an error image  $I_{Error}$  that represents the difference between the actual image as seen from the user perspective and the simulated image  $I_{CVD}$ :

$$I_{Error} = I - I_{CVD} \quad (2)$$

#### CVD correction

The second step is the modulation of the environment to compensate for CVD. For this purpose, we need to precisely modulate the part of the environment seen from the perspective of the user via the OSTHMD. As described in the paper, we investigate two different methods for the CVD correction: 1) RGBShift and 2) LMSShift.

##### RGBShift

For RGB Correction Values  $R_n, G_n, B_n$  were set either by pre-set values based on the work of Tanuwidjaja et al. [32] (but adjusted to meet the requirements of the OSTHMD):

$$\begin{bmatrix} R_n \\ G_n \\ B_n \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1.5 \end{bmatrix} \quad (3)$$

or set by the user.

We then compute the output image by:

$$O_{RGB} = I - I_{Error} \times I + I_{Error} \times \begin{bmatrix} R_n \\ G_n \\ B_n \end{bmatrix} \quad (4)$$

Where  $I_{Error}$  is the computed error provided from simulations (equation 2) and  $I$  is the input image in RGB colour space.

	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.0016**	-	-	-	-
RGBShiftAdj	0.0016**	0.0032**	-	-	-
LMSShift	0.0016**	1	0.0019**	-	-
LMSShiftAdj	0.0016**	0.0022**	1	0.0016**	-
Edges	0.0016**	0.286	0.0032**	0.283	0.0022**

**Table 1. Results Feasibility Study: P-Values Wilcoxon for Success Rate: Significant differences are highlighted in grey. Significance levels are highlighted as \*  $p < 0.05$  and as \*\*  $p < 0.01$ .**

	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.0021**	-	-	-	-
RGBShiftAdj	0.0021**	0.0096**	-	-	-
LMSShift	0.0021**	0.5913	0.0039**	-	-
LMSShiftAdj	0.0021**	0.0039**	0.0096**	0.0021**	-
Edges	0.0021**	0.0195*	0.0023**	0.0109*	0.0021**

**Table 2. Results Feasibility Study: P-Values Wilcoxon for Confidence: Significant difference are highlighted in grey. Significance levels are highlighted as \*  $p < 0.05$  and as \*\*  $p < 0.01$**

Before adding the new colour to the correction a proportional amount of the original colour was removed, to allow for more diverse shifts, whilst still ensuring the new colour was set relevant to the original colour. This allows for the formula of [32] to be created with adjusted values, as well as more diverse user values.

#### LMSShift

For LMS correction, we maintained the same formula as that used for the RGB correction, however replaced  $R_n, G_n, B_n$ , with a set rotation of the original colour in LMS space ( $R_X, R_Y, R_Z$ )

$$M_{rotation} = M_{LMS2RGB} \times R_X \times R_Y \times R_Z \times M_{RGB2LMS} \quad (5)$$

We then compute the output image by using:

$$O_{LMS} = I - I_{error} \times I + I_{error} \times M_{rotation} \times I \quad (6)$$

## STUDY DETAILS

In this section, we provide a few more details about the user studies in the paper. In particular, how the confidence score was computed that was used in study 1 and study 2. Furthermore, we will provide more details about the results of the studies, such as p values and descriptive statistics.

### Details for Computing Confidence Score

We computed a confidence score that is adjusted for correct and incorrect answers. For this purpose, we convert the confidence values to a positive confidence value if their answer was correct and to a negative confidence rate if the answer was incorrect (see supplementary material for equations). This confidence rate allows us to penalize wrong answers with a high confidence and reward correct answers with high confidence:

$$Conf = \frac{1}{n} \sum_{i=1}^n (P_i C_i),$$

where

$$P_i = \begin{cases} 1, & \text{if correct answer} \\ -1, & \text{otherwise} \end{cases}$$

$C_i$  = confidence per item

### Details for study 1: Feasibility study

As described in the paper, we performed a Friedman test that showed significant differences in the success rate ( $\chi^2(5) = 79.154, p - \text{value} < 0.001$ ). Based on this we performed a post-hoc analysis using Wilcoxon signed-rank test (Holm correction). A detailed list of p values is given in Table 1. The results show that all ChromaGlasses correction methods have a significantly higher success rate compared to the uncorrected condition (None). LMSShiftAdj performed best (Table 5). Furthermore, results showed significant differences between all conditions, except RGBShift-LMSShift, RGBShift-Edges, RGBShiftAdj-LMSShiftAdj and LMSShift-Edges.

Furthermore, we performed a Friedman test showing that there are significant differences between confidence scores for the different methods ( $\chi^2(5) = 75.332, p - \text{value} < 0.001$ ). The post-hoc analysis using Wilcoxon signed-rank test (Holm correction) showed that there are significant differences between all correction methods except RGBShift-LMSShift (Table 2) and that all correction methods performed significantly better than the reference method using no correction (None). Again, LMSShiftAdj performed best (Table 6).

In addition, we were interested to look for correlations between the adjustment parameters for the RGBShiftAdj and LMSShiftAdj. For this purpose, we performed a Pearson correlation test between each of the parameters for each plate over all participants (Results in Table 7 and Table 8). The correlation tests did not show any consistent patterns. For most of the parameters we only found a weak correlation. For some we found a moderate correlation, but there was no consistent pattern. For instance, for the green adjustment parameter for RGBShift there was a moderate positive correlation for one pair of plates (Plate1 vs Plate4, Table 7 Middle). For LMSShift we only found weak correlations and also here we

	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.0031**	-	-	-	-
RGBShiftAdj	0.003**	0.0451*	-	-	-
LMSShift	0.0022**	0.0279*	0.5297	-	-
LMSShiftAdj	0.0019**	0.0011**	0.0111*	0.0411*	-
Edges	0.0037**	0.2589	0.0274*	0.0037**	0.0019**

**Table 3. Results Study 2 P-Values Wilcoxon for SuccessRate:** Significant difference are highlighted in grey. Significance levels are highlighted as \*  $p < 0.05$  and as \*\*  $p < 0.01$

	None	RGBShift	RGBShiftAdj	LMSShift	LMSShiftAdj
RGBShift	0.0033**	-	-	-	-
RGBShiftAdj	0.003**	0.0073**	-	-	-
LMSShift	0.003**	0.0033**	0.2917	-	-
LMSShiftAdj	0.003	0.003**	0.003**	0.003**	-
Edges	0.0033**	0.9792	0.0043**	0.0033**	0.003**

**Table 4. Results Study 2 P-Values for Wilcoxon for Confidence:** Significant difference are highlighted in grey. Significance levels are highlighted as \*  $p < 0.05$  and as \*\*  $p < 0.01$

Method	mean	sd	median
None	0.16	0.19	0
RGBShift	0.78	0.11	0.75
RGBShiftAdj	0.96	0.09	1
LMSShift	0.78	0.08	0.75
LMSShiftAdj	0.97	0.08	1
Edges	0.7	0.13	0.75

**Table 5. Detailed results for success rate for study 1.**

Method	mean	sd	median
None	-2.68	1.66	-3.25
RGBShift	2.93	0.85	3
RGBShiftAdj	4.01	0.8	4
LMSShift	2.84	0.58	2.75
LMSShiftAdj	4.43	0.77	4.75
Edges	1.99	0.96	2

**Table 6. Detailed results for confidence rate for study 1.**

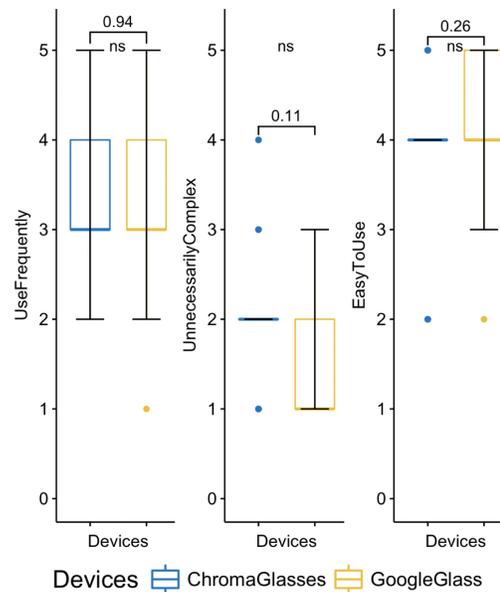
found no consistent patterns. For some of the plates there was no correlation or weak positive and sometimes a weak negative correlation (Table 8).

### Details for study 2: Feasibility wearing ChromaGlasses

With the replication study where participants wear the ChromaGlasses directly, we were able to reproduce similar results as in the feasibility study.

We performed again a Friedman test for the success rate that showed significant differences in the success rate ( $\chi^2(5) = 73.446$ ,  $p$ -value  $< 0.001$ ). The post-hoc analysis using Wilcoxon signed rank test (Holm correction) showed that there are significant differences in the success rate between the uncorrected option (None) and all correction methods (Table 3). Furthermore, we found significant differences between all correction methods, except RGBShift-Edges and LMSShift-RGBShiftAdj. As in study 1, LMSShiftAdj performed best (Table 9).

We also performed a Friedman test for the confidence rate that showed significant differences in the confidence rate ( $\chi^2(5) = 78.926$ ,  $p$ -value  $< 0.001$ ). The post-hoc analysis us-



**Figure 1. Results for the SUS questions for study 3.**

ing Wilcoxon signed rank test (Holm correction) showed that there are significant differences in the confidence rate between the uncorrected option (None) and all correction methods (Table 4). All correction methods showed significant differences to all other correction methods, except for RGBShift-Edges and LMSShift-RGBShiftAdj. Again LMSShiftAdj showed the best confidence score (Table 10).

### Details for study 3: Usability and workload study

In the following, we report on more details for the study comparing our approach (ChromaGlasses) against a Google Glass based approach (Tanuwidjaja et al.). For this purpose, we prepared a set of six different task with two test images each as described in the paper. The tasks were: 1) Identifying red fruits on a fruit-stand, 2) identifying red areas in a landscape, 3) identify the red graph within a complex graph, 4) identify

	P11	P12	P13	P14		P11	P12	P13	P14		P11	P12	P13	P14
P11	1				P11	1				P11	1			
P12	0.366	1			P12	0.201	1			P12	-0.474	1		
P13	-0.258	0.273	1		P13	0.184	0.334	1		P13	0.332	-0.099	1	
P14	0.061	-0.121	-0.056	1	P14	0.544	0.499	0.381	1	P14	0.191	0.054	0.458	1

Table 7. Correlations for RGBAdj parameters. Left) Red adjustment, Middle) Green adjustment, Right) Blue adjustment.

	P11	P12	P13	P14		P11	P12	P13	P14		P11	P12	P13	P14
P11	1				P11	1				P11	1			
P12	0.0270	1			P12	-0.219	1			P12	0.046	1		
P13	0.261	0.167	1		P13	0.402	-0.036	1		P13	1	0.0459	1	
P14	0.156	0.135	-0.159	1	P14	0.185	0.075	0.082	1	P14	0.292	-0.370	0.292	1

Table 8. Correlations for LMSShiftAdj parameters. Left) L adjustment, Middle) M adjustment, Right) S adjustment.

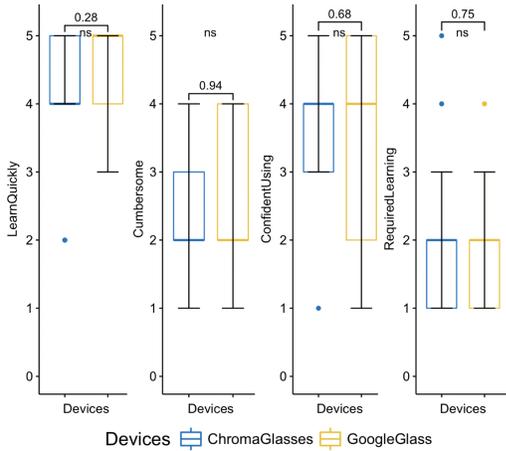


Figure 2. Results for the SUS questions for study 3 continued.

Task	Device	mean	sd	median
Identifying Fruits	ChromaGlasses	34.66	13.84	34.17
	Google Glass	39.61	19.3	30.83
Landscape	ChromaGlasses	32.7	12.57	36.67
	Google Glass	34.27	21.08	25
Identifying Graphs	ChromaGlasses	21.76	8.79	21.67
	Google Glass	34.26	15.89	33.33
Identifying Maps	ChromaGlasses	33.87	13.49	35.83
	Google Glass	40.2	19.42	40.83
Counting Flower	ChromaGlasses	28.19	17.36	23.33
	Google Glass	26.96	14.9	22.5
Counting Fruits	ChromaGlasses	28.19	14.34	27.5
	Google Glass	26.96	13.15	22.5

Table 11. Detailed results for the Raw TLX.

Method	mean	sd	median
None	0.15	0.21	0
RGBShift	0.69	0.14	0.75
RGBShiftAdj	0.83	0.15	0.75
LMSShift	0.86	0.18	1
LMSShiftAdj	1	0	1
Edges	0.64	0.15	0.75

Table 9. Detailed results for success rate for study 2.

Method	mean	sd	median
None	-2.93	1.62	-3.5
RGBShift	1.08	1.15	1.38
RGBShiftAdj	2.54	1.23	2.38
LMSShift	3.04	1.47	3.5
LMSShiftAdj	4.69	0.59	5
Edges	0.94	1.24	1.25

Table 10. Detailed results for confidence rate for study 2.

red areas in maps, approximate number of 5) red flowers or 6) red fruits on a tree. Figure 3 displays the used images.

As discussed in the paper, we found significant differences for the raw TLX overall score between the ChromaGlasses and Google Glass for the tasks on graphs and maps. As shown in Table 11 the score was lower for the ChromaGlass (Graphs: mean=21.76, Maps: mean=33.87) which indicates that the workload is smaller compared to the Google Glass (Graphs: mean=34.26, Maps: mean=40.2).

We also tested for differences in the questions on efficiency (subset of SUS), but we could not find any significant difference between both devices (Figure 1 and Figure 2). Further details about means, medians and standard deviation is available from Table 12.

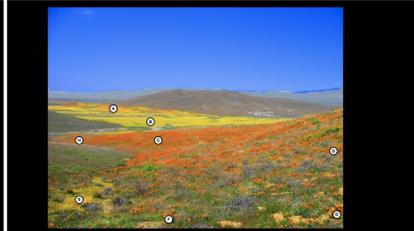
Question	Device	mean	sd	med.
Use Frequently	ChromaGlasses	3.41	0.8	3
	Google Glass	3.41	1.12	3
Unnecessarily Complex	ChromaGlasses	2	0.79	2
	Google Glass	1.59	0.8	1
Easy to use	ChromaGlasses	3.82	0.95	4
	Google Glass	4.18	0.81	4
Learn Quickly	ChromaGlasses	4	1.06	4
	Google Glass	4.41	0.71	5
Cumbersome	ChromaGlasses	2.29	0.92	2
	Google Glass	2.35	1.17	2
Confident Using	ChromaGlasses	3.71	0.99	4
	Google Glass	3.47	1.33	4
Learning Required	ChromaGlasses	1.94	1.14	2
	Google Glass	1.76	0.9	2

Table 12. Detailed results for the SUS questions.

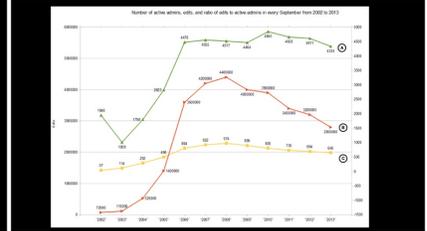
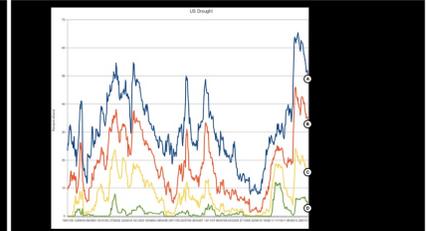
### Identifying Fruits



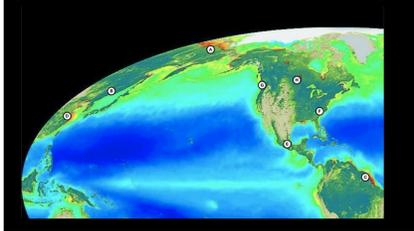
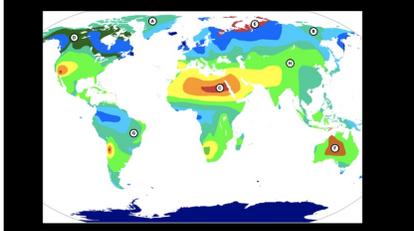
### Identifying Landscape



### Identifying Graphs



### Identifying Maps



### Counting Flowers



### Counting Fruits



Figure 3. Images used for study 3 investigating usability and workload study. For each task there are two images and we randomized.