

ChromaCorrect: Prescription Correction in Virtual Reality Headsets through Perceptual Guidance

Ahmet Güzel*
University of Leeds

Jeanne Beyazian†
University College London

Praneeth Chakravarthula‡
UNC Chapel Hill
Princeton University

Kaan Aksit§
University College London

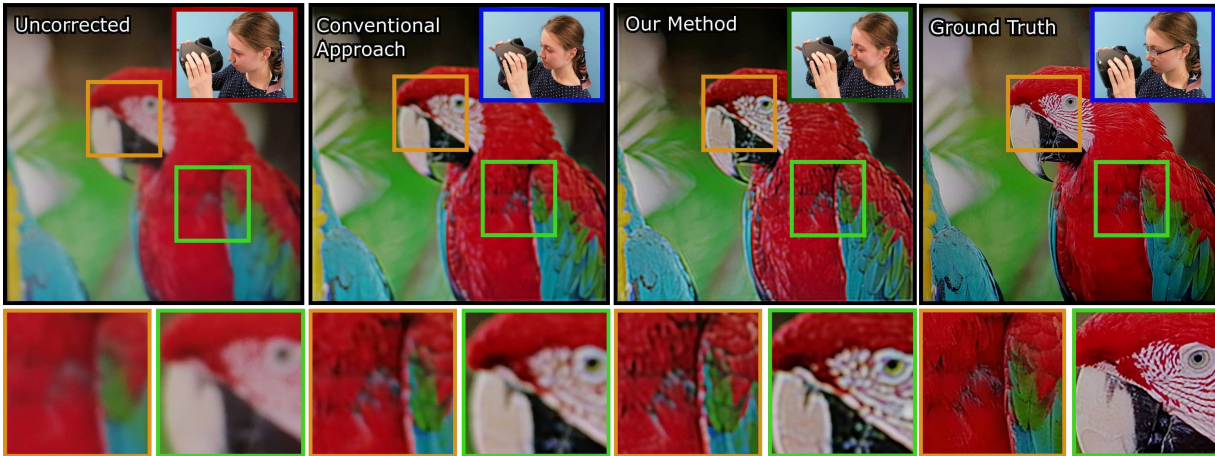


Figure 1: Perceptual Guidance in Prescription Correction. We provide a differentiable perception model for optimizing images that compensate for user prescription and improve the perceived contrast and color in images. Here, we show images as captured by a camera on a prototype Virtual Reality headset where the images are intentionally defocused to mimic an eye with common refractive errors. Without any prescription correction, the perceived images appear blurry due to defocus caused by refractive errors (first column). The second column captures the performance of the conventional algorithmic approach to prescription correction [31] for the same refractive error. Our proposed computational approach to algorithmic prescription compensation improves the perceived images, both in color and contrast, as can be seen in the third column. For reference, we provide a ground truth photograph focused at the display plane as in the fourth column, resembling what a user would see with their prescription lenses incorporated into the virtual reality headset. Source image is from Rich Franzen [12].

ABSTRACT

A large portion of today’s world population suffer from vision impairments and wear prescription eyeglasses. However, eyeglasses causes additional bulk and discomfort when used with augmented and virtual reality headsets, thereby negatively impacting the viewer’s visual experience. In this work, we remedy the usage of prescription eyeglasses in Virtual Reality (VR) headsets by shifting the optical complexity completely into software and propose a prescription-aware rendering approach for providing sharper and immersive VR imagery. To this end, we develop a differentiable display and visual perception model encapsulating display-specific parameters, color and visual acuity of human visual system and the user-specific refractive errors. Using this differentiable visual perception model, we optimize the rendered imagery in the display using stochastic gradient-descent solvers. This way, we provide prescription glasses-free sharper images for a person with vision impairments. We evaluate our approach on various displays, including desktops and VR headsets, and show significant quality and contrast improvements for users with vision impairments.

*e-mail: od20ahg@leeds.ac.uk

†e-mail: jeanne.beyazian.21@ucl.ac.uk

‡e-mail: praneethc@princeton.edu

§e-mail: k.aksit@ucl.ac.uk

Index Terms: Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

Virtual Reality (VR) headsets are becoming increasingly popular amongst consumers. This encouraged researchers to conceptualize and build technologies that would enable fully immersive remote experiences [34]. However, most of the recent developments overlook the prevalence of refractive vision problems such as myopia, hyperopia, or astigmatism among potential VR users especially older than 40 years old, which is at least 23.9%, 8.4% and 33% of population, respectively [33]. Moreover, while the current near-eye display research is focused on miniaturization of the headset to eyeglasses form-factor [21, 29], wearing prescription glasses under a VR headset causes uncomfortable viewing experiences that break the feeling of immersion.

Hardware-driven approaches to prescription correction [9, 23, 48] may lead to VR headsets that are bulkier and expensive while necessitating upgrading components with new devices. On the other hand, algorithmic approaches to prescription correction enable tackling the prescription issue without the need for specialized components and with the benefit of software updates [31].

Our work offers a new perceptually-guided algorithmic approach to prescription correction, while eliminating the need of corrective lenses. To this end, we first study the low-level workings of the Human Visual System (HVS), i.e., how different types of cone cells respond to various wavelengths of light. We then model the display’s

specific light spectrum (e.g. subpixels emitting various wavelengths) and the associated response of cone cells on the retina. Hence, we build an end-to-end differentiable perception model that helps us to simulate how a user with a Point-Spread Function (PSF) model with Zernike polynomials [27] perceives images on a specific display. Finally, our end-to-end perception framework enables optimizing the display rendering to produce an in-focus image for a user with vision impairments. Specifically, our work makes the following contributions:

- **Perceptually guided Prescription Correction.** We incorporate the display specific color perception and PSF of a user into a new differentiable model to ensure that the optimized image’s contrast and color characteristics are distinctly enhanced in visual perception.
- **Learned Prescription Correction.** We train a Convolutional Neural Network (CNN) to estimate optimal images for prescription correction, enabling prescription correction at interactive rates.
- **Evaluation on Actual Displays.** We analyze our findings beyond simulations. Thus, we evaluate our approach to VR headsets and conventional displays and demonstrate real-life use cases.

2 RELATED WORK

Researchers have previously attempted to compensate for refractive vision problems for glasses-free experience in displays. We summarize most relevant papers here in Tbl. 1.

Programmable Prescription Lenses. Utilizing focus-tunable lenses that may be adjusted to the user’s prescription is a common technique, especially in displays such as VR headsets where the users view a display through magnifying lenses [9, 26, 36, 41]. An alternative to these approaches, phase-only spatial light modulators, could also be used to form a programmable prescription correction lens [19]. Beyond requiring customized hardware, these techniques would also require eye-tracking and depth sensor data of a scene to operate, leading to more demands in hardware.

Computational Displays. Altering the display hardware and image acquisition technologies could help with prescription correction [25]. Huang et al. [16] address extreme contrast loss and ringing artifacts in algorithmic correction techniques by utilizing a stack of semi-transparent, light-emitting layers for LCDs. Wu and Kim [49] embed free-form image combiners inside prescription lenses to create customizable Augmented Reality (AR) displays. Pamplona et al. [37] implements 4D light field displays to move the solution to a higher-dimensional (light field) space, where the inverse problem is well-posed. To overcome this limitation in resolutions in Pamplona’s work [37], Huang et al. [17] propose a 4D prefiltering algorithm that can provide higher contrasts and resolutions. The described approach [37] has a significant drawback, namely that the PSF of an eye with refractive errors is typically a low-pass filter and, as such, irrevocably cancels higher frequencies from the original image. Moreover, holographic vision correction [10, 20] is superior to conventional approaches, including light field displays. Curious readers could consult to survey by Aydınoglu et al. [7] for more on these holographic displays.

Algorithmic Prescription Correction. Refractive vision impairments of the eye are commonly modeled by constructing a PSF that represents how the eye as an optical system transmits a point on the object to a point on the retina. The spatially varying PSF is convolved with the image of the object to produce the image formed on the retina. Performing the inverse operation, i.e., deconvolving the image with the retinal PSF, could help produce an image that forms clearly on the retina when observed. Alonso et

Table 1: Comparison of prescription correction techniques. Many of the solutions for prescription correction either fail to provide good image quality or require bulky hardware components affecting user comfort negatively. We take an algorithmic approach utilizing an accurate perception model of the human visual system, leading to improved image quality and real-time image generation. SW refers to Software while HW refers to Hardware in this table.

Name	Method	Perceptual Guidance	Realtime	Image Quality	Display Type
Multi-domain [4]	SW	Preliminary	No	Poor	Desktop
Constrained Total Variation* [31]	SW	Preliminary	No	Poor	Desktop
Tone Mapping [53]	SW	Preliminary	No	Poor	Desktop
Network [43]	SW	No	No	Poor	Desktop
Vision Enhancement [18]	SW	No	No	Poor	AR
SharpView [35]	SW	No	No	Poor	AR
FocusAR [9]	HW	No	Yes	Good	AR
Autofocals [36]	HW	No	Yes	Good	AR
Phase Modulated [19]	HW	No	Yes	Good	AR
RectifEye [26]	HW	No	Yes	Good	VR
Alvarez Lenses [41]	HW	No	Yes	Good	VR
Software [51]	SW	No	Yes	Poor	VR
Ours	SW	Yes	Yes	Fair	VR

*This technique is referred to as the conventional method throughout the paper.

al. [3] verifies the possibility of such an image correction technique by constructing a simple artificial eye and comparing the image it forms when viewing a standard and a corrected image. They also propose an ad-hoc solution to mitigate contrast loss and “ripples” or ringing artifacts [4]. Monalto et al. [31] present constrained total variation to decrease ringing artifacts in the corrected image while sharpening the image’s edges, thereby producing an image with high contrast along sharp edges. Ye et al. [53] focus on finding a ringing-free image with higher contrast in locations important to Human Visual System (HVS), while tolerating more blurriness elsewhere. Tanaka et al. [43] uses a CNN-based pipeline for prescription correction along with Zernike-based visual aberration modeling. Li et al. [28] feed an aberrated image and a map of a PSF for multiple subregions, to account for spatially variant aberrations into a deep neural network and train it for image correction on a variety of lenses. Similar image correction techniques have been applied to VR headsets. Itoh et al. [18] corrects the defocus aberration for optical see-through headsets by overlaying a compensated image in the user’s view. Xu et al. [51] use gradient-based priors to achieve realtime visual aberration correction for VR HMDs. Oshima et al. [35] describe realtime defocus correction for optical see-through HMDs, which is caused by focal rivalry: the simultaneous viewing of real and virtual content.

Perceptual considerations in displays and graphics systems are becoming commonplace in relevant research branches (e.g. consult our supplementary for perceptual considerations in graphics systems). The surveyed research work does not provide a complete model of HVS in their solutions, leading to either poor image quality

or demanding hardware. We believe our work resembles the first attempt to enhance algorithmic solutions in the literature by bridging the gap between perceptual modeling and prescription correction.

3 PERCEPTUALLY GUIDED PRESCRIPTION CORRECTION

We introduce a differentiable framework for modeling the display and human visual perception, encapsulating display-specific parameters, color and visual acuity of human visual system and the user-specific refractive errors. Our framework allows for optimizing prescription compensated rendered imagery on standard displays using a gradient-based policy with novel display-specific perceptually guided loss functions (Section 3.1). We rely on Zernike polynomials (Section 3.2) for describing user-specific retinal point spread functions [10] within the forward model to represent optical aberrations in the HVS (Section 3.3). On overview of our entire display-visual perception model and the optimization process is depicted in Fig. 2.

3.1 Modeling Display-specific Visual Perception

We characterize our target display and device a computational model to transform the displayed imagery on the target display into imagery as perceived by the HVS.

Characterizing target display. A given display has three types of emission spectra, $\lambda_R, \lambda_G, \lambda_B$, for their red, green, and blue channel pixels, respectively. As these emission spectra vary for each display system, we calibrate the spectra using a spectrometer by measuring the spectral bands of the target display at various pixel levels. More details on the spectra measurement and display calibration process are discussed in the Supplementary Material. We then fit a proxy function to determine the display color primaries from the spectral measurements. While a simple Gaussian mixture model with weighted sum of Gaussians can be used for such a proxy color primary function, we learn this function using a multi-layer perceptron network that act as general function approximator. Implementation of this proxy function fitting can be found in (See `odak.learn.tools.multi_layer_perceptron()` in [2]). Once we fit a proxy function for the color primaries, we utilize it to investigate the color perception responses of the HVS.

Converting color primaries to perceived colors. Human retinal cells can be broadly classified into rods and cones. Cone cells, which are primarily responsible for color perception in the HVS, are of three different subtypes: Short (S), Medium (M), and Long (L) cells. Each of them differs in its sensitivity to different wavelengths of light. Please refer to our Supplementary Material for a detailed discussion this. The L, M, and S cones reduce wavelengths of incoming light into trichromat values by integrating them over their response functions [50]. Note that perception in HVS is contrary to modeling general camera or display response where red, green and blue wavelengths are independently measured on the camera sensor or the human retina. The following steps show how to convert an input color image displayed on a target display to the corresponding cone response:

$$\begin{bmatrix} I_L \\ I_M \\ I_S \end{bmatrix} = \begin{bmatrix} L_R & L_G & L_B \\ M_R & M_G & M_B \\ S_R & S_G & S_B \end{bmatrix} \begin{bmatrix} I_R \\ I_G \\ I_B \end{bmatrix}, \quad (1)$$

where I_R, I_G, I_B represents red, green and blue pixel values of an input image, and I_L, I_M, I_S represents L, M and S cone activation values for each pixel of the displayed image. From the generalized formula above, we provide a sample conversion for L_R as in the following equation,

$$\sum_{\lambda_R} \lambda_L \lambda_R = L_R, \quad (2)$$

where λ_L represents L cone sensitivity function, λ_R represents red pixel emission spectrum function for a targeted display, and L_R

represents L cone output for the displayed red pixel. Similarly, L cone sensitivity functions for green and blue pixel emissions can be computed. Thus, L, M and S cone sensitivity functions can be computed for the three different subpixel emissions. After computing the cone sensitivity functions, we apply the conversion from the color opponency model proposed by Schmidt et al. [40] to represent a complete perception model,

$$\begin{bmatrix} I_{(M+S)-L} \\ I_{(L+S)-M} \\ I_{(L+M)+S} \end{bmatrix} = \begin{bmatrix} (I_M + I_S) - I_L \\ (I_L + I_S) - I_M \\ (I_L, I_M, I_S) \end{bmatrix}, \quad (3)$$

where $I_{(M+S)-L}, I_{(L+S)-M}, I_{(L+M)+S}$ represents the three channels of the image sensed in the color-opponency space.

3.2 Computing Point Spread Functions from Color Primaries

The point spread function for the HVS with visual aberrations can be defined over several wavelengths of light (see Supplementary Material for equations). Therefore, we can sample a set of wavelengths from each color primary, calculate PSFs for each and use a weighted sum of the PSFs to obtain a single, combined PSF for each color primary,

$$PSF(x, y, c) = \sum_{\lambda_c} w_{\lambda_{c_i}} PSF(x, y, \lambda_{c_i}) \quad (4)$$

where c represents a particular color primary, $PSF(x, y, c)$ is the PSF for a particular color primary, $PSF(x, y, \lambda_{c_i})$ the PSF for a sampled wavelength in the color primary and $w_{\lambda_{c_i}}$ is the weight for that sampled wavelength. The above PSF kernel can be utilized in RGB, or color opponency spaces, depending on designers choices. In our method, we introduce color opponency based PSF formulation (perceptually guided) to improve the perceptual characteristics (contrast, quality) of the retinal image. Eq. 4 is extended to formulate LMS based kernel,

$$PSF_{lms}(x, y, \lambda_{c_i}) = A * PSF(x, y, \lambda_{c_i}) \quad (5)$$

$$PSF_{lms}(x, y, c) = \sum_{\lambda_c} w_{\lambda_{c_i}} PSF_{lms}(x, y, \lambda_{c_i}) \quad (6)$$

where A is the conversion matrix defined in Eq. 1, $PSF_{lms}(x, y, c)$ is the PSF for a particular color primary with LMS components. Similarly, we modelled the digital camera color primary decoding by using measurements from the display and captured images from the digital camera. In this way, we are able to use digital camera captured images to represent our work in this paper. In the Eq. 5 and Eq. 6, PSF_{lms} is represented for both the HVS and digital camera RGB decoding. We can now compute the retinal image $r(x, y, c)$ in the LMS space, by convolving PSF_{lms} with the input image $s(x, y, c)$,

$$r(x, y, c) = PSF_{lms}(x, y, c) * s(x, y, c). \quad (7)$$

3.3 Optimizing Images for Prescription Correction

In the final step, we aim to optimize an image which, after passing through the eye's optical system (modelled as a convolution in Eq. 7), is intended to produce a retinal image that is as close as possible to the ground truth image. This is done by solving the optimization problem,

$$s' \leftarrow \underset{s \neq \emptyset}{\operatorname{argmin}} \mathcal{L}(PSF * s, t) \quad (8)$$

where t is the the ground truth image and s' is the input image optimized for a user's eye, PSF is kernel defined in Eq. 4. In our method, we reformulate Eq. 8 to incorporate color opponency space optimization,

$$s' \leftarrow \underset{s \neq \emptyset}{\operatorname{argmin}} \mathcal{L}(PSF_{lms} * s, t_{lms}) \quad (9)$$

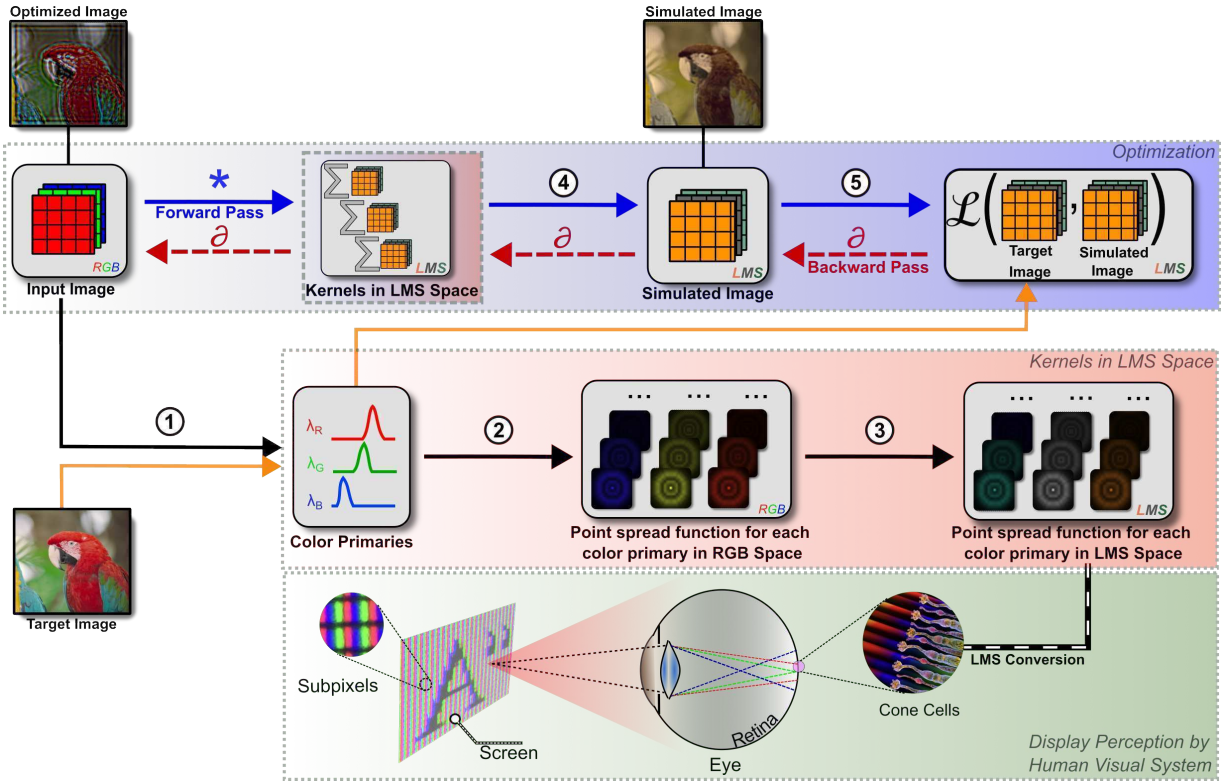


Figure 2: Prescription correction using a perceptually guided computational model and a differentiable optimization pipeline. (1) A screen with color primaries (RGB) displays an input image. (2) A viewer’s eye images the displayed image onto the retina with a unique Point Spread Function (PSF) describing the optical aberrations of that person’s eye. (3) Retinal cells convert the aberrated RGB image to a trichromat sensation, also known as Long-Medium-Short (LMS) cone perception [42]. (4) Our optimization pipeline relies on the perceptually guided model described in previous steps (1-3). Thus, the optimization pipeline converts a given RGB image to LMS space at each optimization step while accounting for the PSFs of a viewer modelled using Zernike polynomials. (5) Our loss function penalizes the simulated image derived from the perceptually guided model against a target image in LMS space. Finally, our differentiable optimization pipeline identifies proper input RGB images using a Stochastic Gradient Descent solver [38].

where t_{lms} is the the ground truth image in LMS space and s' is the input image optimized for a user’s eye, PSF_{lms} is kernel defined in Eq. 6. To perform the above optimization, we compare images using a loss function (e.g. least-squared error) to calculate the error between the ground truth image and the retinal image, $\mathcal{L}(r(x, y, c), t(x, y, c))$, where x and y represent image coordinates and c the color channels, which could be in RGB or LMS color opponency spaces. Note that we have also built a learned equivalent of our approach, which we will detail in the Sec. 4.

4 IMPLEMENTATION

Our approach is comprised of three primary elements: a color perception model, a prescription correction optimization pipeline and a learned model that demonstrates that our differentiable pipeline can be learnt. All of these components are implemented on PyTorch [38].

4.1 Color Perception Model

Firstly, we identify the emitted wavelengths from the subpixels of a target display device. For that purpose, we acquire the spectrometer data for a target display consisting of discrete wavelengths and their corresponding intensity values normalized between zero and one. We use Multilayer Perceptron (MLP) to fit a curve on this discrete data to achieve a vector representation of our intensity profile of color primaries with respect to wavelength. Our MLP has 64 hidden layers and converges over 1000 iterations in training with a learning rate of 0.0005. Once we have numerically identified the normalized

```

1 import torch
2
3 def get_LMS_kernel(spectrum):
4     [ $\lambda_R, \lambda_G, \lambda_B$ ] = get_display_spectrum()
5     for  $i, \lambda_i$  in enumerate( $[\lambda_R, \lambda_G, \lambda_B]$ ):
6         wavelengths=torch.arange(400, 701)
7         psf = generate_psf( $\lambda_i$ )
8         weighted_psf = spectrum[ $\lambda_i$ ] * psf
9         LMS_kernel += convert_to_lms(
10            weighted_psf,  $\lambda_i$ , spectrum[ $\lambda_i$ ])
11     return LMS_kernel
12
13 kernel = torch.zeros(3, H, W, 3)
14 kernel[0] = [H, W, get_LMS_kernel(red_spectrum)]
15 kernel[1] = [H, W, get_LMS_kernel(green_spectrum)]
16 kernel[2] = [H, W, get_LMS_kernel(blue_spectrum)]

```

Listing 1: Computing L, M, S kernel triple for each of the R, G, B channels and create forward model kernel in 4D tensor form. The abstraction is Pythonic.

intensity of each color primaries as a function of wavelength, we use these 2D (intensity, wavelength) vectors to create our color perception based kernel in LMS space. For each color primary, we create the set of PSF based on our zernike polynomial generator by

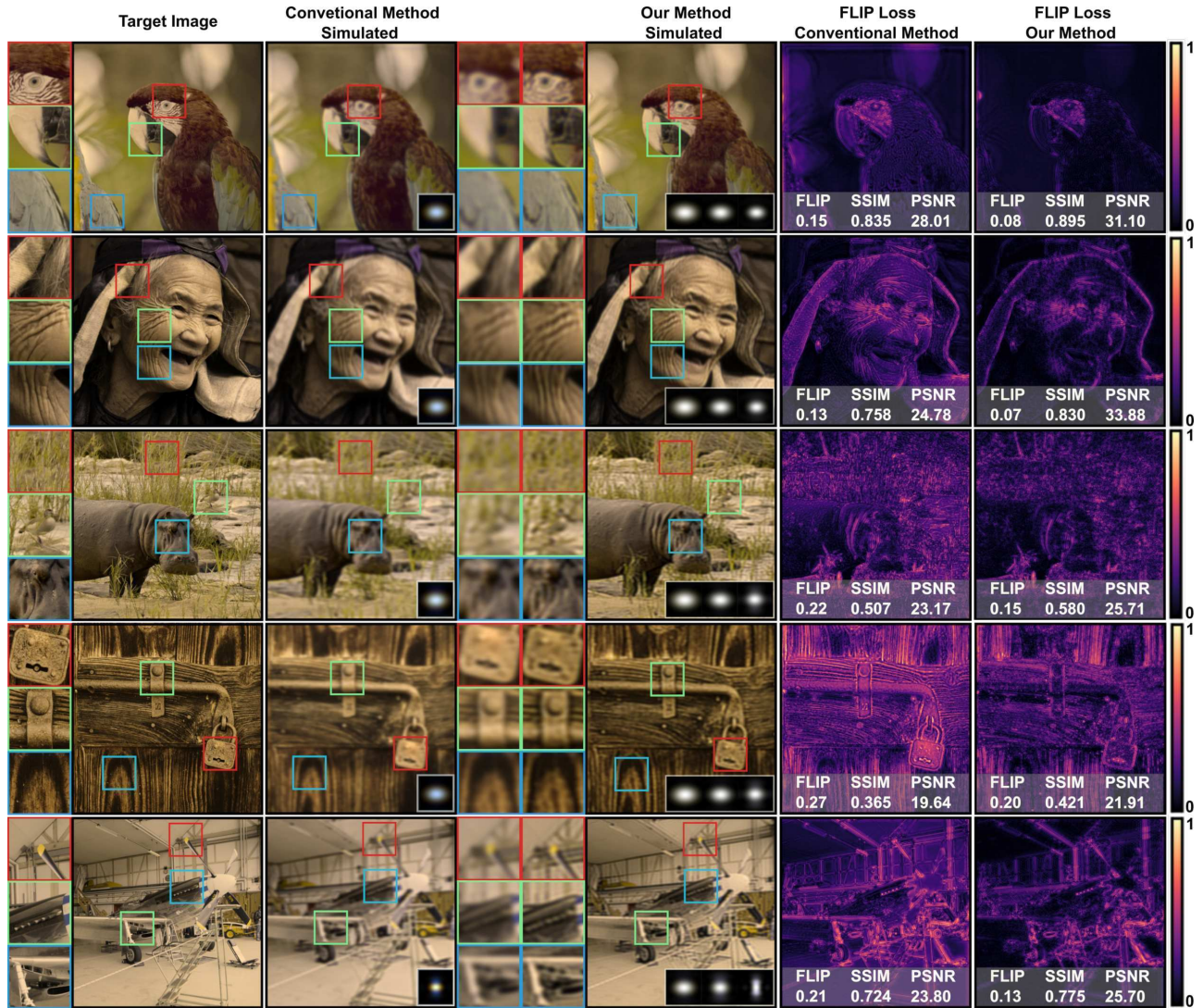


Figure 3: Here we compare outputs from five different refractive vision problems (myopia, hyperopia, hyperopic astigmatism, myopic astigmatism, and myopia with hyperopic astigmatism) for five sample input images. We provide simulated LMS space representations of target image, conventional method output, and our method. FLIP per-pixel difference along with its mean value (lower is better), SSIM and PSNR are provided to compare performance of methods. Our method shows better loss numbers for each image quality metrics for each experiment in simulated LMS space. The contrast improvement by using our method against conventional method also can be observed perceptually. Source images are from DIV2K image dataset [1].

sampling wavelengths from 400 to 700 with 1 nm intervals. During each sampling step, we create weighted kernels by multiplying the created PSF with the intensity value based on corresponding wavelength from our created 2D vectors for each color primary. After creating the weighted kernel in each sampling step, we obtain LMS cone responses of weighted kernel using the same intensity and wavelength data. To compute LMS cone responses, we use the method explained in section 3.1. In the last step, the set of weighted kernels are summed up to create our color perception based kernel for each color primary forming a 4D tensor as $[Color\ Primary, H, W, LMS\ Response]$. Our method differs from the conventional method both in terms of kernel type, and convolution operation.

In conventional method, kernel is a 3D tensor with RGB channels while in our method we use 4D tensor. In this 4D tensor formed kernel, each color primary has LMS triple separately as $[3, H, W, 3]$. The LMS based kernel convolves the image's each color channel with corresponding each display spectrum LMS responses. This

operation computationally more expensive compared to conventional method, since more matrix operation is needed. We provide a pseudo-code for constructing our LMS based kernel as in Listings 1.

4.2 Optimization Pipeline

We implement a prescription correction optimization pipeline using a modern machine learning library with automatic differentiation [38]. Source code of our implementation is publicly available at [GitHub : complight/learned-prescription](https://github.com/complight/learned-prescription) [13] and [GitHub : complight/learned-prescription_model](https://github.com/complight/learned-prescription_model) [8].

Optimization loop: The differentiable input RGB image initialized from the our target RGB image, and it is passed through the forward model during optimization loop. In forward model, each color channel of initialized input RGB image convolved with the LMS kernel created in computational color pipeline. For example, red channel of input RGB image is convolved with L, M, S channel of red spectrum kernel in LMS space. Other color channels of input

RGB image are convolved with the same method. The resulting simulated image represents the image formed on the retina from L, M, S cone activations. The target image is converted to LMS space to calculate L2 loss against the simulated image in LMS space, which is back-propagated through the optimization model to the input RGB image. Our results are obtained using Stochastic Gradient Descent with ADAM [24] as the optimizer. Our pipeline is available to be used in NVIDIA GPU accelerated computer.

4.3 Learned Model

We implement a semi-supervised deep learning model capable of reconstructing optimized images from their original RGB versions. We use a U-Net architecture [39] for this purpose. Such solution is more suitable than an iterative process for achieving real-time applications. But it trades the image quality for a faster rendering speed. Our model comprises of 2 outer layers linked to 8 convolutional hidden layers symmetrically connected by skip connections. Each layer on the contractive path of the model are formed by a double convolution and a max pooling operation. On the expanding path, an up-sampling operation with bilinear interpolation initiates each convolution. During training, batch normalization and ReLU activation are used.

Our model was evaluated on a machine with an NVIDIA GeForce RTX 2070 GPU. The training dataset comprises of 20 images of dimension 512×512 pixels, the RGB images were obtained from Zhang et al.'s color image processing dataset [54] and the target optimized images were generated using our iterative method. A learning rate of 1×10^{-4} was used for the training phase and a conventional mean-squared-error loss function guides the stochastic gradient descent optimization. With convolutional kernels of size 3×3 , each input image sees its channels expand from 3 to 92 and all the way up to 1472 at the latent space. The results in Figure 5 shows the comparison of the corrected image between our original pipeline and the neural network's prediction after over 800 epochs of training. The average time to generate a single corrected image is 0.0029 seconds with the model as opposed to 8.127 seconds using the original method, a tremendous decrease.

5 EVALUATION

We divided our experiments in to two sections. In the first part, we use real hardware to test our methods for defocus prescription. We used Oculus Quest 1 virtual reality headset, and we placed a defocus lens to create artificial prescription for a camera shot. In our experiments we use fixed pose, focus camera to capture images to demonstrate the method's performance.

Figure 4 shows our experimental setup for defocus experiments. Figure 1 shows results from the first part of our experiments. We modeled the myopia defocus, since in this way we can use defocus lenses to replicate eye prescription. Experiments shows that we improved contrast and color compared to conventional method. In fact, our method is not able to produce same quality with the target image.

In the second part, we evaluated our method with different prescriptions to model different refractive eye problems in simulated retinal image representation. Thus, all the images used in this part are evaluated in simulated LMS space. Selected images are aimed to have both high frequency and low frequency features. Four common different prescriptions are chosen which are myopia, hyperopia, myopic astigmatism, hyperopic astigmatism to test our method against the conventional model. Also, we tested our method for myopia with hyperopic astigmatism as a complicated refractive eye problem which is not trivial for eyeglass correction. In each refractive eye problem modelling, $\pm 1.5D$ refractive error is used to model prescriptions. Results are visualized in Figure 3. We use different image quality measures to compare our method against the conventional method. Our primary chosen image quality metric is FLIP which

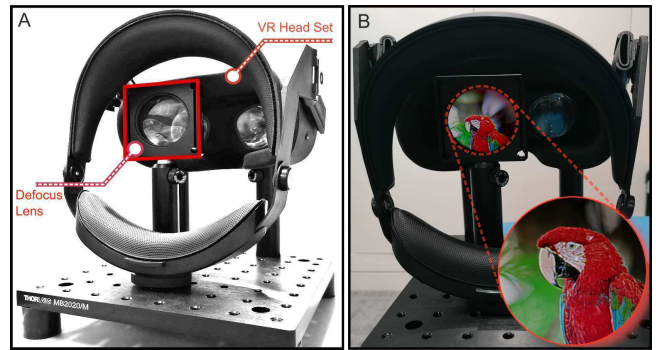


Figure 4: Testbed used in our evaluations. (A) We use a virtual reality headset and a camera to capture images from our virtual reality headset. To emulate a prescription problem in the visual system, we use a defocus lens. (B) We take pictures with fixed pose and camera focus from behind the defocus lens to evaluate reconstructed images.

compares the images by using principles of human perception [5]. FLIP allows per-pixel difference loss maps in magma color which is used in our evaluation images to visualize difference in each pixel against the ground truth image. In each pixel comparison FLIP counts both color and edge differences based on models of HVS. Therefore, we believe that this metric fits with our work. Although many research on this area has been used SSIM or PSNR loss, however FLIP is adventegous as it is adhering human visual system while others are not [32]. In addition to FLIP, we use SSIM and PSNR to compare our method agains to conventional method to be stayed relevant with the research community. Figure 3 demonstrates the comparison of our method against the naive method with our perceptually guided color modelling.

Results shows that color opponency based kernel modelling improves the contrast of retinal output image. Selected 3 areas are magnified to show visibility of improvements in a detailed way. Fom the per-pixel difference loss maps, we found that our method is better in low frequency features while our method provides slight improvement in high frequency parts of images. In overall, it is shown that perceptually guided color based kernel has better contrast compared to conventional method.

6 DISCUSSION

Our method could be potentially integrated with Mandl et al. [30] to support a broader user base with refractive vision impairments. To the best of our knowledge, we provide encouraging results improving the conventional method in the literature.

Spatially Varying PSF. Our method does not account for spatially varying natures of PSF in the HVS, which often arrives with computational cost and complexity [15]. We designed our implementation in constant resolution displays instead of varying resolution ones like foveated displays. As an alternative, the deep learning methods can help support spatially varying PSF convolutions in the modeling [52] with lesser computational cost but with demand in data for training. Thus, our method can benefit from these techniques in the future for precision modeling.

Chromatic Aberrations In A Human Eye. We use PSF created by the same Zernike coefficients for each wavelength in our forward model. However, optics of HVS contain chromatic aberrations that are wavelength-dependent. As a future work, we can further improve the accuracy of our modeling for a human observer by taking into account the chromatic aberrations in the HVS. In the meantime, curious readers can find greater details regarding chromatic aberrations in work by Cholewiak et al. [11].

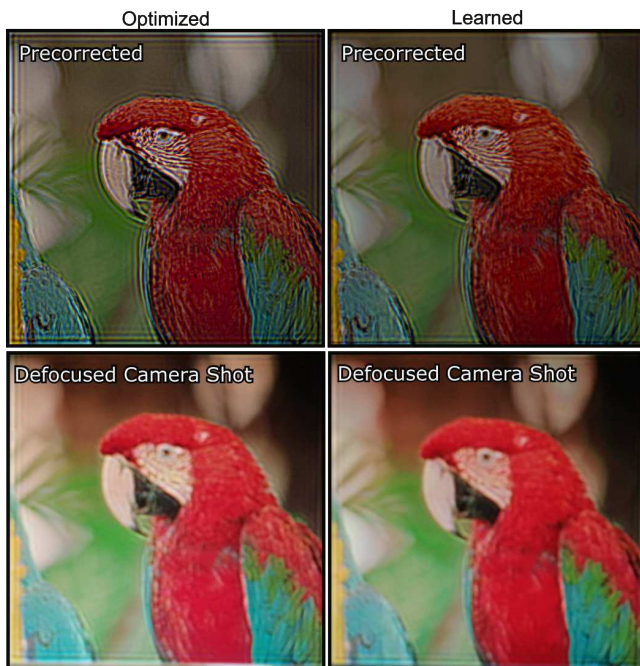


Figure 5: Results from our learned model. We compare our optimization pipeline against our learned model. The top row shows precorrected images reconstructed by our optimizer and learned model. The bottom row shows photographs for each case when captured with a defocused camera.

Image Quality. Approaches for prescription correction with additive displays are fundamentally limited. This limit stems from the fact that PSF, the non-negative transfer function of an additive display, could support a limited range of frequencies and cause contrast loss. Our work could be made to be complementary to holographic displays [7, 10, 46], which promise a unique solution for this issue originating from non-negativity in additive displays.

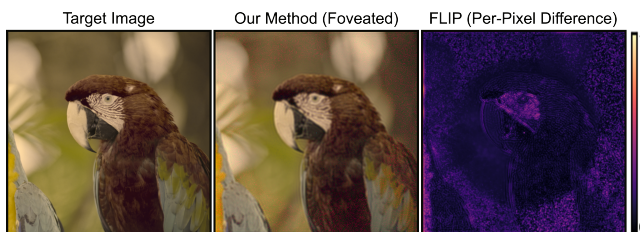


Figure 6: We reconstructed image in our method with addition of foveation. Foveated rendered area is in the center of reconstructed image. FLIP per-pixel difference map highlights the foveation.

Foveated Rendering. Foveated rendering in graphics [45] and displays [22] has garnered interest in the VR and AR research community. We believe that our method can also benefit from this trend by accounting for trends in chromatic and achromatic contrast sensitivity [14, 44, 47] in the HVS. Moreover, we could add a rod’s response to cone responses by reformulating the LMS response to improve color difference predictions [6]. We will explore this path in our future work (See Figure 6 for our early results.)

7 CONCLUSION

Identifying means to help display users with their vision impairments is an essential aspect of graphics systems. As we focus on this critical

issue, we present a new rendering approach that provides sharp images when viewed by users with vision impairments without their prescription glasses. Specifically, our rendering approach uniquely merged key insights from HVS. It showed that it could help improve visual experiences and comfort in VR headsets by enhancing color and contrast in the displayed images. The future will likely bring more principled approaches in AR/VR displays (e.g. holographic displays), which could enable future research investigations based on findings from this work.

ACKNOWLEDGMENTS

The authors would like to thank Praveen Selvaraj for the engineering support through early phases of this project. Kaan Akşit is supported by Meta Reality Labs inclusive rendering initiative for building the rendering pipeline.

REFERENCES

- [1] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.
- [2] K. Akşit, A. S. Karadeniz, P. Chakravarthula, W. Yujie, K. Kavaklı, Y. Itoh, D. R. Walton, and R. Zhu. Odak, May 2022. If you use this software, please cite it as below. doi: 10.5281/zenodo.6528486
- [3] M. Alonso, A. Barreto, J. A. Jacko, and M. Adjouadi. Verification of computer display pre-compensation for visual aberrations in an artificial eye. In *Assets '05: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pp. 210–211. Association for Computing Machinery, New York, NY, USA, Oct. 2005. doi: 10.1145/1090785.1090834
- [4] M. Alonso, A. Barreto, J. A. Jacko, and M. Adjouadi. A multi-domain approach for enhancing text display for users with visual aberrations. In *Assets '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, pp. 34–39. Association for Computing Machinery, New York, NY, USA, Oct. 2006. doi: 10.1145/1168987.1168995
- [5] P. Andersson, J. Nilsson, T. Akenine-Möller, M. Oskarsson, K. Åström, and M. D. Fairchild. Flip: A difference evaluator for alternating images. *Proc. ACM Comput. Graph. Interact. Tech.*, 3(2), aug 2020. doi: 10.1145/3406183
- [6] R. K. M.-G. F. A. K. Ashraf, Maliha and S. Wuerger. Colour difference formula for photopic and mesopic vision incorporating cone and rod responses. Society for Imaging Sciences and Technology, 2022.
- [7] G. Aydınođan, K. Kavaklı, A. Şahin, P. Artal, and H. Ürey. Applications of augmented reality in ophthalmology. *Biomedical optics express*, 12(1):511–538, 2021.
- [8] J. Beyazian and K. Akşit. complight/learned_prescription_model. https://github.com/complight/learned_prescription_model, 2022.
- [9] P. Chakravarthula, D. Dunn, K. Akşit, and H. Fuchs. Focusar: Auto-focus augmented reality eyeglasses for both real and virtual. *IEEE transactions on visualization and computer graphics*, 2018.
- [10] P. Chakravarthula, Z. Zhang, O. Tursun, P. Didyk, Q. Sun, and H. Fuchs. Gaze-contingent retinal speckle suppression for perceptually-matched foveated holographic displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4194–4203, 2021.
- [11] S. A. Cholewiak, G. S. Love, P. P. Srinivasan, R. Ng, and M. S. Banks. Chromablur: Rendering chromatic eye aberration improves accommodation and realism. *ACM Transactions on Graphics (TOG)*, 36, 2017. doi: 10.1145/3130800.3130815
- [12] R. Franzen. Kodak lossless true color image suite. source: <http://r0k.us/graphics/kodak>, 4(2), 1999.
- [13] A. Güzel and K. Akşit. complight/learned_prescription. https://github.com/complight/learned_prescription, 2022.
- [14] T. Hansen, L. Pracejus, and K. R. Gegenfurtner. Color perception in the intermediate periphery of the visual field. *Journal of vision*, 9(4):26–26, 2009.

- [15] F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb. High-quality computational imaging through simple lenses. *ACM Transactions on Graphics (TOG)*, 32(5):1–14, 2013.
- [16] F.-C. Huang, D. Lanman, B. A. Barsky, and R. Raskar. Correcting for optical aberrations using multilayer displays. *ACM Trans. Graphics*, 31(6):1–12, Nov. 2012. doi: 10.1145/2366145.2366204
- [17] F.-C. Huang, G. Wetzstein, B. A. Barsky, and R. Raskar. Eyeglasses-free display: towards correcting visual aberrations with computational light field displays. *ACM Trans. Graphics*, 33(4):1–12, July 2014. doi: 10.1145/2601097.2601122
- [18] Y. Itoh and G. Klinker. Vision enhancement: defocus correction via optical see-through head-mounted displays. In *AH '15: Proceedings of the 6th Augmented Human International Conference*, pp. 1–8. Association for Computing Machinery, New York, NY, USA, Mar. 2015. doi: 10.1145/2735711.2735787
- [19] Y. Itoh, T. Langlotz, S. Zollmann, D. Iwai, K. Kiyoshi, and T. Amano. Computational phase-modulated eyeglasses. *IEEE Transactions on Visualization and Computer Graphics*, 27(3):1916–1928, 2021. doi: 10.1109/TVCG.2019.2947038
- [20] D. Kim, S.-W. Nam, K. Bang, B. Lee, S. Lee, Y. Jeong, J.-M. Seo, and B. Lee. Vision-correcting holographic display: evaluation of aberration correcting hologram. *Biomed. Opt. Express*, 12(8):5179–5195, Aug 2021. doi: 10.1364/BOE.433919
- [21] J. Kim, M. Gopakumar, S. Choi, Y. Peng, W. Lopes, and G. Wetzstein. Holographic glasses for virtual reality. In *ACM SIGGRAPH 2022 Conference Proceedings*, pp. 1–9, 2022.
- [22] J. Kim, Y. Jeong, M. Stengel, K. Aksit, R. A. Albert, B. Boudaoud, T. Greer, J. Kim, W. Lopes, Z. Majercik, et al. Foveated ar: dynamically-foveated augmented reality display. *ACM Trans. Graph.*, 38(4):99–1, 2019.
- [23] M. S. J.-Y. W. B. B. J. S. K. A. R. A. e. a. Kim, Jonghyun. Matching prescription & visual acuity: Towards ar for humans. *ACM SIGGRAPH 2019 Emerging Technologies*, pp. 1–2, 2019.
- [24] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *arXiv*, Dec. 2014. doi: 10.48550/arXiv.1412.6980
- [25] G. A. Koulieris, K. Akşit, M. Stengel, R. K. Mantiuk, K. Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. In *Computer Graphics Forum*, vol. 38, pp. 493–519. Wiley Online Library, 2019.
- [26] P.-Y. Laffont, T. Martin, M. Gross, W. D. Tan, C. Lim, A. Au, and R. Wong. Rectifyeye: A vision-correcting system for virtual reality. In *SIGGRAPH ASIA 2016 VR Showcase*, SA '16. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2996376.2996382
- [27] V. Lakshminarayanan and A. Fleck. Zernike polynomials: a guide. *Journal of Modern Optics*, 58(7):545–561, 2011.
- [28] X. Li, J. Suo, W. Zhang, X. Yuan, and Q. Dai. Universal and flexible optical aberration correction using deep-prior based deconvolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2613–2621, 2021.
- [29] A. Maimone and J. Wang. Holographic optics for thin and lightweight virtual reality. *ACM Trans. Graphics*, 39(4):1–67, July 2020. doi: 10.1145/3386569.3392416
- [30] D. Mandl, P. M. Roth, T. Langlotz, C. Ebner, S. Mori, S. Zollmann, P. Mohr, and D. Kalkofen. Neural Cameras: Learning Camera Characteristics for Coherent Mixed Reality Rendering. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 508–516. IEEE, Oct. 2021. doi: 10.1109/ISMAR52148.2021.00068
- [31] C. Montalto, I. Garcia-Dorado, D. Aliaga, M. M. Oliveira, and F. Meng. A Total Variation Approach for Customizing Imagery to Improve Visual Acuity. *ACM Trans. Graphics*, 34(3):1–16, May 2015. doi: 10.1145/2717307
- [32] J. Nilsson and T. Akenine-Möller. Understanding ssim, 2020. doi: 10.48550/ARXIV.2006.13846
- [33] A. A. of Ophthalmology. Eye health statistics, Oct 2019.
- [34] J. Orlosky, M. Sra, K. Bektaş, H. Peng, J. Kim, N. Kos'myna, T. Höllerer, A. Steed, K. Kiyokawa, and K. Akşit. Telelife: The Future of Remote Living. *Front. Virtual Real.*, 0, 2021. doi: 10.3389/frvir.2021.763340
- [35] K. Oshima, K. R. Moser, D. C. Rompapas, J. E. Swan, S. Ikeda, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato. Sharpview: Improved clarity of defocused content on optical see-through head-mounted displays. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 173–181. IEEE, 2016.
- [36] N. Padmanaban, R. Konrad, and G. Wetzstein. Autofocals: Evaluating gaze-contingent eyeglasses for presbyopes. *Science Advances*, 5:eaav6187, 06 2019. doi: 10.1126/sciadv.aav6187
- [37] V. F. Pamplona, M. M. Oliveira, D. G. Aliaga, and R. Raskar. Tailored displays to compensate for visual aberrations. *ACM Trans. Graphics*, 31(4):1–12, July 2012. doi: 10.1145/2185520.2185577
- [38] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. 2017.
- [39] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. vol. 9351, pp. 234–241, 10 2015. doi: 10.1007/978-3-319-24574-4_28
- [40] B. P. Schmidt, M. Neitz, and J. Neitz. Neurobiological hypothesis of color appearance and hue perception. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 31 4:A195–207, 2014.
- [41] R. E. Stevens, D. P. Rhodes, A. Hasnain, and P.-Y. Laffont. Varifocal technologies providing prescription and VAC mitigation in HMDs using Alvarez lenses. In B. C. Kress, W. Osten, and H. Stolle, eds., *Digital Optics for Immersive Displays*, vol. 10676, pp. 142 – 158. International Society for Optics and Photonics, SPIE, 2018. doi: 10.1117/12.2318397
- [42] A. Stockman and L. T. Sharpe. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000. doi: 10.1016/S0042-6989(00)00021-3
- [43] H. Tanaka and H. Kawano. Image correction for improving visual acuity using zernike-based vision simulation. In *2021 20th International Symposium on Communications and Information Technologies (ISCIT)*, pp. 32–36. IEEE, 2021.
- [44] A. P. Venkataraman, P. Papadogiannis, D. Romashchenko, S. Winter, P. Unso, and L. Lundström. Peripheral resolution and contrast sensitivity: effects of monochromatic and chromatic aberrations. *JOSA A*, 36(4):B52–B57, 2019.
- [45] D. R. Walton, R. K. Dos Anjos, S. Friston, D. Swapp, K. Akşit, A. Steed, and T. Ritschel. Beyond blur: Real-time ventral metamers for foveated rendering. *ACM Transactions on Graphics*, 40(4):1–14, 2021.
- [46] D. R. Walton, K. Kavaklı, R. K. Dos Anjos, D. Swapp, T. Weyrich, H. Urey, A. Steed, T. Ritschel, and K. Akşit. Metameric varifocal holograms. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 746–755. IEEE, 2022.
- [47] D. Williams, N. Sekiguchi, and D. Brainard. Color, contrast sensitivity, and the cone mosaic. *Proceedings of the National Academy of Sciences*, 90(21):9770–9777, 1993.
- [48] J.-Y. Wu and J. Kim. Prescription ar: a fully-customized prescription-embedded augmented reality display. *Optics Express*, 28(5):6225–6241, 2020.
- [49] J.-Y. Wu and J. Kim. Prescription ar: a fully-customized prescription-embedded augmented reality display. *Opt. Express*, 28(5):6225–6241, Mar 2020. doi: 10.1364/OE.380945
- [50] S. Wuerger. Colour perception. In *London Imaging Meeting*, vol. 2022, pp. YZ–MK. Society for Imaging Science and Technology, 2022.
- [51] F. Xu and D. Li. Software based visual aberration correction for hmds. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 246–250. IEEE, 2018.
- [52] K. Yanny, K. Monakhova, R. W. Shuai, and L. Waller. Deep learning for fast spatially varying deconvolution. *Optica*, 9(1):96–99, Jan 2022. doi: 10.1364/OPTICA.442438
- [53] J. Ye, Y. Ji, M. Zhou, S. B. Kang, and J. Yu. Content aware image pre-compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7):1545–1558, 2018.
- [54] L. Zhang, X. Wu, A. Buades, and X. li. Color demosaicking by local directional interpolation and nonlocal adaptive thresholding. *Journal of Electronic Imaging - J ELECTRON IMAGING*, 20, 01 2011.

ChromaCorrect: Prescription Correction in Virtual Reality Headsets through Perceptual Guidance: Supplementary Material

Ahmet Güzel*
University of Leeds

Jeanne Beyazian†
University College London

Praneeth Chakravarthula‡
UNC Chapel Hill
Princeton University

Kaan Akşit§
University College London

Index Terms: Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 CONTRIBUTIONS

Here we list contributions from each author in this specific research work:

Ahmet Güzel:

- Coding an entire computational pipeline for color model and optimization.
- Compiling extensive study on evaluation.
- Compiling majority of the figures.
- Writing, review and edits for the entire documentation.

Jeanne Beyazian:

- Compiling a dataset and development and training of machine learning model.
- Helping with the figures.
- Writing sections related to machine learning.

Praneeth Chakravarthula:

- Mentorship (Ahmet Güzel).
- Steering the project towards Virtual Reality applications.
- Review and edits for the entire documentation.

Kaan Akşit:

- Initiation of project idea, this includes blueprints for projector and algorithmic designs, and early experimentations.
- Literature review and mentorship (Ahmet Güzel and Jeanne Beyazian).
- Writing, review and edits for the entire documentation.

*e-mail: od20ahg@leeds.ac.uk

†e-mail: jeanne.beyazian.21@ucl.ac.uk

‡e-mail: praneethc@princeton.edu

§e-mail: k.aksit@ucl.ac.uk

2 COLOR OPPONENCY

Human retinal cells could be broadly classified into rods and cones. Cone cells are primarily responsible for color perception in Human Visual System (HVS), and the cone density on our retinas peak at our fovea and drops sharply towards larger eccentricities [4]. Cone cells have three subtypes known as Short (S), Medium (M), and Long (L) cells, where each differs in sensitivity to wavelengths of light [6, 7]. These L, M, and S cones reduce wavelengths of incoming light into trichromat values by integrating them over their response functions [8]. According to earlier color opponency studies, HVS relies on comparison of L, M and S cone activation outputs [3]. A widely used modeling method in color opponency is S versus (M+L) for blue-yellow (BY) channel, and L versus M for red-green (RG) channel [2]. Schmidt *et al.* [5] proposes L versus (M+S) opponency for BY channel and M versus (L+S) opponency for RG channel.

3 LEARNED MODEL

Figure 1 shows further results for comparison of the corrected image between our open-box model and the neural network’s prediction after over 800 epochs of training.

4 FURTHER EVALUATION

Figure 2 shows further results for different myopia starting from -2.0 dioptres to -4.5 dioptres. We show that our method’s reconstructed images have better contrast than conventional method’s reconstructed images for each case. Images are simulated in LMS space.

Figure 3 shows further results for different images from DIV2K image dataset [1]. In each refractive eye problem modelling, +/-1.5D refractive error is used to model prescriptions.

REFERENCES

- [1] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.
- [2] D. H. B. Andrew Stockman. Color vision mechanisms. In M. Bass, ed., *Vision and Vision Optics 3rd*, pp. 1–104. McGraw-Hill, New York, 2009.
- [3] L. M. Hurvich and D. Jameson. An opponent-process theory of color vision. *Psychological review*, 64, Part 1 6:384–404, 1957.
- [4] J. B. Jonas, U. Schneider, and G. O. Naumann. Count and density of human retinal photoreceptors. *Graefes Archive for Clinical and Experimental Ophthalmology*, 230(6):505–510, 1992.
- [5] B. P. Schmidt, M. Neitz, and J. Neitz. Neurobiological hypothesis of color appearance and hue perception. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 31 4:A195–207, 2014.
- [6] A. Stockman and L. T. Sharpe. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000. doi: 10.1016/S0042-6989(00)00021-3
- [7] A. Stockman, L. T. Sharpe, and C. Fach. The spectral sensitivity of the human short-wavelength sensitive cones derived from thresholds and color matches. *Vision Research*, 39(17):2901–2927, 1999. doi: 10.1016/S0042-6989(98)00225-9
- [8] S. Wuerger. Colour perception. In *London Imaging Meeting*, vol. 2022, pp. YZ–MK. Society for Imaging Science and Technology, 2022.

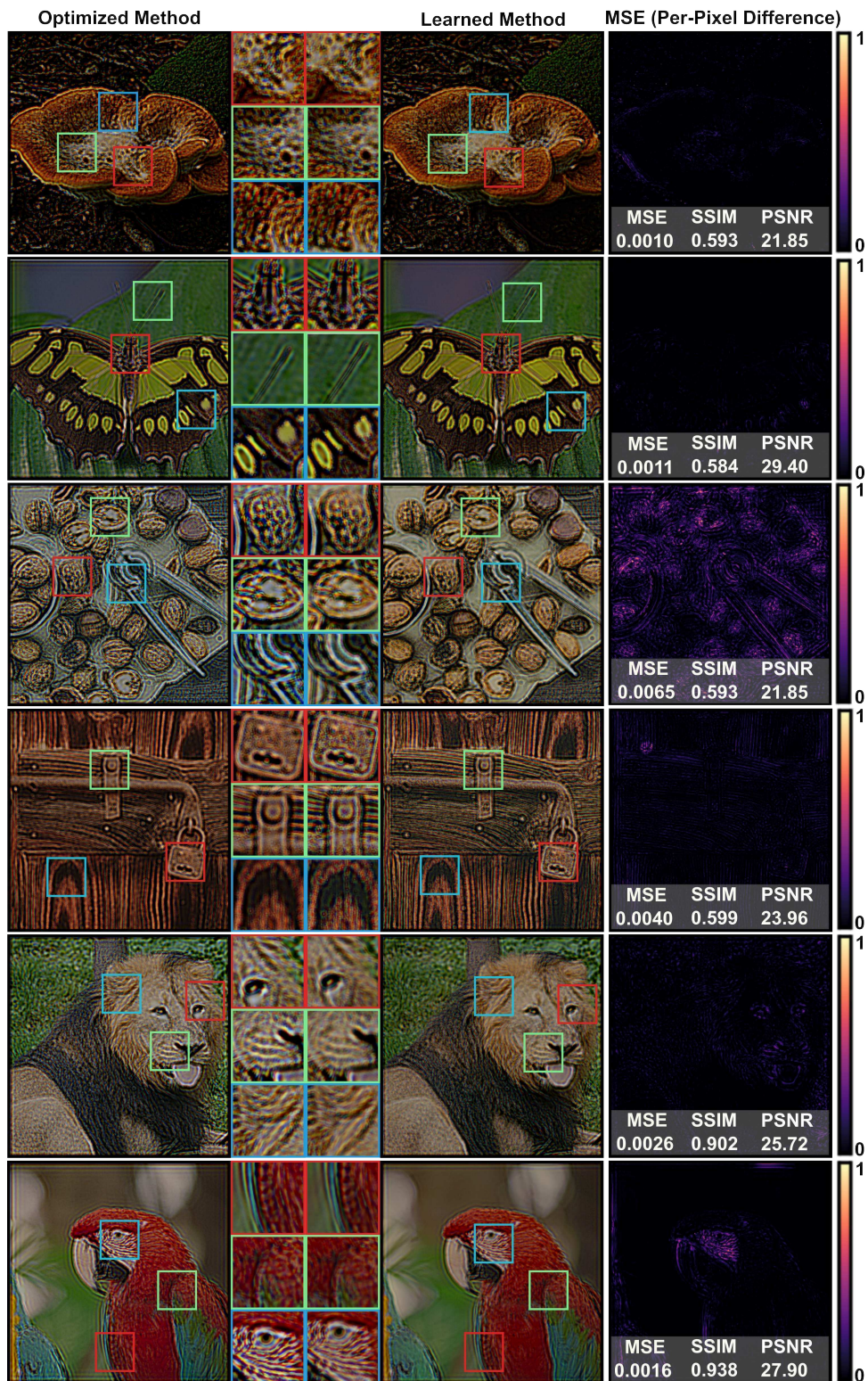


Figure 1: Additional images for learned method. Images are selected from DIV2K image dataset [1].

Uncorrected	Target Image	Conventional Method	Our Method
		 FLIP SSIM PSNR 0.15 0.835 28.01	 FLIP SSIM PSNR 0.08 0.895 31.10
		 FLIP SSIM PSNR 0.16 0.829 27.83	 FLIP SSIM PSNR 0.09 0.886 30.69
		 FLIP SSIM PSNR 0.17 0.825 27.62	 FLIP SSIM PSNR 0.11 0.879 30.21
		 FLIP SSIM PSNR 0.19 0.802 26.48	 FLIP SSIM PSNR 0.13 0.856 29.17
		 FLIP SSIM PSNR 0.21 0.798 26.17	 FLIP SSIM PSNR 0.14 0.849 28.67

Figure 2: Comparison of conventional method and our approach with different myopia cases. Images are simulated in LMS space.


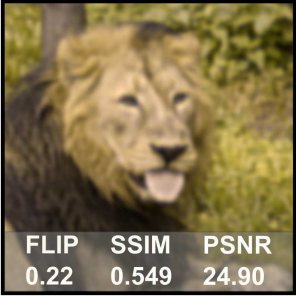

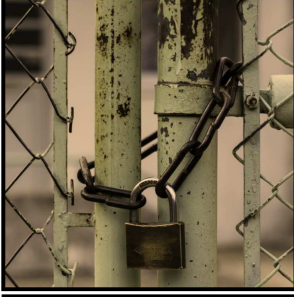


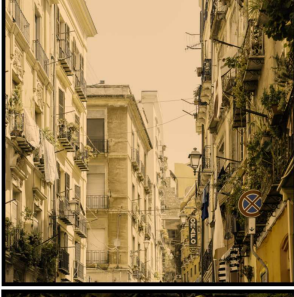

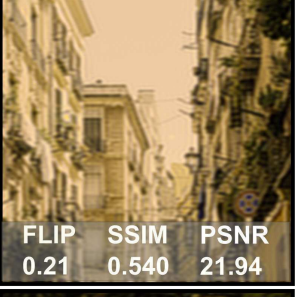
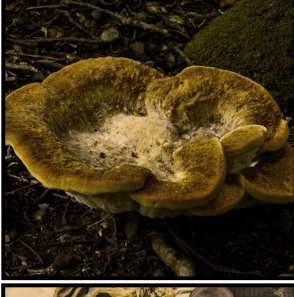
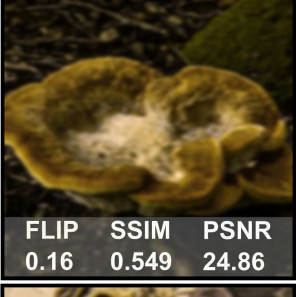


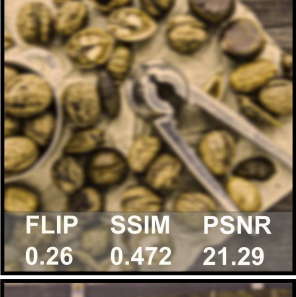
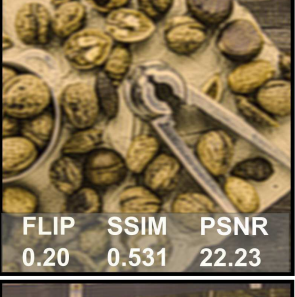


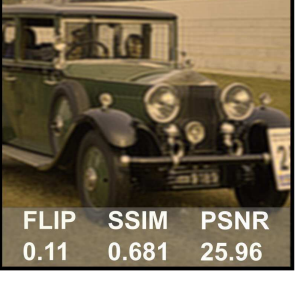
Target Image	Conventional Method	Our Method
	 FLIP SSIM PSNR 0.22 0.549 24.90	 FLIP SSIM PSNR 0.16 0.608 26.28
	 FLIP SSIM PSNR 0.22 0.515 23.54	 FLIP SSIM PSNR 0.16 0.586 25.10
	 FLIP SSIM PSNR 0.28 0.492 20.41	 FLIP SSIM PSNR 0.21 0.540 21.94
	 FLIP SSIM PSNR 0.16 0.549 24.86	 FLIP SSIM PSNR 0.13 0.587 25.65
	 FLIP SSIM PSNR 0.26 0.472 21.29	 FLIP SSIM PSNR 0.20 0.531 22.23
	 FLIP SSIM PSNR 0.17 0.639 24.27	 FLIP SSIM PSNR 0.11 0.681 25.96

Figure 3: Additional images for the evaluation section. Images are simulated in LMS space.