THESIS

EMPIRICAL EVALUATION OF VIRTUAL FONTS WITH ENHANCED OUT-OF-FOCUS LEGIBILITY IN THE PRESENCE OF INACCURATE OR INCOMPLETE EYE ACCOMMODATION DEMAND

Submitted by

Junhwan Kim

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2026

Master's Committee:

Advisor: Mohammed Safayet Arefin

Carlos Montalto
Anne Cleary

ABSTRACT

EMPIRICAL EVALUATION OF VIRTUAL FONTS WITH ENHANCED OUT-OF-FOCUS LEGIBILITY IN THE PRESENCE OF INACCURATE OR INCOMPLETE EYE ACCOMMODATION DEMAND

One of the key challenges in optical see-through (OST) augmented reality (AR) environments is accommodation mismatch caused by the fixed focal plane of commercial head-mounted displays (HMDs). Devices such as the Microsoft HoloLens 2 render all virtual content at a fixed distance of two meters. So, retinal blur occurs when users attempt to view virtual content rendered at different distances. This blur significantly impairs text legibility and can lead to serious consequences in fields where accurate information transmission is critical, such as healthcare, manufacturing, and navigation. SharpView addresses this issue through a pre-compensation method that combines Zernike polynomial-based blur modeling with Total Variation (TV) optimization, ensuring text remains sharp even under blurring conditions. This thesis further improved the algorithm in two respects. First, the fourth-order Zernike term was incorporated into the OTF generation process to model the spherical aberration of the human eye more accurately. Second, a white masking technique was introduced alongside the existing black masking method to eliminate experimental bias caused by brightness differences between SharpView images and the Control. Using this improved algorithm and the Microsoft HoloLens 2, a commercial OST AR device, a human subject experiment with 16 participants was conducted under two accommodation mismatch conditions. Incomplete Accommodation refers to a condition where virtual content appears before the eyes have fully focused on the new distance, while Incorrect Accommodation refers to a condition where the eyes focus on the display's fixed focal plane rather than the actual rendering distance of the virtual content, resulting in persistent blurriness. The experiment considered three letters (A, E, and O), three levels of CL (%) (70%, 80%, 90%), and two levels of τ (high, low). Two focal disparities were

considered (3.5Diopters and 4.5Diopters). The experimental results showed an overall SharpView selection rate of 32.1% in Experiment 1 (Incomplete Accommodation) and 22.6% in Experiment 2 (Incorrect Accommodation), with participants generally preferring the Control. This tendency is consistent with the mere-exposure effect, in which familiarity with a stimulus increases preference independent of its objective quality. Nevertheless, under certain parameter combinations, specifically CL levels of 70% and 80%, High τ, and the letters A and E, SharpView's selection rate exceeded chance level, although these differences did not reach statistical significance after Bonferroni correction. High τ consistently increased the SharpView selection rate across both aberration conditions and both experiments, whereas a CL of 90% consistently reduced it due to excessive ringing. SharpView was more effective for letters composed of straight or diagonal lines (A, E) than for the curved letter O, although this pattern was not confirmed under the 4.5D condition of Experiment 1. These findings indicate that SharpView, in its current form, did not clearly outperform the Control in either experiment. However, by identifying the conditions under which SharpView is ineffective, this study provides practical parameter guidelines for improving text legibility in future OST AR systems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Augmented Reality

Augmented Reality (AR) is a technology that overlays virtual information and objects onto the real world for the user to see. AR can be implemented in various forms, such as smartphones, tablets, and Head-Mounted Displays (HMDs). Among these, HMDs, worn on the user's head, offer the most immersive AR experience [1].

There are two main methods for displaying virtual information on an HMD. The first is the video see-through (VST) method, which overlays virtual information onto the video feed of the real-world environment captured by an external camera. While this approach offers the advantage of freely combining virtual information with the real environment, it has the downside of potential latency or reduced realism, as the user views the real world indirectly through the camera.

The second method is optical see-through (OST), which allows users to view the real world directly through semi-transparent optical lenses while simultaneously seeing virtual information superimposed upon it. Because OST enables a direct view of the real world, it offers the advantages of high realism and zero latency. The Microsoft HoloLens 2 is a prominent commercial OST HMD and was used in this study.

1.2 Augmented Reality in Real Life

AR is utilized in various real-world fields and enhances work efficiency and accuracy by providing users with useful information in real time.

In the medical field, augmented reality (AR) provides surgeons with real-time information by overlaying 3D data, including MRI, CT scans, and ultrasound images, onto the actual surgical environment. This facilitates minimally invasive surgery by enabling X-ray vision, a capability that allows virtual information to be displayed behind opaque physical objects.

In the field of education and training, AR provides a more realistic training experience by overlaying 3D models of complex equipment onto the real-world environment. Compared to traditional training methods that rely on text or images, AR delivers superior learning outcomes by enabling direct observation and interaction.

In the fields of manufacturing and maintenance, AR enables workers to view work instructions overlaid on the equipment while inspecting it. This enhances work accuracy and efficiency while reducing error rates.

In the field of navigation, AR helps users travel efficiently to their destinations by overlaying route guidance information onto the real-world environment of indoor and outdoor spaces.

A common and critical factor across all the aforementioned fields is the visibility, or legibility, of the text and information displayed via AR. Poor text legibility in AR systems can lead users to miss or misread information, potentially resulting in catastrophic consequences in sectors where accuracy is paramount, such as healthcare, manufacturing, and navigation. Consequently, enhancing text legibility in AR environments is a vital area of research.

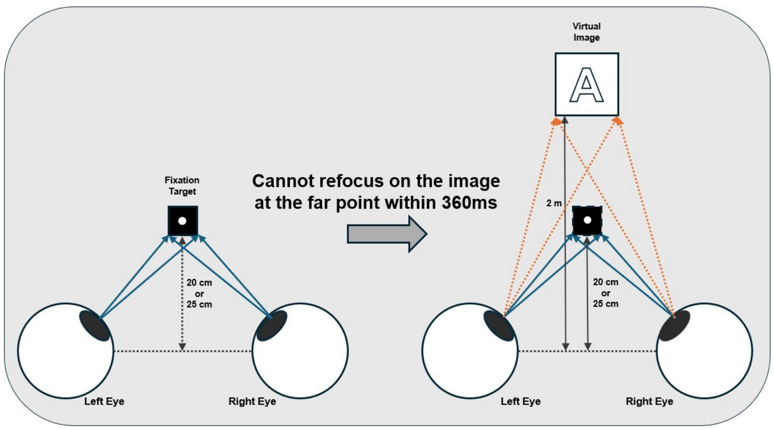
1.3 Accommodation Mismatch

Commercial OST HMDs (e.g., Microsoft HoloLens 2, Magic Leap 2) feature a single fixed focal plane. This means that virtual information is always rendered at a specific, fixed distance. For instance, the HoloLens 2 has a fixed focal plane of 2 meters, so all virtual information is rendered at that distance.

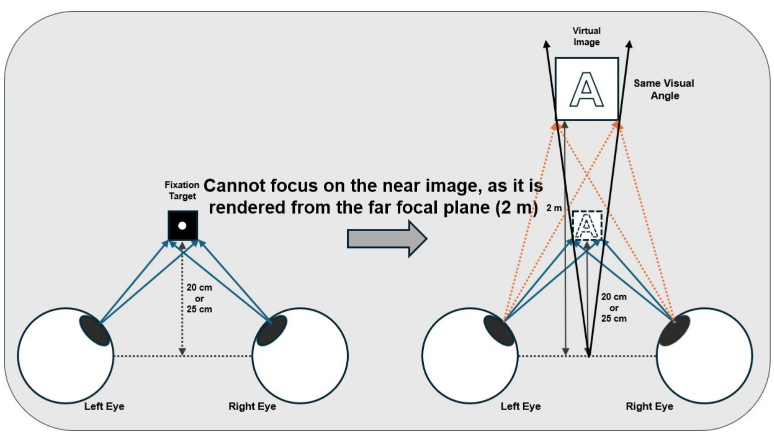
The human eye possesses the ability to automatically adjust its focus based on distance, a process known as accommodation. While the eye can naturally focus on objects at varying distances, the fixed focal plane in an OST HMD environment leads to a phenomenon called accommodation mismatch. This accommodation mismatch causes the retinal image to blur, as the eye attempts to focus at one distance while the display renders at another, spreading light across the retina and reducing the sharpness of displayed text.

1.4 Two Types of Accommodation Mismatch

This accommodation mismatch manifests in two forms. Incomplete Accommodation refers to a situation where virtual information is displayed before the eyes have fully adjusted their focus to the new distance. Achieving full accommodation takes approximately 360–380 ms. Incorrect Accommodation refers to a situation where, even given sufficient time, the eyes focus on the display's fixed focal plane rather than the actual rendering distance, resulting in persistent blurriness. These two forms of accommodation mismatch are illustrated in Figure 1.1.



(a) Incomplete Accommodation.



(b) Incorrect Accommodation.

Figure 1.1: Two forms of accommodation mismatch in OST AR environments. (a) Incomplete Accommodation: the participant fixates on the fixation target at 20 cm or 25 cm. When the virtual image is displayed at the HoloLens 2 focal plane (2 m), the eyes cannot refocus on the image at the far point within 360 ms, resulting in transient blur. (b) Incorrect Accommodation: although the participant fixates at 20 cm or 25 cm, the virtual image is rendered from the far focal plane (2 m) at the same visual angle. The eyes cannot focus on the near image, as it is rendered from the far focal plane (2 m), resulting in persistent blur.

1.5 Problem Statement

In fields where the accurate transmission of information is critical, such as healthcare, education, and manufacturing, impaired text legibility can lead to disastrous consequences. Therefore, addressing the issue of reduced legibility caused by accommodation mismatch in OST HMDs is a significant research challenge. Hardware-based approaches, such as varifocal lenses and light-field displays, have been proposed as existing solutions. While effective, these methods face limitations, including high costs and implementation complexity, that make them difficult to apply to commercial devices. More fundamentally, hardware-based solutions are limited by their dependence on specific devices, preventing them from being universally applied to other commercial devices. Software-based approaches have been proposed as an alternative [2, 3].

Arefin et al. [4] proposed an improved SharpView font by combining out-of-focus blur modeling using Zernike polynomials with a correction algorithm based on constrained TV optimization. SharpView models blur using Zernike polynomials and corrects images through Total Variation (TV)-based deconvolution, thereby making text appear clearer even under out-of-focus conditions. However, while Oshima et al. [3] and Cook et al. [5] conducted experiments with human subjects using the initial SharpView algorithm, the improved SharpView proposed by Arefin et al. [4] was validated solely through camera-based evaluation, without experiments involving human subjects using commercial OST AR devices. Camera-based evaluation is commonly used in the initial validation stage because the experimental environment is simple and reproducible, and the characteristics of the optical system can be measured objectively. Yet, since cameras cannot perfectly replicate the biological characteristics of the human eye, results obtained via cameras cannot be directly applied to actual human vision. Therefore, it is necessary to verify whether SharpView is effective in improving actual human legibility.

1.6 Research Goals

This study aims to evaluate the effect of the SharpView font on human legibility. To this end, experiments were conducted in an OST AR environment using the Microsoft HoloLens 2 under two conditions of accommodation mismatch: Incomplete Accommodation and Incorrect Accommodation. The findings of this study are expected to provide practical guidelines for improving text legibility in OST AR systems. To achieve these objectives, the following two research questions were formulated.

RQ1: Does the SharpView font improve human legibility under Incomplete Accommodation and Incorrect Accommodation conditions?

RQ2: If the SharpView font improves legibility, which combination of algorithmic parameters and letter shapes is most effective?

By answering these two research questions, we will be able to validate the effectiveness of the SharpView font and propose an optimal parameter combination to enhance text legibility in future OST AR systems. Furthermore, the findings of this study are expected to contribute to increasing the practical utility of AR technology across various fields, such as healthcare, education, and manufacturing.

Chapter 2

Background and Related Works

2.1 OST AR Displays

An OST AR display is a device that allows users to view the real world directly through semi-transparent optical lenses while simultaneously seeing virtual information superimposed upon it [1], as introduced in Section 1.1. Representative commercial OST HMDs include the Microsoft HoloLens 2 and Magic Leap 2.

However, current commercial OST HMDs suffer from a fundamental limitation: they possess a single, fixed focal plane. For instance, the HoloLens 2 renders all virtual information at a distance of 2 meters [6]. Consequently, when users attempt to view objects at distances other than 2 meters, the virtual information appears blurred. In an experiment involving precision manual tasks using the HoloLens, Condino et al. [7] demonstrated that this fixed focal plane caused a significant decline in task accuracy compared to viewing with the naked eye. This suggests that the fixed focal plane issue in OST AR can have a serious impact on actual task performance, extending beyond mere inconvenience.

2.2 Accommodation Mismatch in OST AR

The human eye possesses the ability to automatically adjust its focus based on distance, a process known as accommodation. Accommodation is achieved through the ciliary muscles surrounding the lens, which alter the lens's thickness. While vergence and accommodation operate in tandem in natural visual environments, this coupling is disrupted in OST AR environments. This occurs because the display's fixed focal plane requires the eyes to accommodate to a constant distance, whereas vergence must converge to align with the rendering distance of virtual objects. This phenomenon is known as the vergence-accommodation conflict (VAC) [8] and is known to cause visual fatigue, discomfort, and reduced performance [8, 9].

As introduced in Section 1.4, accommodation mismatch in OST AR manifests as Incomplete Accommodation and Incorrect Accommodation. According to Campbell and Westheimer [10] and Gabbard et al. [9], it takes approximately 360 ms for the human eye to accommodate from far to near, and approximately 380 ms from near to far, when shifting focus from one distance to another. Consequently, if virtual information is presented for only a very brief period, the eyes must process the information without having fully achieved focus. Arefin et al. [11] demonstrated that this transient focal blur significantly affects user performance in text-based visual search tasks.

Regarding Incorrect Accommodation, Gabbard et al. [9] experimentally demonstrated the impact of focal distance switching on user performance, providing empirical support for the persistent blur that arises when the eyes accommodate to the display's fixed focal plane rather than the actual rendering distance of virtual content.

Both conditions impair the legibility of AR text and can lead to serious consequences in fields where the accurate transmission of information is critical, such as healthcare, manufacturing, and navigation. These two forms of accommodation mismatch are illustrated in Figure 1.1.

2.3 Hardware-Based Approaches to Address Accommodation Mismatch

Hardware-based approaches to resolving the accommodation mismatch are broadly categorized into multifocal displays, varifocal displays, and light field displays [12].

A varifocal display is a system that adjusts the focal plane in real time based on the user's gaze direction. Chakravarthula et al. [13] proposed FocusAR, which utilizes a liquid lens to create AR glasses capable of automatic focus adjustment for both the real world and virtual information. Similarly, Xia et al. [14] proposed a switchable AR/VR display that simultaneously addresses the accommodation-vergence conflict and supports prescription vision correction.

A light field display is a system that simultaneously presents light rays from multiple directions, allowing the eyes to focus at a desired distance naturally. Maimone and Fuchs [15] applied this approach to an AR environment using computational AR eyeglasses.

However, these hardware-based approaches share common limitations, including high costs, implementation complexity, a small eye box, and low resolution [12], and have largely remained at the optical workbench prototype stage, making them difficult to apply to commercial devices.

2.4 Software-Based Approaches to Address Accommodation Mismatch

As introduced in Section 1.5, software-based methods offer an alternative to the cost and complexity of hardware-based solutions. These methods work by mathematically modeling optical blur and rendering a pre-corrected image that reverses the blur, thereby ensuring the image appears sharp even when out of focus.

Montalto et al. [2] proposed an image deconvolution approach based on constrained Total Variation (TV) as a method for customizing images for individuals with refractive vision problems. This method demonstrated greater effectiveness in enhancing the visual clarity of text information compared to conventional techniques such as Wiener filtering. Subsequently, it served as a key technique in research on AR out-of-focus correction.

Software-based approaches for AR out-of-focus correction remain a relatively underexplored area. Oshima et al. [3] proposed the initial SharpView algorithm based on a Gaussian PSF and Wiener filtering. However, it failed to adequately account for the optical characteristics of human vision and did not yield significant results for short AR text [5]. Cook et al. [5] conducted user preference experiments and found a significant user preference for SharpView-enhanced text under non-fixated viewing conditions. However, the preference for SharpView-enhanced text was limited, particularly for short AR text labels. Meanwhile, in a VR environment, Xu and Li [16] proposed a software-based visual aberration correction method combining Zernike polynomials and TV deconvolution.

Based on these prior studies, Arefin et al. [4] developed the SharpView font, which combines out-of-focus blur modeling using Zernike polynomials with a correction algorithm based on constrained TV optimization. In camera-based evaluations, the SharpView font demonstrated a

40%–44% improvement in sharpness through synthetic simulation, and a 24%–32% improvement through camera-based evaluation, compared to standard fonts. However, since cameras cannot fully replicate the biological characteristics of the human eye, there are limitations to directly applying camera-based results to actual human vision. Therefore, validation involving human subjects is necessary, and this serves as the primary motivation for the current study.

Chapter 3

SharpView Algorithm

3.1 Algorithm Overview

The images used in this study were generated based on the SharpView technique proposed by Arefin et al. [4]. While previous research validated the effectiveness of SharpView solely through camera-based evaluation, this study modified and improved the existing algorithm in two respects to conduct experiments involving human participants. First, the spherical aberration of the eye was modeled more accurately by incorporating the fourth-order Zernike term into the Optical Transfer Function (OTF) generation process. Second, a white masking technique was added to the existing black masking method to eliminate experimental bias caused by brightness differences between the SharpView and Control images (original Arial Bold font). The overall algorithm proceeds as follows: wavefront aberration modeling using Zernike polynomials, OTF conversion, generation of a pre-corrected image via TV optimization, masking, and final image selection. The pseudocode for the SharpView image generation algorithm used in this study is presented in Algorithm 1. Examples comparing a Control image and a SharpView image generated using this algorithm are shown in Figure 3.1. The SharpView image incorporates ringing artifacts, intentionally introduced to compensate for out-of-focus blur, and is designed to be perceived more clearly under conditions of accommodation mismatch.

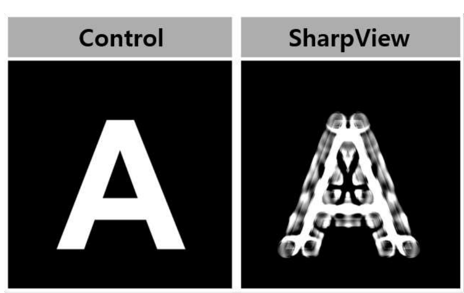


Figure 3.1: Comparison of Control (left) and SharpView-processed image (right) for the letter A. The SharpView image contains ringing artifacts designed to compensate for out-of-focus blur, resulting in a sharper percept when viewed under accommodation mismatch conditions.

Algorithm 1 SharpView Image Generation

```
procedure Main(I, focusDistance, defocusDistance)
    for each optical condition (pupil_diam, CL) do
        H ← ComputeOTF(focusDistance, defocusDistance, pupil_diam)
        tc ← AdjustContrast(I, CL)
        for each target t = 0 → 50 do
            u_PrecorrectionImageGeneration ← PrecorrectionImageGeneration(tc, H, θ_min)
            [u_test, u_without_mask] ← Masking(u_PrecorrectionImageGeneration, I, focusDistance,
                                           defocusDistance, pupil_diam, α, threshold)
            τ_max ← |TV(u_test) − TV(I)|/TV(I)
            if τ_max < t − 0.25 then
                break
            end if
            repeat
                u_PrecorrectionImageGeneration ← PrecorrectionImageGeneration(tc, H, θ)
                [u, u_without_mask] ← Masking(u_PrecorrectionImageGeneration, I, focusDistance,
                                          defocusDistance, pupil_diam, α, threshold)
                τ ← |TV(u) − TV(I)|/TV(I)
                if τ < t then
                    decrease θ
                else if τ > t then
                    increase θ
                end if
            until |τ − t| ≤ 0.25
            SaveResult(u, τ, θ, pupil_diam, CL)
        end for
    end for
end procedure
```

3.2 Point Spread Function (PSF)

The PSF represents the image formed on the retina when light from a point light source passes through the eye's optical system. It indicates the quality of the retinal image and serves as a mathematical representation of the eye's optical aberrations [17]. The degree of blur depends on the pupil diameter and the focal distance mismatch; the larger the pupil and the greater the difference between the actual focal distance and the rendering distance, the more the PSF spreads and the more severe the blur becomes.

In this study, the PSF is calculated during the blur modeling stage of the SharpView algorithm and combined with the Zernike polynomial-based wavefront aberration model, described in the following sections (Section 3.3 and Section 3.4), to generate the OTF.

3.3 Zernike Polynomials

Zernike polynomials are continuous, orthogonal polynomials defined on the unit circle that provide a mathematical representation of wavefront aberrations in the eye. Each Zernike term corresponds to a specific type of optical aberration (e.g., defocus, astigmatism, coma, spherical aberration, etc.), and the hierarchical structure of these polynomials is illustrated in Figure 3.2.

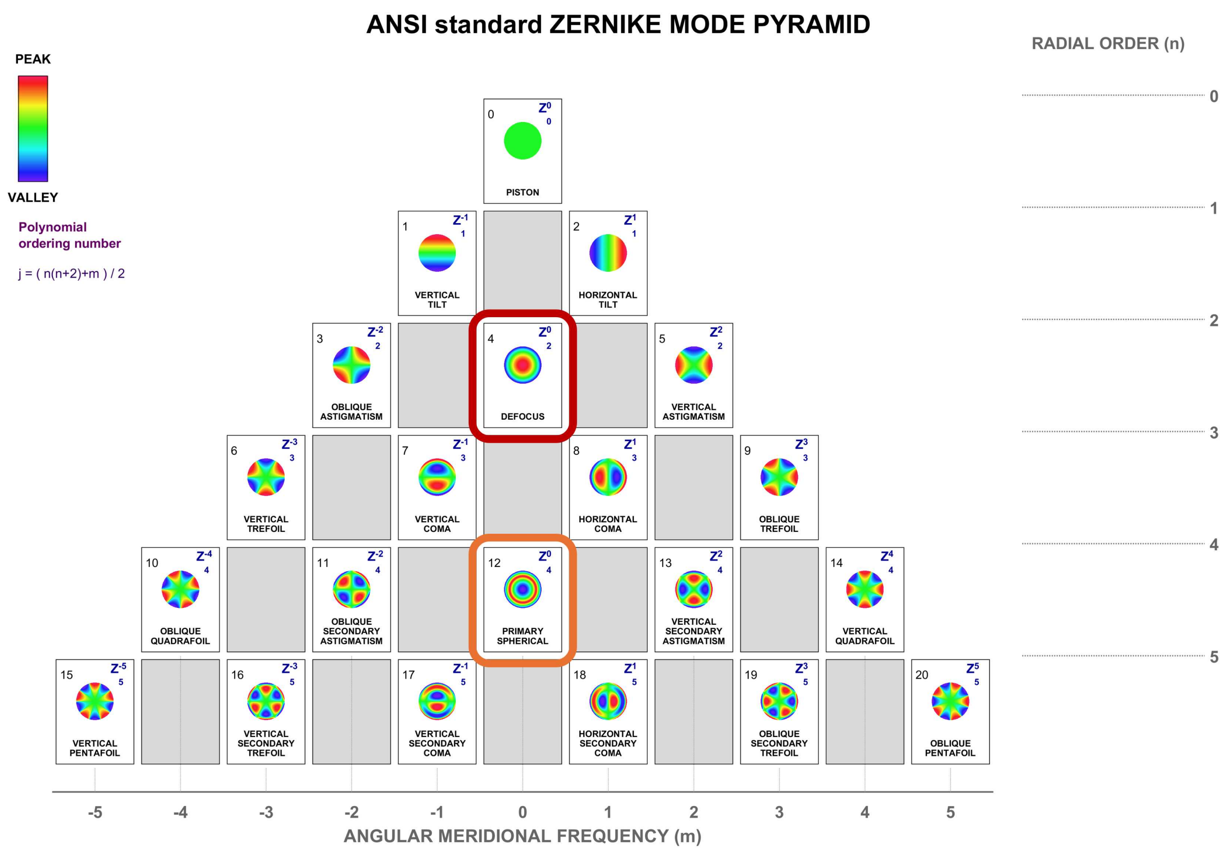


Figure 3.2: ANSI standard Zernike mode pyramid, showing Zernike terms organized by radial order and angular meridional frequency [18].

Zernike polynomials are classified based on radial order n and angular frequency m, and this ordering system serves as a basis for distinguishing various ocular aberrations. Aberrations with a radial order n of less than 3, such as defocus and astigmatism, are termed low-order aberrations and are the primary causes of refractive errors like myopia, hyperopia, and astigmatism. Aberrations with an n of 3 or greater, such as coma, trefoil, and spherical aberration, are referred to as high-order aberrations [18, 19].

Despite their various advantages, Zernike polynomials have certain limitations. Since Zernike coefficients are valid only for the specific pupil diameter used in their calculation, computing coefficients for every possible pupil diameter entails high computational cost. Furthermore, because Zernike polynomials are orthogonal over the unit circle, they are unsuitable for optical systems that are non-circular or irregular (e.g., keratoconus) [20]. Consequently, Zernike polynomials are best suited for optical systems that can be represented as circular, such as the human eye, telescopes, and microscopes.

Beyond representing ocular aberrations, Zernike polynomials have been utilized in various fields. Notably, they are widely used to measure image moments (weighted averages of pixel intensities), referred to as Zernike moments, which are useful for feature extraction and shape description. These moments have contributed significantly to cancer research by enabling the analysis and quantification of cancer cell morphology [21]. They are also extensively employed in astronomy to fit wavefronts distorted by atmospheric turbulence [22]. Thus, Zernike polynomials have established themselves as a standard tool not only in ophthalmology and vision science but also in numerous research areas requiring the measurement of optical aberrations and optical testing.

Traditionally, Zernike defocus coefficients are calculated from an individual's spectacle prescription data, specifically the sphere, cylinder, and axis values [18, 19]. However, this method requires specific prescription data, which may be unavailable or irrelevant, to model the artificial defocus caused by focal distance shifts in AR. To address this, Xu and Li developed a method to calculate the defocus coefficient (S-value) using only the viewing distance or focused distance, rather than prescription data, to model low-order visual aberrations in head-mounted displays [16].

Building on this distance-based approach, Arefin developed a method to derive defocus coefficients using the thin-lens formula based on the focused distance and the out-of-focus distance, specifically to address the out-of-focus issue in AR.

The x and y coordinates of the wavefront aberration function are referenced to the pupil center and normalized by the pupil radius [23]. Since aberrations in the left and right eyes tend to exhibit similar characteristics, the distance-based defocus model employed in this study is not specific to a particular eye and can be applied to both eyes using the same methodology [24].

The existing method (Arefin et al. [4]) modeled wavefront aberration using only the second-order Zernike term ($C2$, defocus). However, the eye exhibits spherical aberration in addition to defocus. Spherical aberration is a phenomenon in which light is refracted differently at the center and periphery of the lens (the eye's lens), and it becomes more pronounced as the pupil size increases. Therefore, this study incorporated the fourth-order Zernike term ($C4$) to model ocular blur more accurately. $C2$ is calculated as follows:

$$C2 = 10^{-3} \times pupilRadius^2 \times diopter/(4\sqrt{3}) \tag{3.1}$$

Here, the diopter represents the difference in refractive power between $focusDistance$ and $defocusDistance$. $C4$ is calculated as follows:

$$C4 = -10^{-9} \times \sqrt{5}/240 \times pupilRadius^4 \times (D1^3 - D2^3) \tag{3.2}$$

$D1$ is the reciprocal of $focusDistance$, and $D2$ is the reciprocal of $defocusDistance$. The final wavefront aberration function W is calculated as the sum of the weighted second- and fourth-order Zernike terms:

$$W = (C2 \times z2) + (C4 \times z4) \tag{3.3}$$

Here, $z2$ and $z4$ are, respectively, the second- and fourth-order Zernike basis functions.

3.4 OTF Generation

To use the wavefront aberration W defined in Section 3.3, one must first determine the normalized frequency coordinates (F_u, F_v) in which the Zernike basis functions $z2$ and $z4$ are defined. The size of the image formed on the retina is calculated using the display size and the eye-to-retina distance (d_0, approximately 22 mm) based on the principle of triangle similarity:

$$iRetinaSizeWidth = DisplayWidth \times d_0 / focusDistance \tag{3.4}$$

Based on this retinal image size, the frequency domain is discretized to generate f_u and f_v. Next, the cutoff frequency f_0, representing the diffraction limit, is calculated using the pupil diameter:

$$f_0 = pupilDiameter / (\lambda \times d_0) \tag{3.5}$$

f_u and f_v are normalized using this f_0 to obtain F_u and F_v:

$$F_u = f_u / f_0, \quad F_v = f_v / f_0 \tag{3.6}$$

On the normalized F_u and F_v, the Zernike basis functions $z2$ and $z4$ (Section 3.3) are calculated in the form defined on the unit circle, and the wavefront aberration function W is thereby obtained. Next, the pupil aperture A is defined as the region where F_u and F_v lie within the unit circle. The pupil function P is calculated using this pupil aperture and the wavefront aberration W as follows:

$$P = A \cdot e^{-i \cdot 2\pi W / \lambda} \tag{3.7}$$

Here, A is the pupil aperture (the area through which light passes), and λ is the wavelength of the light. The PSF is calculated as the squared magnitude of the Fourier transform of the pupil function:

$$PSF = |\mathcal{F}(P)|^2 \tag{3.8}$$

Finally, the PSF is normalized in the frequency domain to obtain the OTF:

$$OTF = PSF / \sum PSF \tag{3.9}$$

This normalization is intended to set the total energy of the PSF to 1, ensuring that the image brightness remains unchanged. The OTF calculated in this manner is used as the blur kernel k for the TV optimization described in Section 3.6.

Algorithm 2 ComputeOTF

function COMPUTEOTF($focusDistance, defocusDistance, pupil\_diam$)
    $pupilRadius \leftarrow pupil\_diam/2$
    $diopter \leftarrow |1/focusDistance - 1/defocusDistance|$
    $C2 \leftarrow 10^{-3} \times pupilRadius^2 \times diopter \ / \ (4\sqrt{3})$
    $D1 \leftarrow 1/focusDistance, \quad D2 \leftarrow 1/defocusDistance$
    $C4 \leftarrow -10^{-9} \times \sqrt{5}/240 \times pupilRadius^4 \times (D1^3 - D2^3)$
    $iRetinaSizeWidth \leftarrow DisplayWidth \times d_0/focusDistance$
    $iRetinaSizeHeight \leftarrow DisplayHeight \times d_0/focusDistance$
    $(f_u, f_v) \leftarrow$ discretized frequency coordinates with spacing $1/iRetinaSizeWidth$, $1/iRetinaSizeHeight$ ▷ Retinal image discretization
    $f_0 \leftarrow pupil\_diam \ / \ (\lambda \times d_0)$ ▷ Diffraction cutoff frequency
    $F_u \leftarrow f_u/f_0, \quad F_v \leftarrow f_v/f_0$ ▷ Normalized frequency coordinates
    $W \leftarrow C2 \cdot z2(F_u, F_v) + C4 \cdot z4(F_u, F_v)$ ▷ Wavefront aberration
    $A \leftarrow \mathbf{1}[\sqrt{F_u^2 + F_v^2} \leq 1]$ ▷ Pupil aperture
    $P \leftarrow A \cdot e^{-i \cdot 2\pi W/\lambda}$ ▷ Pupil function
    $PSF \leftarrow |\mathcal{F}(P)|^2$
    $H \leftarrow PSF / \sum PSF$ ▷ Normalized OTF
    return H
end function

3.5 Blur Simulation Results

To illustrate the degree of blur actually produced by the OTF (PSF) calculated in Section 3.4, this section presents a simulation showing the result of convolving the original image with the PSF. This simulation is not part of the SharpView image generation pipeline but serves as a reference to illustrate the extent of blur caused by the OTF. The result of convolving the original image I with the PSF can be expressed as follows:

$$I_{blurred} = I \otimes PSF \tag{3.10}$$

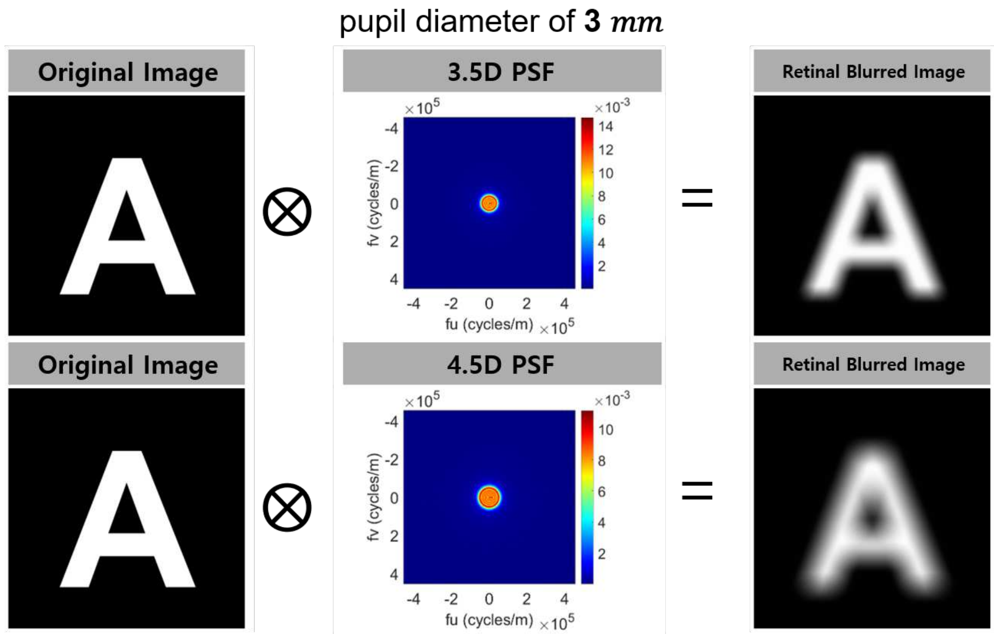


Figure 3.3: Simulation of retinal blur for the letter A at a pupil diameter of 3 mm, provided for illustrative purposes. The original image is convolved with the PSF computed for the 3.5D condition (top, $focusDistance = 2$ m, $defocusDistance = 0.25$ m) and the 4.5D condition (bottom, $focusDistance = 2$ m, $defocusDistance = 0.20$ m), producing the corresponding retinal blurred image.

For the letter A and a pupil diameter of 3 mm, the results of convolution with the PSF under aberration conditions of 3.5D and 4.5D are shown in Figure 3.3. In the PSF visualization, blue (values close to 0) indicates a state with little to no blur, while higher values (approaching red) indicate increasingly severe blur.

The results of the same simulation, performed across all combinations of three letters (A, E, O), two aberration conditions (3.5D, 4.5D), and three pupil diameters (3 mm, 4 mm, 5 mm), are shown in Figure 3.4. A trend of increasing blur is observed as the pupil diameter and the degree of defocus (aberration) increase. These simulation results intuitively demonstrate the extent to

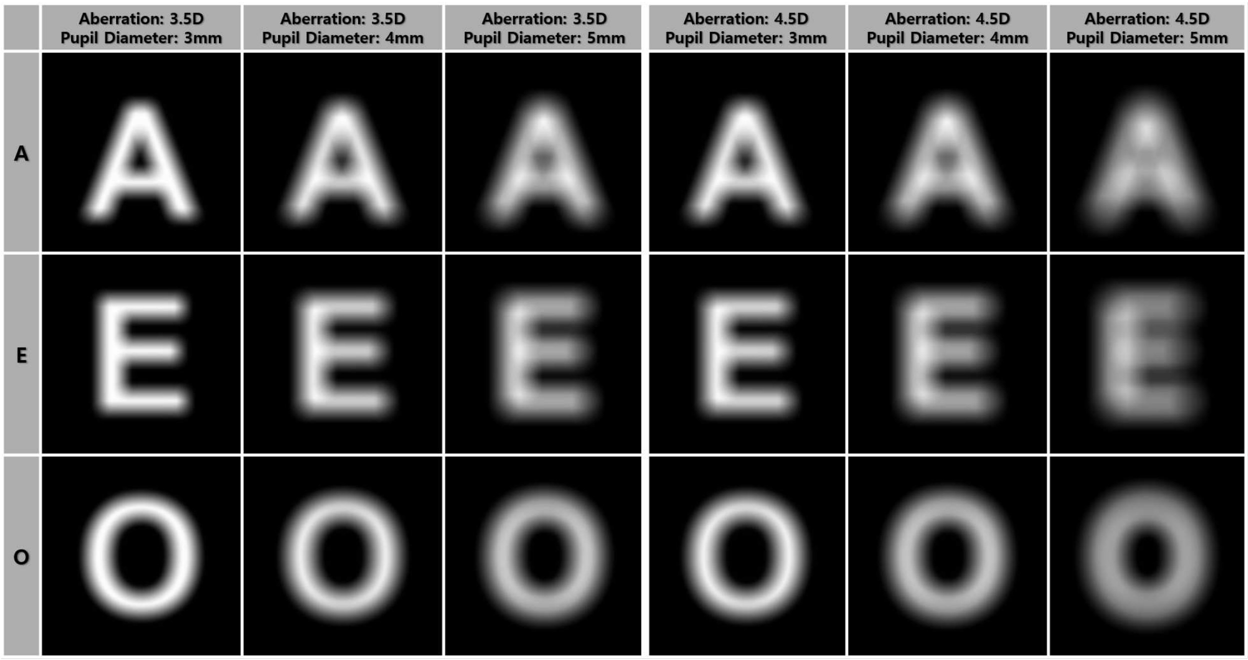


Figure 3.4: Simulated retinal blur for letters A, E, and O across aberration conditions (3.5D, 4.5D) and pupil diameters (3 mm, 4 mm, 5 mm), provided for illustrative purposes. Blur increases with both aberration magnitude and pupil diameter.

which an uncorrected original image is perceived as blurred under conditions of accommodation mismatch.

3.6 TV Optimization via FISTA

Wiener filtering and total variation (TV) methods have been the most widely used deconvolution techniques for correcting image focus issues. Alonso and Barreto first proposed an image-based vision correction technique, using Wiener filtering as the deconvolution algorithm, to pre-process images for low-vision users viewing computer screens [25]. Subsequently, Yellott and Yellott refined this method to facilitate the reading of text blurred by presbyopia, though they noted the inherent limitation of Wiener filtering-based approaches regarding contrast loss [26].

The concept of total variation was originally introduced by Rudin, Osher, and Fatemi for image denoising. They obtained an optimal solution for the denoised image by minimizing the total variation norm within a numerical algorithm [27]. Subsequently, various modifications of TV-based algorithms have been developed. Bioucas-Dias et al. proposed a TV-based image deconvolution algorithm for scenarios involving additive white Gaussian noise [28]. Additionally, Bioucas-Dias and Figueiredo introduced a two-step iterative thresholding algorithm that achieves a significantly faster convergence rate than existing TV algorithms [29]. Another TV-based approach is the primal-dual method, in which one optimization step is performed in the primal domain and another in the dual domain. According to Chambolle and Pock, TV-based techniques, formulated as convex optimization problems that account for sharp image discontinuities, offer the advantage of preserving and representing the intrinsic characteristics of edges [30].

To overcome these limitations, Montalto et al. proposed an image deconvolution method based on constrained total variation (TV) for users with low vision [2]. TV-based deconvolution is formulated as a convex optimization problem utilizing the L1-norm of the image gradient as a regularization term, thereby introducing a new parameter to manage the trade-off between ringing artifacts and contrast. Montalto et al. directly compared the Wiener filtering-based approach with the TV-based approach using both synthetic images and camera-captured images; they confirmed that, under a blur condition of -2.5D, the constrained TV-based system yielded superior visual acuity improvements compared to the Wiener filtering-based Gaussian PSF approach. Notably, in user experiments, 86% of participants were able to read text pre-processed using the TV-based method, whereas only 9% could read text pre-processed using the Wiener filtering-based method [2].

Based on these results, this study adopted a constrained TV-based deconvolution approach utilizing the monotone Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) proposed by Beck and Teboulle [31]. However, the pre-corrected images resulting from this focus-correction-based deconvolution algorithm differ in pixel values from the original images and sometimes contain negative values or extremely high-intensity values that are unsuitable for human visual perception [32].

The Contrast Loss (CL) adjustment and TV optimization procedure described in this section follow the method proposed by Arefin et al. [4, 33], which in turn adopted the contrast reduction technique of Montalto et al. [2]. Before generating the pre-corrected image, the contrast of the original image is intentionally reduced. CL is a parameter that intentionally narrows the pixel brightness range of the original image, calculated as:

$$tc = I \times (C_{high} - C_{low}) + C_{low}, \quad C_{high} = 1 - CL/200, \quad C_{low} = CL/200 \tag{3.11}$$

A higher CL value results in a greater reduction of the original image's contrast, which intensifies the ringing phenomenon in the uncorrected image. However, since excessive ringing can impair legibility, selecting an appropriate CL level is important. Details regarding the CL conditions are described in Section 4.6.1. Examples of contrast-adjustment results for the letter 'A' at CL levels of 70%, 80%, and 90% are shown in Figure 3.5.

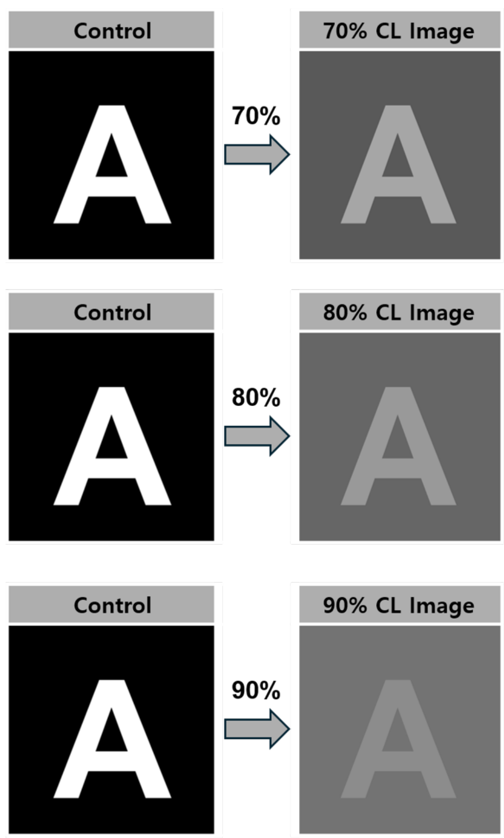


Figure 3.5: Contrast-adjusted images for the letter A at CL 70%, 80%, 90%. As CL increases, the contrast is further reduced, narrowing the brightness range of the original image.

The contrast-adjusted image tc serves as the input for TV optimization. TV optimization is a technique that controls noise and ringing by minimizing the L1-norm (sum of absolute gradient magnitudes), thereby generating a pre-corrected image u. This optimization is performed using the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [31]. The objective function for TV optimization is as follows:

$$p(\theta, I_c) = \arg\min_{0 \leq \|p\| \leq 1} \left(\|k * p - I_c\|_{L^2} + \theta \|\nabla p\|_{L^1} \right) \tag{3.12}$$

Here, p is the pre-corrected image, k is the blur kernel (OTF), I_c is the contrast-adjusted image (tc), and θ is the regularization parameter. The $\|k * p - I_c\|_{L^2}$ term (deconvolution term) ensures visual similarity between the corrected image and the original, while the $\theta\|\nabla p\|_{L^1}$ term (regularization term) controls ringing artifacts. In this process, the pixel values of p are constrained to the range [0, 1] (constrained TV optimization). θ is the TV regularization parameter and a key factor controlling the degree of ringing in the pre-corrected image; a high θ value reduces ringing and softens the image but decreases sharpness, whereas a low θ value increases ringing but produces a sharper image. FISTA is an algorithm that iteratively minimizes this TV-optimized objective function. The pseudocode for PrecorrectionImageGeneration is presented in Algorithm 3.

Algorithm 3 PrecorrectionImageGeneration

function PRECORRECTIONIMAGEGENERATION(tc, H, θ)
    Initialize $u \leftarrow tc$
    repeat
        $\nabla f(u) \leftarrow \mathcal{F}^{-1}\big(\bar{H} \cdot (H \cdot \mathcal{F}(u) - \mathcal{F}(tc))\big)$
        $u \leftarrow \mathrm{prox}_{\theta \cdot TV}(u - \alpha \nabla f(u))$ ▷ FISTA proximal step
        $u \leftarrow \Pi_{[0,1]}(u)$ ▷ Box constraint projection
    until convergence
    return u
end function

The pre-corrected image generated via Algorithm 3 has not yet undergone masking. At this stage, unwanted ringing artifacts appear around the characters. Examples of pre-corrected images before masking are shown in Figure 3.6.

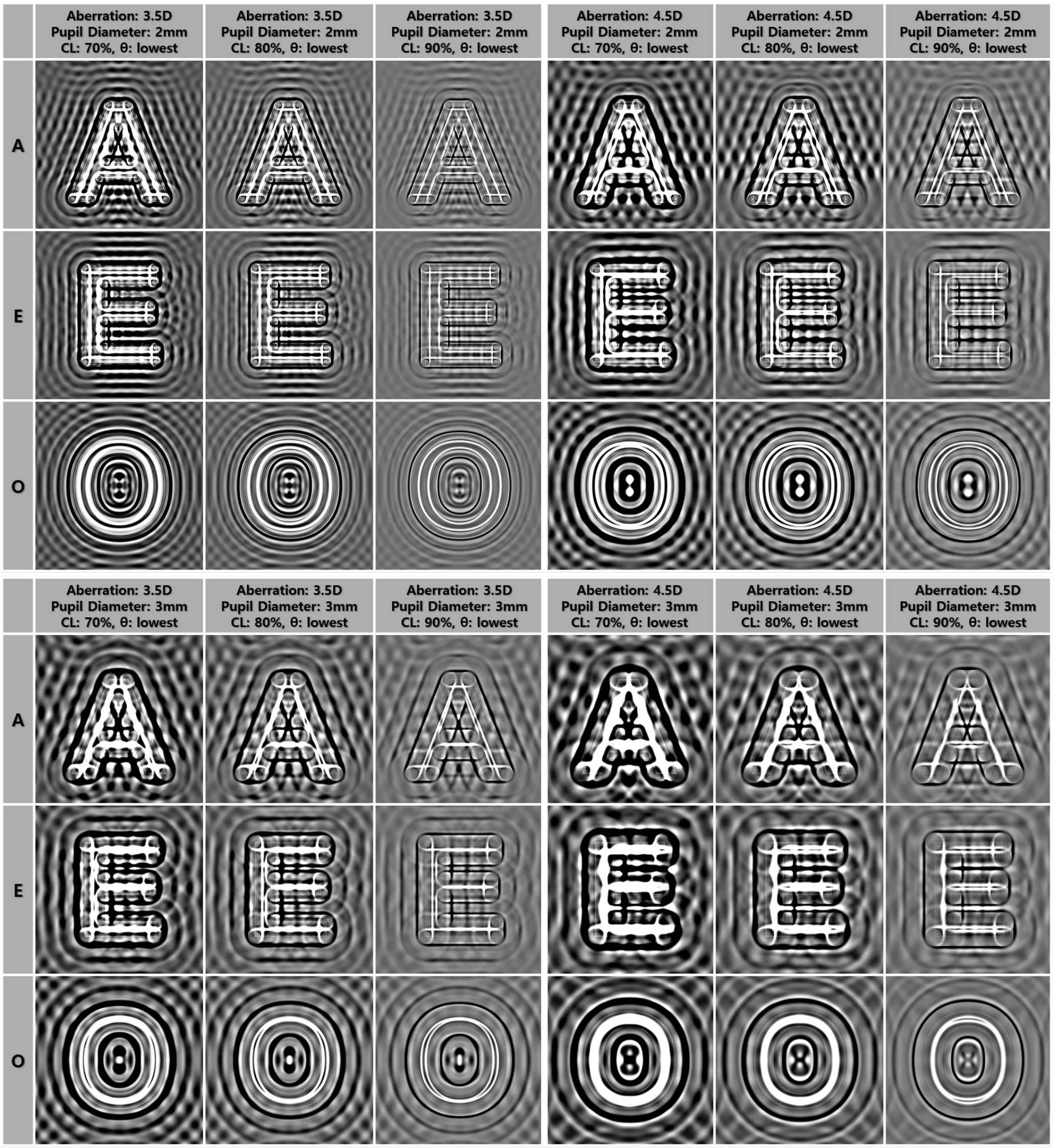


Figure 3.6: Pre-corrected images for letters A, E, and O before masking, across aberration conditions (3.5D, 4.5D) and pupil diameters (2 mm, 3 mm) at lowest θ (0.00001). Ringing artifacts are visible around the character edges.

3.7 Black and White Masking

In the existing method (Arefin et al. [4]), only black masking was applied to remove unwanted ringing artifacts appearing around the characters. The black mask is generated through the following process: first, the diameter of the Zernike PSF is estimated based on a threshold value (set to 0.01 of the PSF's maximum value in this study). This diameter is estimated by determining the bounding box of the region exceeding the threshold. Next, the mask diameter is determined by multiplying the estimated PSF diameter by an alpha value ($\alpha = 0.3$). Finally, a circular mask with this diameter is created and convolved with the original image to produce the final mask. This mask is applied to the pre-corrected image via pixel-wise multiplication to eliminate unwanted ringing artifacts appearing outside the characters.

However, applying the SharpView algorithm during the TV optimization process reduces contrast, causing the interior of the characters to darken. This creates a difference in brightness between the SharpView image and the Control image. Consequently, participants can distinguish between the two images based on brightness rather than legibility, potentially introducing bias into the experimental results. To address this issue, a white mask was introduced. This mask is generated by identifying the interior regions of the characters using a morphological erosion operation and setting the pixel values in those regions to 1 (white). By maintaining the brightness within the characters at a level comparable to that of the Control image in this way, bias resulting from brightness differences is eliminated. This masking process is illustrated in Figure 3.7.

Algorithm 4 Masking

function MASKING($u, I, focusDistance, defocusDistance, pupil\_diam, \alpha, threshold$)
    $diameterPSF \leftarrow$ ESTIMATEZERNIKEDIAMETER($focusDistance, defocusDistance, pupil\_diam, threshold$)
    $mask\_diameter \leftarrow \alpha \times diameterPSF$
    $mask \leftarrow$ CREATECIRCLESMASK($I, [N/2, M/2], mask\_diameter$)
    $mask_{original} \leftarrow$ CONV2($I, mask$) ▷ Convolve with original image
    $mask_{target} \leftarrow mask_{original} > 0$ ▷ Convert to binary mask
    $u_{without\_mask} \leftarrow u$
    $u \leftarrow u \times mask_{target}$ ▷ Black masking
    $letter\_boundary \leftarrow I > 0.5$
    $se \leftarrow$ STREL($disk, round(\alpha \times diameterPSF)$)
    $interior\_mask \leftarrow$ IMERODE($letter\_boundary, se$)
    $u[interior\_mask] \leftarrow 1$ ▷ White masking
    return $u, u_{without\_mask}$
end function

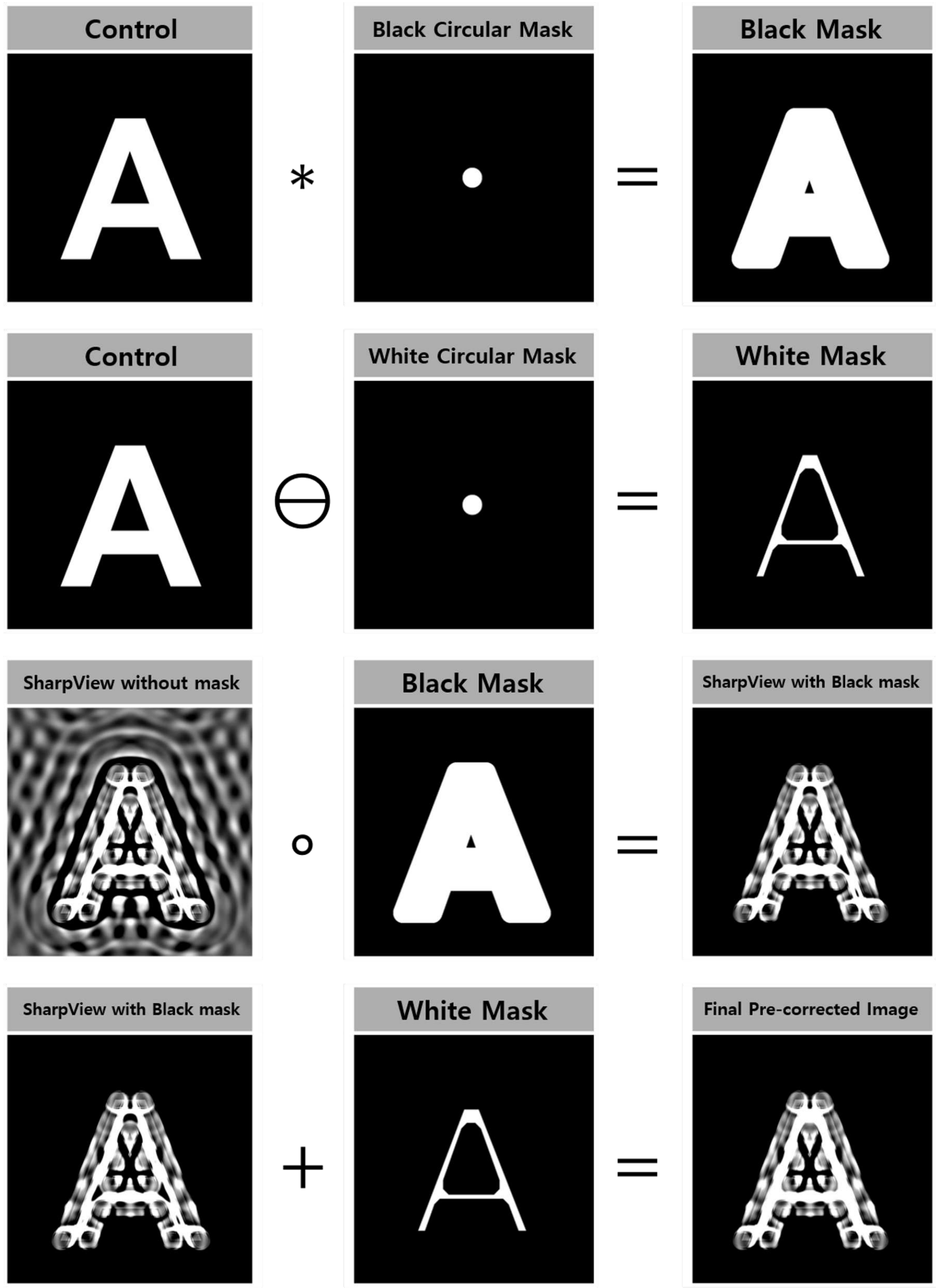


Figure 3.7: Process of generating the final pre-corrected image. (Row 1) The black mask is obtained by convolving the Control image I with a black circular mask ($*$). (Row 2) The white mask is obtained by applying morphological erosion ($\ominus$) to the Control image I with a white circular structuring element. (Row 3) The SharpView image without masking is element-wise multiplied ($\circ$) by the black mask to produce the SharpView image with black mask. (Row 4) The white mask is added to the SharpView image with black mask to produce the final pre-corrected image, where values exceeding 1 are clipped to 1.

3.8 Final Image Selection (CL × τ)

All calculated τ values for each combination of character (A, E, O) and CL (70%, 80%, 90%) are sorted in descending order. Here, τ is calculated based on the final image u that has undergone the masking process described in Section 3.7. Since a lower θ value increases the degree of correction applied during TV optimization (Section 3.6), it results in a higher τ value. Conversely, a higher θ value results in a lower τ value. For a given combination, the image with the highest τ value is designated as the "High τ image." To select the "Low τ image," the remaining τ values are examined sequentially in descending order, starting from the High τ image. The first image exhibiting an L2 norm difference of 15% or greater compared to the High τ image is selected as the Low τ image. The L2 norm difference is calculated as follows:

$$pct = \frac{\|I_{cand} - I_{ref}\|_F}{\|I_{ref}\|_F} \times 100 \tag{3.13}$$

Here, I_{cand} is the candidate image and I_{ref} is the reference image (High τ image). If pct is 15% or higher, the image is selected as the Low τ image.

Examples of SharpView images generated under conditions of 70%, 80%, 90% CL and High τ for pupil diameters of 2 mm and 3 mm and aberration conditions of 3.5D and 4.5D are shown in Figure 3.8 and Figure 3.9.

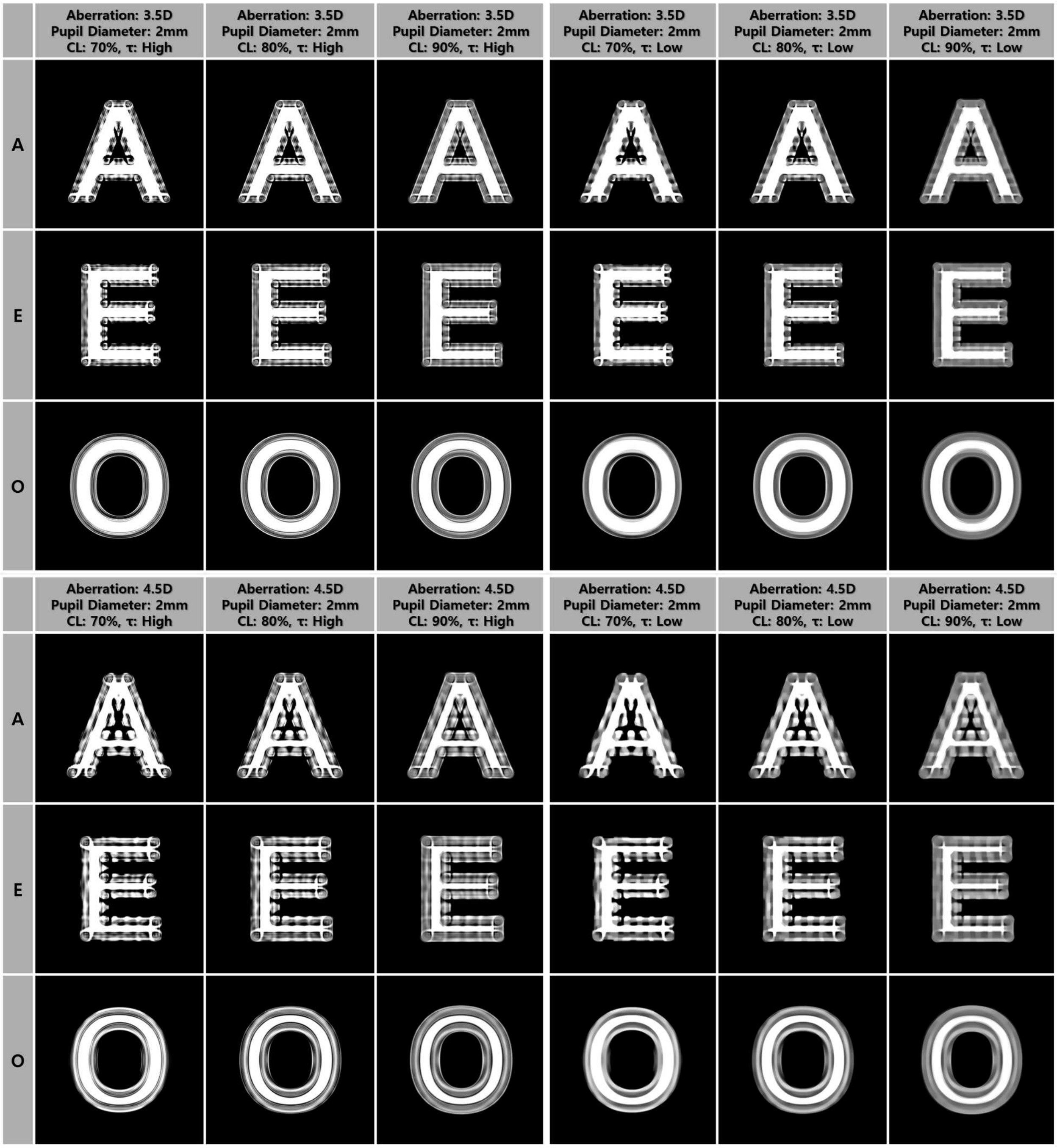


Figure 3.8: Final SharpView images for letters A, E, and O at CL 70%, 80%, 90%, High τ, and a pupil diameter of 2 mm, under the 3.5D and 4.5D aberration conditions.

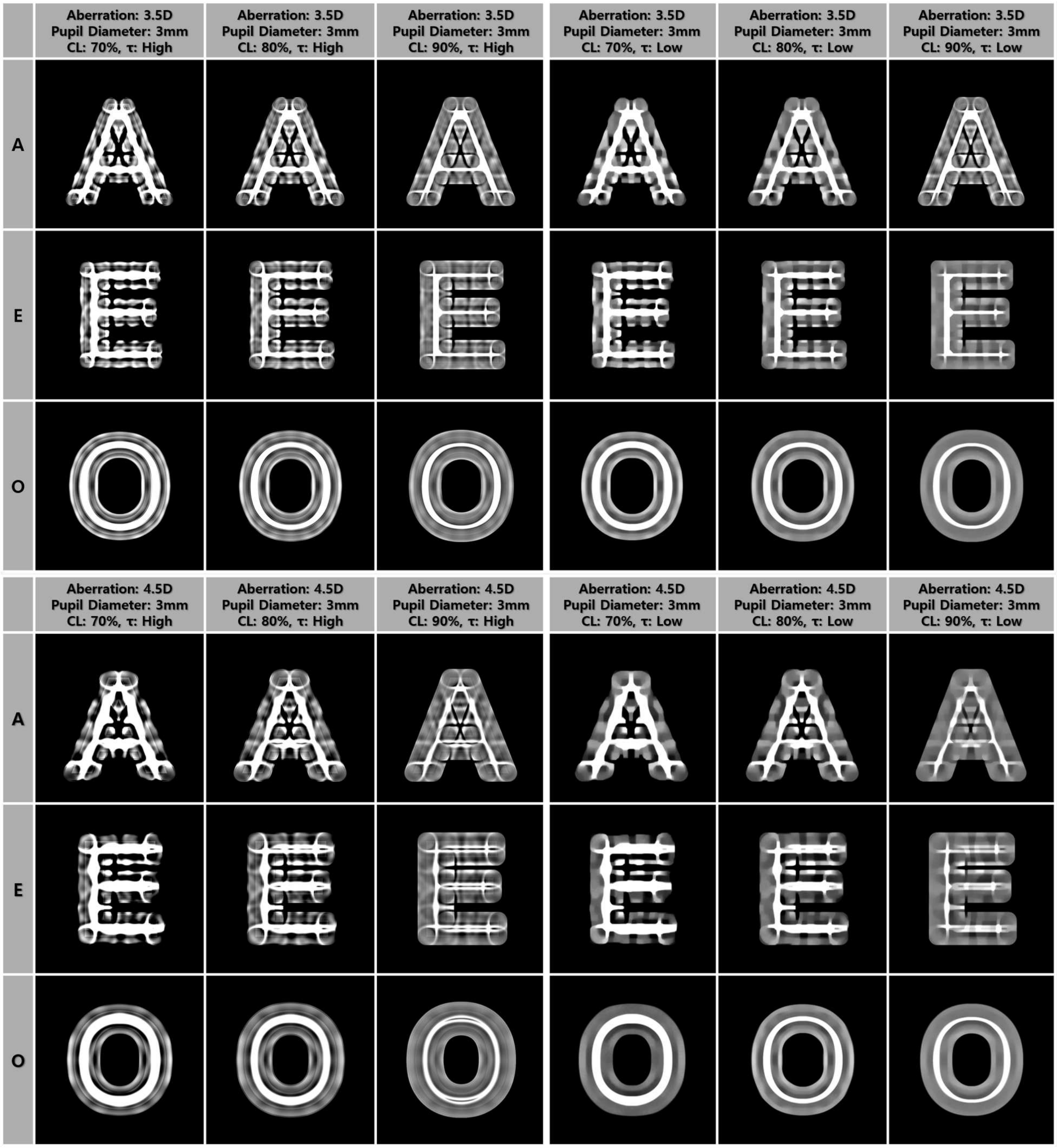


Figure 3.9: Final SharpView images for letters A, E, and O at CL 70%, 80%, 90%, High τ, and a pupil diameter of 3 mm, under the 3.5D and 4.5D aberration conditions.

Chapter 4

Experiment

4.1 Experimental Overview

Existing fonts designed to enhance legibility under conditions of defocus have primarily been validated through camera-based testing. This study evaluates the effectiveness of SharpView from the perspective of human legibility using the Microsoft HoloLens 2, a commercial OST AR device. The experiment involves two conditions: Incomplete Accommodation and Incorrect Accommodation. Participants perform a forced-choice task in which they compare SharpView-processed images with Control references and select the more readable image. This study modified the original SharpView algorithm and optimized its parameters for the experimental conditions. The following section details the methods, variables, and experimental setup employed to achieve these research objectives.

4.2 Hypothesis

H1: SharpView image will have a higher preference under Incomplete Accommodation and Incorrect Accommodation conditions.

H2: Since Arefin et al. [4] showed that the combination of Low θ (High τ; Section 3.8) and high CL produced the best SharpView effect, the combination of High τ and High CL is expected to show the greatest SharpView effect in this study.

H3: The SharpView font algorithm will be more effective for fonts with edges rather than for fonts with curves, since SharpView's correction mechanism generates ringing artifacts along letter edges [4].

4.3 Experimental Task and Procedure

In this experiment, a forced-choice task was employed to allow observers to evaluate the legibility of the SharpView font. This method involves selecting the more readable option between two presented images, offering the advantage of eliciting clear, quantifiable preference judgments. In each trial, two images were presented sequentially. One was displayed in the SharpView font, and the other served as the Control. Participants viewed both images in succession and selected which of the two was easier to read. A two-second interval, determined through pilot testing, was maintained between the images to prevent the afterimage of the first from influencing the judgment of the second, while the specific display duration was determined by the experimental conditions (detailed in Section 4.3.1 and Section 4.3.2).

Participants were not informed which images were SharpView and which were Control. This was to ensure that they could make judgments based purely on legibility, without any preconceived notions.

4.3.1 Experiment 1: Incomplete Accommodation

Before the experiment began, participants completed a consent form. Subsequently, their pupil diameters were measured using Pupil Labs Neon. Next, the built-in eye-tracking calibration for the HoloLens 2 was performed. Once calibration was complete, participants sat at a chin rest and put on the HoloLens 2. Participants observed a continuously presented X image in Arial Bold font via the HoloLens 2, while keeping their gaze fixed on a target directly in front of them. This allowed them to verify the image's size and position, experience the blur effect, and familiarize themselves with the experimental conditions.

The HoloLens 2's built-in eye-tracking system remains active throughout the experiment to verify that the participant is precisely focused on the fixation target. This verification is essential because, to induce incomplete accommodation, the virtual text must appear at the exact moment the participant's eyes are actually fixated on the target. Failure to detect this moment via eye-tracking could result in the image appearing while the participant is not properly fixating on the target. In such a case, the intended state of accommodation would not be guaranteed, potentially compromising the experimental manipulation itself.

The image was displayed for 150 ms. As noted in Section 1.4, full accommodation to a new distance requires approximately 360–380 ms [9, 10]. Since 150 ms is insufficient time for the participant to refocus at the new distance within this window, a state of Incomplete Accommodation was induced. A two-second interval was maintained between the two images to prevent afterimage effects. After viewing the two images sequentially, participants selected the image that appeared sharper by saying "one" or "two," or by indicating "1" or "2" with their fingers. The experimenter recorded the participants' responses using a Bluetooth keyboard. A one-minute rest period was provided after each set. In cases where Experiment 1 was conducted first, a 10-minute rest period was allowed after the experiment concluded. The experimental condition is illustrated in Figure 4.1.

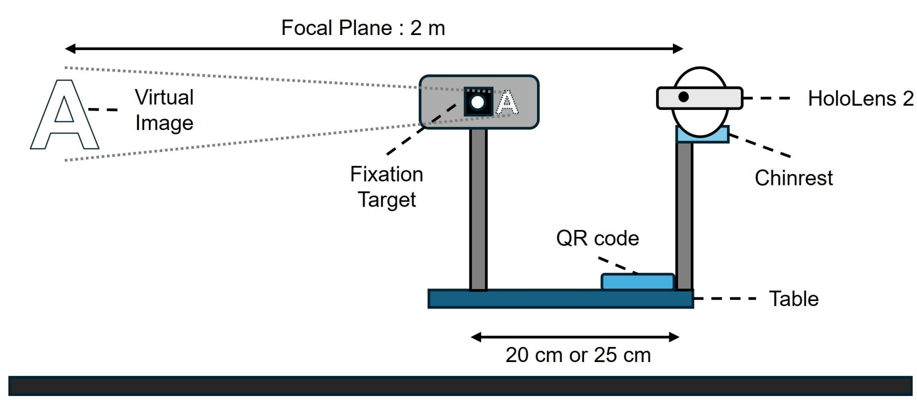


Figure 4.1: Experiment 1: Incomplete Accommodation, where the virtual image is rendered at the HoloLens 2 focal plane (2 m) while the participant fixates at 20 cm or 25 cm.

4.3.2 Experiment 2: Incorrect Accommodation

A calibration phase was also conducted in Experiment 2. While fixating on a point directly in front of them, participants viewed a continuously displayed X image in the Arial Bold font through the HoloLens 2. This allowed them to familiarize themselves with the experimental conditions while verifying the size and position of the image.

In this experiment, participants maintained a fixed distance while gazing at specific targets with their heads stabilized on a chin rest. Eye-tracking was not used in Experiment 2. Unlike Experiment 1, the manipulation in Experiment 2 (Incorrect Accommodation) does not depend on the precise moment the participant views the fixation target. Incorrect Accommodation is a state of continuous blur caused by the HoloLens 2's fixed focal plane (2 m), persisting throughout the entire duration (2 seconds) during which the participant views the image. Therefore, there is no need to verify momentary fixation as in Experiment 1. It is sufficient to simply maintain a stable viewing position at the designated distance (20 cm or 25 cm) using a chin rest.

Although the virtual character was rendered at a distance of 20 cm or 25 cm, the focal plane of the HoloLens 2 was fixed at 2 m. Consequently, even when a participant focused on the character, the actual image was generated at a distance of 2 m, resulting in blurriness. Each image was presented for 2 seconds, with a 2-second interval between images to prevent afterimage effects, a duration determined through pilot testing. After viewing the two images sequentially, participants selected the one that appeared sharper by saying "1" or "2" or by indicating the number with their fingers. The experimenter recorded the participants' responses using a Bluetooth keyboard. In Experiment 2, as well, a one-minute rest period was provided after the completion of each set. Feedback regarding the experiment was collected from the participants after all experimental procedures were concluded. The experimental condition is illustrated in Figure 4.2.

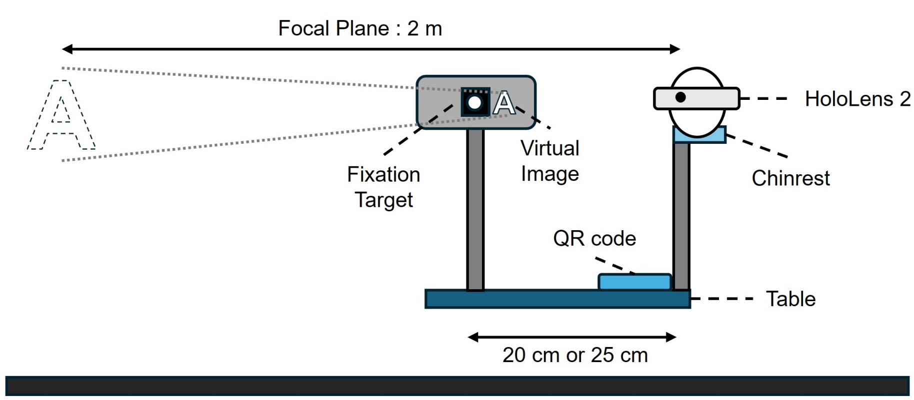


Figure 4.2: Experiment 2: Incorrect Accommodation, where the participant fixates at 20 cm or 25 cm, but the virtual image is rendered at the HoloLens 2 focal plane (2 m).

Two pilot experiments were conducted before this experiment. The first pilot experiment with 10 participants was conducted without white masking. Participants showed a strong tendency to select the Control regardless of the SharpView effect because of the low brightness of the SharpView images. White masking was added to compensate for this difference in brightness. A second pilot experiment was conducted with four participants under the condition of incorporating this masking. Based on the second pilot study, participants consistently selected the Control when the SharpView effect was absent and even when the effects were comparable.

4.4 Stimuli

Two types of visual stimuli were used in this experiment: a fixation target to lock the participants' gaze, and a comparison image pair, which itself consisted of two versions, SharpView and Control. The fixation target consisted of a white dot centered within a black square printed on paper, a design chosen to maximize fixation stability [34].

The comparison image pair consisted of an image generated using the SharpView algorithm and the Control image. Since the Control condition consists of the original image without any corrections, it serves as a baseline for evaluating the effects of SharpView's processing. In other words, this condition is rather a reference point representing how the image appears to participants without any correction. By comparing how SharpView is perceived relative to this baseline, whether better, similar, or worse, one can isolate and evaluate the effectiveness of the correction algorithm itself.

Both stimuli utilized the three letters A, E, and O. These letters were rendered in the Arial Bold font and generated as 1000×1000 pixel images featuring white text on a black background. Since the AR system was rendered black as transparent, only the white letters were actually displayed. The letter sizes were set to 1 cm at a distance of 20 cm and 1.25 cm at a distance of 25 cm, ensuring a constant visual angle of approximately 2.86° across both conditions.

The SharpView images were individualized according to each participant's pupil size. For each participant, 18 images (calculated as letters (3) × CL (3) × τ (2)) were used per aberration condition. Consequently, a total of 36 SharpView images were used across the two aberration conditions (3.5D and 4.5D). Details regarding the selection of τ and characters are provided in Section 3.8 and Section 4.6.1. Examples of the stimuli presented to participants are shown in Figure 4.3.

Figure 4.3: Stimuli presented to participants, consisting of a fixation target (left) and a SharpView-processed virtual image (right).

4.5 Apparatus and Setup

4.5.1 Devices (Microsoft HoloLens 2, Pupil Labs Neon)

This experiment was conducted using the Microsoft HoloLens 2, a commercial OST AR device. The device specifications include a fixed focal plane of 2 m, a diagonal field of view (FOV) of 52°, a resolution of 1440×936, a refresh rate of 60 Hz, and a weight of 566 g [6]. To prevent fatigue caused by excessive brightness and to avoid missing text due to insufficient brightness, the device's brightness was set to a medium level (5/11). The built-in eye-tracking capability of HoloLens 2 was used to determine whether participants were fixating in Experiment 1 (Section 4.3.1).

Each participant's pupil diameter was measured using Pupil Labs Neon. A total of 1,000 frames were collected using code written in Python 3.12, and values below 2 mm were excluded as errors. The average of the measurements from the left and right eyes was calculated in 0.1 mm increments and used as a parameter to generate each participant's SharpView image. The apparatus used in this study is shown in Figure 4.4.

Figure 4.4: Apparatus used in this study: Microsoft HoloLens 2 (left) and Pupil Labs Neon (right).

4.5.2 Experimental Setup

A chinrest was used to stabilize the participants' heads and maintain a constant distance from the HoloLens 2, ensuring that all participants took part in the experiment under identical conditions. The virtual image was positioned immediately to the right of the fixation target, and its size was set to subtend the same visual angle as the black square of the fixation target. To render images at the precise distance, a QR code was placed on the desk, and its top-left corner was designated as the coordinate origin. The experimental setup is shown in Figure 4.5.

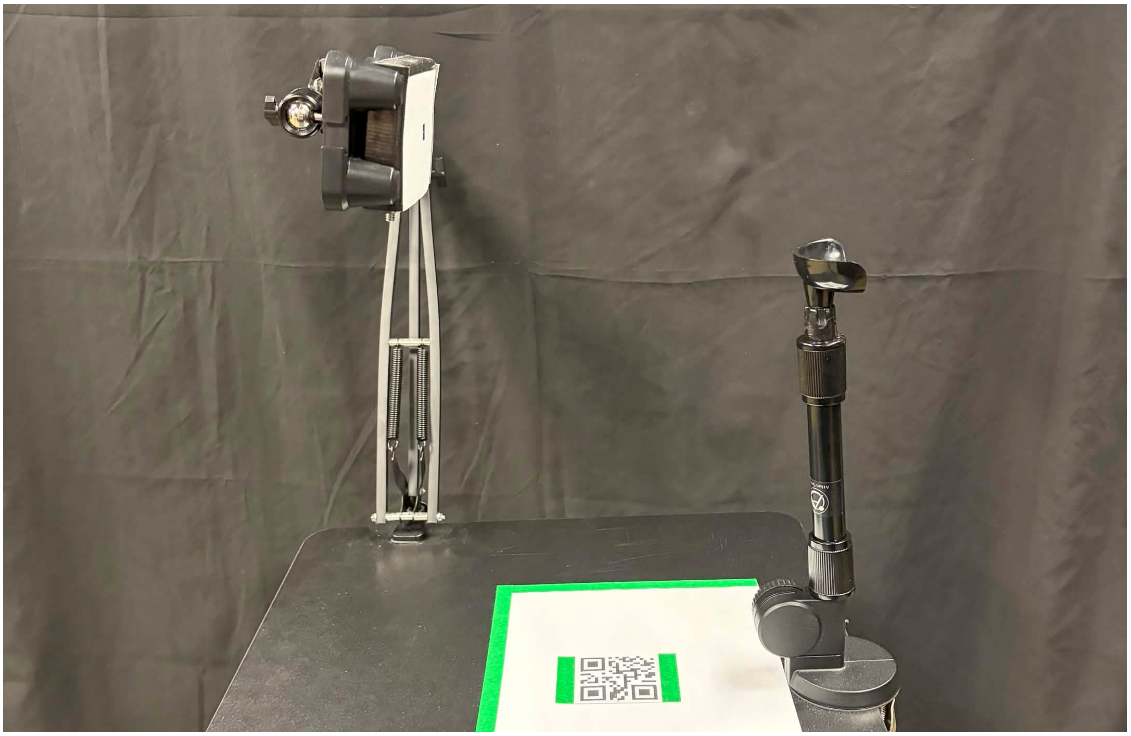


Figure 4.5: Experimental setup consisting of a chinrest, fixation target stand, and QR code on the table.

4.5.3 Software

This experimental app was developed using Unity 2022.3.22f1. To conduct the experiment, a hotspot was enabled on a computer, and the HoloLens 2 was configured to connect to it via Wi-Fi. In this setup, Unity's "Build and Run" function was used to remotely install and launch the app on the HoloLens 2. Additionally, a Bluetooth keyboard was connected to the HoloLens 2, allowing the experimenter to control the experiment and record participant responses. The collected data was saved as a CSV file in the HoloLens 2's local storage.

4.6 Variables

4.6.1 Independent Variables

These five independent variables (font type, CL, τ, aberration, and letter) were common to both Experiment 1 and Experiment 2.

Table 4.1: Independent variables used in this experiment.

Independent Variables	Levels
Font Type	SharpView, Control
CL (Contrast Loss)	70%, 80%, 90%
τ (Tau)	High, Low
Aberration	3.5D, 4.5D
Letter	A, E, O

Font Type (SharpView, Control)

The font type conditions consist of two categories: SharpView, a pre-corrected image developed in this study, and Control, an undistorted original Arial Bold font image. Participants selected the image that appeared sharper between the two.

CL (70%, 80%, 90%)

CL stands for Contrast Loss. It is a parameter that intentionally reduces contrast by narrowing the pixel brightness range of the original image to [C_low, C_high]. It is calculated as CL = (C_low + (1 - C_high)) × 100. A higher CL value results in a greater reduction of the original image's contrast, thereby increasing the ringing effect in the uncorrected image and making it appear sharper, even when out of focus. This study employed three conditions: 70%, 80%, 90%. The minimum value was set at 70% based on pilot data indicating that participants did not select SharpView images under conditions below 70%, making it impossible to obtain meaningful results.

τ (High, Low)

In the study by Montalto et al. [2], τ was defined based on the contrast-corrected image tc to manage the trade-off between mitigating the ringing phenomenon and enhancing contrast. In the present study, this method was modified to calculate τ based on the original image I, thereby using it as a measure of correction intensity that reflects the overall degree of image alteration relative to the original, independent of CL conditions. τ is calculated as $\tau = |TV(u) - TV(I)|/TV(I)$. A higher τ value indicates a greater degree of alteration in the image compared to the original. To evaluate the significance of the τ effect, two conditions, High τ and Low τ, were established in this study. As described in Section 3.8, a lower θ value corresponds to a higher τ value.

Aberration (3.5D, 4.5D)

Aberration is a variable representing the degree of defocus. A value of 3.5D corresponds to a distance of 25 cm relative to a 2 m focal distance, while 4.5D corresponds to 20 cm. Compared to the 2 m focal plane, both distances provide sufficient blur while remaining within the range where participants can maintain focus and the HoloLens 2 can render images. These specific distances were selected because placing images too close can quickly cause visual fatigue for participants.

Letter (A, E, O)

In this experiment, the three letters A, E, and O were used as independent variables. Among the most frequently used English letters, namely E, T, A, and O, the letter T was excluded because, like E, it consists solely of straight lines. Consequently, A (composed of diagonal lines), E (composed of straight lines), and O (composed of curved lines) were selected to represent a variety of letter shapes.

4.6.2 Dependent Variables

Table 4.2: Dependent variable used in this experiment.

Dependent Variable	Levels
Legibility	Selection of the image that is easier to read

The dependent variable in this experiment, legibility, is operationalized as the image selected by the participant in each trial and was measured in both Experiment 1 and Experiment 2.

Participants chose which of the two presented images was more visible (the first or the second), and based on these responses and the image presentation order (First_Image), it was determined whether the SharpView image was selected. Detailed data processing methods are described in Section 5.1.

4.7 Experimental Design

Presenting image patterns in a fixed order could lead to participant bias. To prevent this, each experiment consisted of four sets, with the 3.5D and 4.5D aberration conditions each presented twice. Within each set, combinations of CL, τ, and characters were presented in a randomized order. If the SharpView image was presented first in the first set, the Control image was presented first in the second set. The order of aberration conditions was also counterbalanced. Specifically, participants who viewed the 3.5D condition first proceeded in the order of 3.5D $\rightarrow$ 4.5D $\rightarrow$ 3.5D $\rightarrow$ 4.5D, whereas those who viewed the 4.5D condition first proceeded in the order of 4.5D $\rightarrow$ 3.5D $\rightarrow$ 4.5D $\rightarrow$ 3.5D. Furthermore, the experimental sequence was counterbalanced such that if one participant performed Experiment 1 first, the next participant performed Experiment 2 first.

This experiment employed a within-subject design in which all participants experienced every condition. Each experiment consisted of a total of 72 trials, comprising factors of CL (3) × τ (2) × aberration (2) × Letter (3) × repetition (2), with each condition repeated twice to enhance the reliability of the results. Combining Experiments 1 and 2, each participant performed a total of 144 trials. Within each set, the image presentation order was randomized, while the order of aberration conditions and experiments was counterbalanced using a Latin square design. Combining Experiments 1 and 2, participants completed the entire study in a single session, with a total duration of approximately 75 minutes.

4.8 Participants

An a priori power analysis was conducted using G*Power to determine the sample size. For a within-subjects design involving a single group and 36 repeated measurements, the analysis, applying a statistical power of 0.80, a small-to-medium effect size (f = 0.20), a correlation coefficient of 0.5 between repeated measurements, and a non-sphericity correction factor of 0.5, indicated that a minimum of 15 participants was required.

Sixteen participants (10 men, 6 women) were recruited via email and personal contacts. The mean age of the participants was 24.6 years (range: 20–32 years). All participants had normal or corrected-to-normal vision. Individuals who wore glasses or contact lenses were eligible to participate. All participants signed an informed consent form prior to the experiment and received a digital gift card worth $20 upon completion.

Chapter 5

Results and Discussion

5.1 Data Pre-Processing

For each trial, the selection of SharpView was determined based on the participant's response (Legibility) and the image presentation order (First_Image). When First_Image was 1, the first image presented was SharpView, and a Legibility score of 1 was recorded as a SharpView selection (1), while a score of 2 was recorded as a Control selection (0). When First_Image was 0, the first image presented was Control, and a Legibility score of 2 was recorded as a SharpView selection (1), while a score of 1 was recorded as a Control selection (0).

The SharpView preference rate (percnt_SV_pref) was calculated for each participant based on the experimental conditions of CL, τ, Aberration, and Letter. This value represents the number of times a participant selected SharpView under a specific condition divided by the total number of trials, resulting in a value between 0 and 1.

Since eye-tracking calibration was completed by all participants before Experiment 1, no trials were excluded due to failed gaze detection in either experiment. Therefore, all data were included in the analysis.

5.2 Analysis

In this study, three statistical analyses were conducted to analyze the effectiveness of SharpView.

Firstly, a binomial test was used to assess whether the SharpView selection rate deviated significantly from the chance level of 50%. If participants were to choose randomly, both SharpView and Control would be selected with a probability of 50% in a forced-choice task. In addition to the overall analysis, binomial tests were performed for each parameter combination (CL, τ, Letter) to identify the specific conditions under which SharpView was significantly preferred. The binomial test was selected because the forced-choice task inherently yields a dichotomous outcome (choosing one of two options) and provides a clear baseline (50%) for random guessing. This method is more directly suited to the situation than the t-test, which assumes a normal distribution and continuous data.

Secondly, a repeated measures ANOVA was used to examine whether CL, τ, Letter, and Aberration influenced the selection of SharpView. Since all participants experienced every condition, CL, τ, Letter, and Aberration were defined as within-subject variables. Effect sizes were reported as generalized eta squared (η^2), with small ($\eta^2 = 0.01$), medium (η^2 = 0.06), and large (η^2 = 0.14) as threshold. All analyses used $p < 0.05$ as the threshold for statistical significance. Repeated-measures ANOVA was used because the study employed a within-subject design in which every participant experienced all combinations of CL, τ, Letter, and Aberration. Consequently, observations across conditions are not independent, and repeated-measures ANOVA explicitly accounts for this within-participant correlation.

Third, for variables showing significant effects in the ANOVA, a Bonferroni post hoc test was conducted to examine the specific differences between conditions. Together with the condition-specific binomial tests, these analyses were used to determine the parameter combinations under which SharpView was most effective. Bonferroni correction was applied due to the large number of multiple pairwise comparisons involving combinations of CL, τ, and Letter. Without correction, the family-wise error rate increases as the number of comparisons grows. All analyses were performed using the ez and DescTools packages in R.

5.3 H1: SharpView Preference under Accommodation Mismatch

5.3.1 3.5D Aberration Condition

In the 3.5D condition of Experiment 1, SharpView was selected in only 180 out of 576 trials (31.3%). This indicates that, in the majority of cases, participants judged the Control condition to be easier to read than SharpView, a statistically significant difference (binomial test, $p < .001$). In the 3.5D condition of Experiment 2, SharpView was selected even less frequently, in 129 trials (22.4%), demonstrating an even more pronounced preference for the Control condition ($p < .001$).

However, since this overall selection rate aggregates data across all parameter combinations, different patterns may emerge under specific conditions. In fact, focusing specifically on letters A and E, SharpView's selection rate actually exceeded 50% in conditions where the CL was relatively low (70% or 80%). In Experiment 1, SharpView was selected more frequently than the Control for letter A (59.4%) and E (62.5%) at 70% CL, and for letter A (53.1%) and E (56.2%) at 80% CL. A similar trend was observed in Experiment 2 for letter E (59.4%) at 70% CL and letter A (56.2%) at 80% CL. However, these differences did not reach statistical significance after Bonferroni correction. While these results do not allow us to definitively conclude that SharpView is clearly superior to the Control in these specific combinations, they suggest that it does not underperform relative to the Control. The results of the binomial tests for each combination are presented in Table 5.1 (Experiment 1) and Table 5.2 (Experiment 2), and this pattern is illustrated in Figure 5.1.

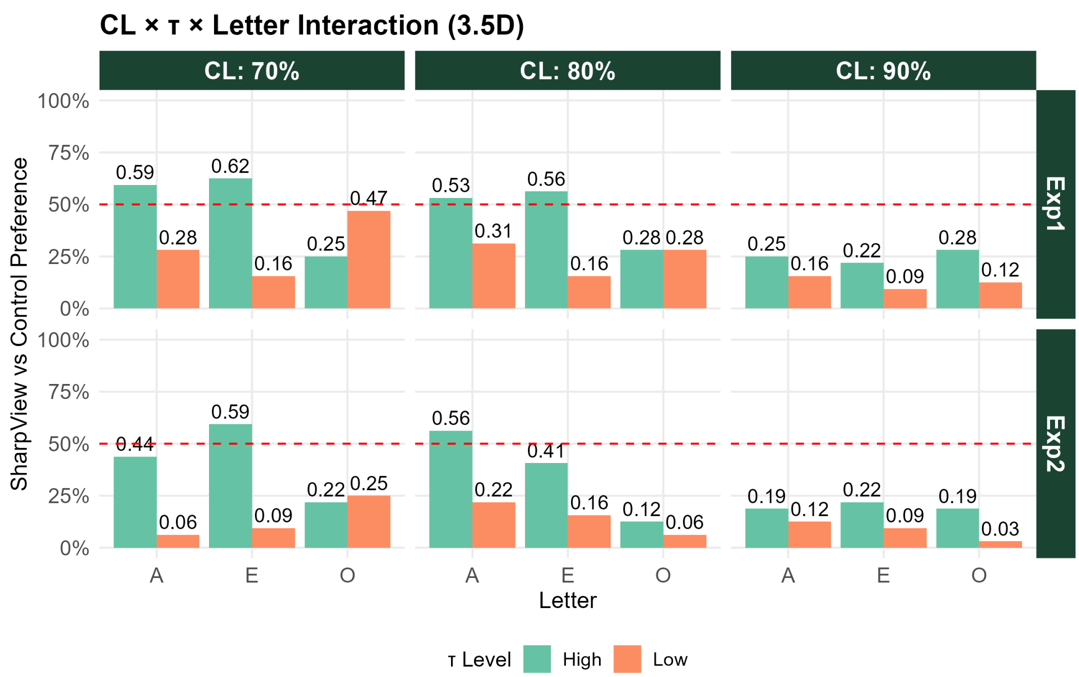


Figure 5.1: SharpView preference rate by CL, τ, and Letter under the 3.5D aberration condition, shown separately for Experiment 1 and Experiment 2. The dashed line indicates the chance level (50%).

Table 5.1: Binomial test results for each CL x τ x Letter combination under the 3.5D aberration condition, Experiment 1. Asterisks indicate significant preference for Control ($p < .05$).

CL	τ	Letter	n_{SV}	n_{total}	Pref. Rate	p
70%	High	A	19	32	59.4%	.377
70%	High	E	20	32	62.5%	.215
70%	High	O	8	32	25.0%	.007*
70%	Low	A	9	32	28.1%	.020*
70%	Low	E	5	32	15.6%	<.001*
70%	Low	O	15	32	46.9%	.860
80%	High	A	17	32	53.1%	.860
80%	High	E	18	32	56.2%	.597
80%	High	O	9	32	28.1%	.020*
80%	Low	A	10	32	31.2%	.050
80%	Low	E	5	32	15.6%	<.001*
80%	Low	O	9	32	28.1%	.020*
90%	High	A	8	32	25.0%	.007*
90%	High	E	7	32	21.9%	.002*
90%	High	O	9	32	28.1%	.020*
90%	Low	A	5	32	15.6%	<.001*
90%	Low	E	3	32	9.4%	<.001*
90%	Low	O	4	32	12.5%	<.001*

Table 5.2: Binomial test results for each CL x τ x Letter combination under the 3.5D aberration condition, Experiment 2. Asterisks indicate significant preference for Control ($p < .05$).

CL	τ	Letter	n_{SV}	n_{total}	Pref. Rate	p
70%	High	A	14	32	43.8%	.597
70%	High	E	19	32	59.4%	.377
70%	High	O	7	32	21.9%	.002*
70%	Low	A	2	32	6.2%	<.001*
70%	Low	E	3	32	9.4%	<.001*
70%	Low	O	8	32	25.0%	.007*
80%	High	A	18	32	56.2%	.597
80%	High	E	13	32	40.6%	.377
80%	High	O	4	32	12.5%	<.001*
80%	Low	A	7	32	21.9%	.002*
80%	Low	E	5	32	15.6%	<.001*
80%	Low	O	2	32	6.2%	<.001*
90%	High	A	6	32	18.8%	<.001*
90%	High	E	7	32	21.9%	.002*
90%	High	O	6	32	18.8%	<.001*
90%	Low	A	4	32	12.5%	<.001*
90%	Low	E	3	32	9.4%	<.001*
90%	Low	O	1	32	3.1%	<.001*

5.3.2 4.5D Aberration Condition

In the 4.5D condition of Experiment 1, SharpView was selected in only 190 out of 576 trials (33.0%), a pattern similar to the 3.5D condition in which the Control was preferred ($p < .001$). In the 4.5D condition of Experiment 2, it was selected 131 times (22.7%), again indicating a strong preference for the Control ($p < .001$).

As with the 3.5D condition, SharpView's selection rate exceeded 50% for certain letter-CL combinations in the 4.5D condition as well. Specifically, SharpView was selected more frequently than the Control for letter A at 70% CL in Experiment 1 (56.2%) and for letter A at 70% CL in Experiment 2 (62.5%). For letter E at 80% CL in Experiment 1, the rate was 50.0%, exactly matching the chance level. These results suggest that while SharpView cannot be definitively said to outperform the Control in these combinations, it does not perform worse. The results of the binomial tests for each combination are presented in Table 5.3 (Experiment 1) and Table 5.4 (Experiment 2), and this pattern is illustrated in Figure 5.2.

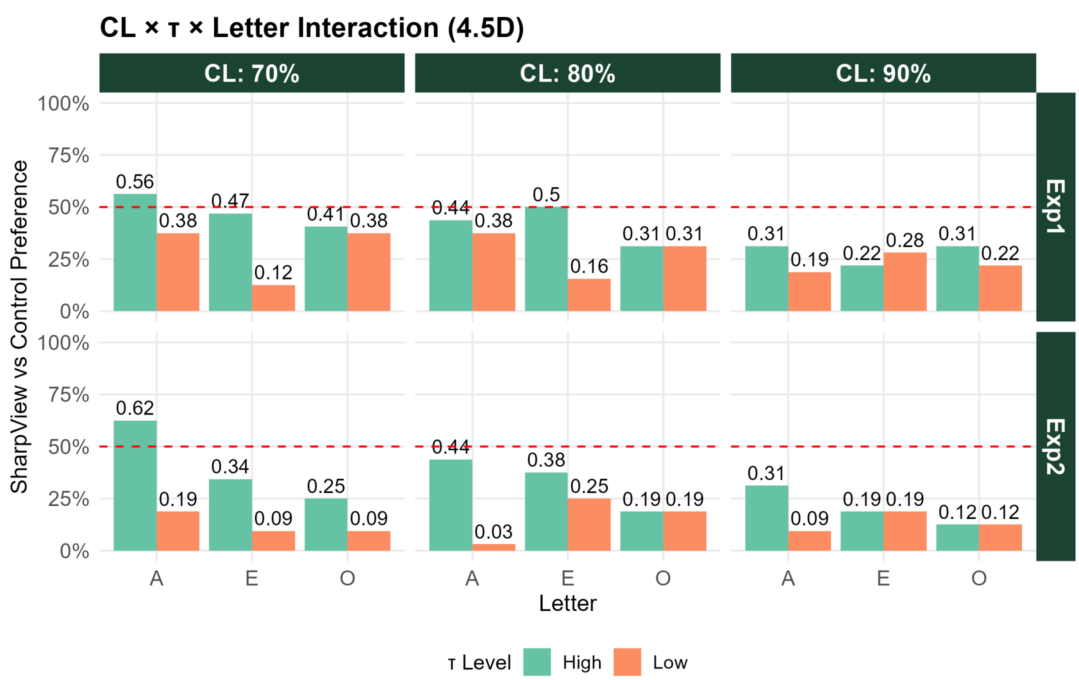


Figure 5.2: SharpView preference rate by CL, τ, and Letter under the 4.5D aberration condition, shown separately for Experiment 1 and Experiment 2. The dashed line indicates chance level (50%).

Table 5.3: Binomial test results for each CL x τ x Letter combination under the 4.5D aberration condition, Experiment 1. Asterisks indicate significant preference for Control ($p < .05$).

CL	τ	Letter	n_{SV}	n_{total}	Pref. Rate	p
70%	High	A	18	32	56.2%	.597
70%	High	E	15	32	46.9%	.860
70%	High	O	13	32	40.6%	.377
70%	Low	A	12	32	37.5%	.215
70%	Low	E	4	32	12.5%	<.001*
70%	Low	O	12	32	37.5%	.215
80%	High	A	14	32	43.8%	.597
80%	High	E	16	32	50.0%	1.000
80%	High	O	10	32	31.2%	.050
80%	Low	A	12	32	37.5%	.215
80%	Low	E	5	32	15.6%	<.001*
80%	Low	O	10	32	31.2%	.050
90%	High	A	10	32	31.2%	.050
90%	High	E	7	32	21.9%	.002*
90%	High	O	10	32	31.2%	.050
90%	Low	A	6	32	18.8%	<.001*
90%	Low	E	9	32	28.1%	.020*
90%	Low	O	7	32	21.9%	.002*

Table 5.4: Binomial test results for each CL x τ x Letter combination under the 4.5D aberration condition, Experiment 2. Asterisks indicate significant preference for Control ($p < .05$).

CL	τ	Letter	n_{SV}	n_{total}	Pref. Rate	p
70%	High	A	20	32	62.5%	.215
70%	High	E	11	32	34.4%	.110
70%	High	O	8	32	25.0%	.007*
70%	Low	A	6	32	18.8%	<.001*
70%	Low	E	3	32	9.4%	<.001*
70%	Low	O	3	32	9.4%	<.001*
80%	High	A	14	32	43.8%	.597
80%	High	E	12	32	37.5%	.215
80%	High	O	6	32	18.8%	<.001*
80%	Low	A	1	32	3.1%	<.001*
80%	Low	E	8	32	25.0%	.007*
80%	Low	O	6	32	18.8%	<.001*
90%	High	A	10	32	31.2%	.050
90%	High	E	6	32	18.8%	<.001*
90%	High	O	4	32	12.5%	<.001*
90%	Low	A	3	32	9.4%	<.001*
90%	Low	E	6	32	18.8%	<.001*
90%	Low	O	4	32	12.5%	<.001*

5.3.3 Overall

Under both 3.5D and 4.5D conditions, participants generally tended to prefer the Control over SharpView. A repeated-measures ANOVA revealed that the main effect of aberration was not significant in either experiment (Experiment 1: $F_{1,15} = 0.415, p = .529$; Experiment 2: $F_{1,15} = 0.02, p = .883$). In other words, there was no difference in the overall preference for SharpView between the two aberration conditions.

This general preference for the Control condition should not be interpreted simply as a failure of SharpView. This tendency was already observed in the pilot study, where participants consistently chose the more familiar Control even in conditions where the SharpView effect was virtually absent (Section 4.3.2). This pattern aligns with the mere-exposure effect, in which initial wariness toward unfamiliar stimuli gradually shifts to liking through repeated exposure [35]. Specifically within the context of typography, Beier and Larson [36] found that readers exhibited a distinct aversion and low preference for unfamiliar typefaces, even though those typefaces did not significantly affect actual reading speed after sufficient exposure. In other words, the negative reaction to unfamiliar letter forms was a matter of subjective preference rather than legibility. This suggests that in the forced-choice task of the current experiment, the participants' choices likely reflected their familiarity with the fonts rather than actual differences in legibility.

Given this context, the fact that the SharpView selection rate exceeded 50% in certain CL-letter combinations, as observed in Section 5.3.1 and Section 5.3.2, takes on greater significance. Despite a general bias among participants to prefer the Control condition due to familiarity, SharpView was selected more frequently in these specific combinations, effectively overcoming that bias. Notably, such combinations appeared more frequently in Experiment 1 (four cases at 3.5D and one at 4.5D) than in Experiment 2 (two cases at 3.5D and one at 4.5D). This discrepancy may stem from the differing mechanisms underlying the blur observed in Incomplete Accommodation (Experiment 1) versus Incorrect Accommodation (Experiment 2). The blur in Experiment 1 is a transient phenomenon occurring while the eye is in the process of focusing on a new distance, whereas the blur in Experiment 2 persists even after the eye has fully focused on the target distance, due to the display's fixed focal plane. These differences imply that the nature of the optical distortion actually experienced by the eye varies between the two conditions, which may explain why the corrective effect of SharpView manifested differently across the experiments.

Overall, H1 was generally not supported, as participants preferred Control over SharpView. However, as observed in Section 5.3.1 and Section 5.3.2, the SharpView selection rate exceeded the chance level for certain combinations of CL, τ, and letters. Although this difference did not reach statistical significance, it suggests potential for improvement through future parameter optimization.

5.4 H2: Effect of CL and τ on SharpView Preference

5.4.1 3.5D Aberration Condition

In the 3.5D condition of Experiment 1, repeated-measures ANOVA results showed that both CL ($F_{2,30} = 13.85, p < .001$) and τ ($F_{1,15} = 17.01, p < .001$) had a significant effect on the SharpView selection rate. The same pattern was observed in Experiment 2 (CL: $F_{2,30} = 7.33, p = .003$; τ: $F_{1,15} = 22.13, p < .001$).

Regarding the effect of τ, the selection rate for the High τ condition was consistently higher than that for the Low τ condition across all CL levels. In Experiment 1, the rates for High τ were 49.0% at CL 70%, 45.8% at CL 80%, and 25.0% at CL 90%, whereas the rates for Low τ were only 30.2%, 25.0%, and 12.5%, respectively. The same trend was observed in Experiment 2 (High τ: 41.7%, 36.5%, 19.8%; Low τ: 13.5%, 14.6%, 8.3%). As shown in Figure 5.3, the bars representing High τ (green) are consistently higher than those representing Low τ (orange) at all CL levels.

Regarding the effect of CL, the SharpView selection rate was higher when CL was lower (70% or 80%) and dropped markedly when CL increased to 90%. This trend was consistent across both High τ and Low τ conditions. Bonferroni post-hoc tests revealed no significant difference between CL levels of 70% and 80% (Experiment 1: $p = 1.000$; Experiment 2: $p = 1.000$), whereas the 90% CL condition differed significantly from both the 70% and 80% conditions (Experiment 1: 70 vs.

90, $p < .001$; 80 vs. 90, $p = .003$; Experiment 2: 70 vs. 90, $p = .014$; 80 vs. 90, $p = .049$). Figure 5.3 also clearly illustrates a pattern in which the selection rate decreases in stages as CL increases across both experiments. The Bonferroni test results for each CL combination are presented in Table 5.5.

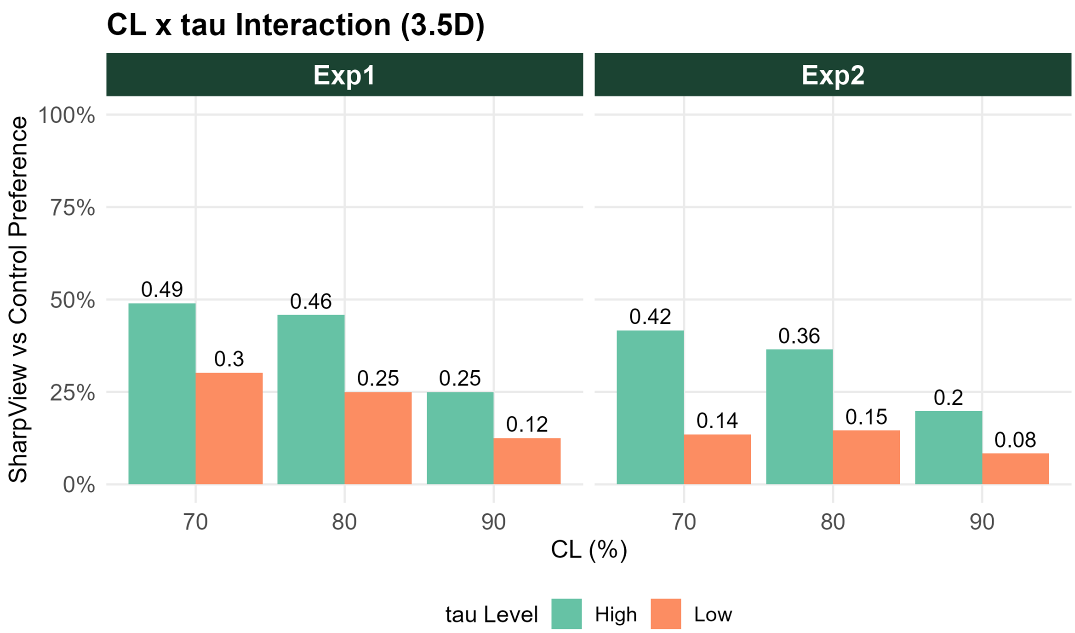


Figure 5.3: SharpView preference rate by CL and τ under the 3.5D aberration condition, shown separately for Experiment 1 and Experiment 2.

Table 5.5: Bonferroni pairwise comparison results for CL under the 3.5D aberration condition, Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
70% vs 80%	1.000	1.000
70% vs 90%	<.001*	.014*
80% vs 90%	.003*	.049*

5.4.2 4.5D Aberration Condition

In the 4.5D condition of Experiment 1, repeated-measures ANOVA results showed that both CL ($F_{2,30} = 4.34, p = .022$) and τ ($F_{1,15} = 5.55, p = .033$) had a significant effect on the SharpView selection rate. However, in Experiment 2, while the effect of τ remained strongly significant ($F_{1,15} = 23.35, p < .001$), the effect of CL was not significant ($F_{2,30} = 2.91, p = .070$). This pattern contrasts with the 3.5D condition, where CL was consistently significant across both experiments.

The effect of τ was consistent with the results observed in the 3.5D condition. In Experiment 1, the selection rates for High τ were 47.9% at CL 70%, 41.7% at CL 80%, and 28.1% at CL 90%, whereas the rates for Low τ were 29.2%, 28.1%, and 22.9%, respectively. In Experiment 2, the selection rates for High τ were also higher than those for Low τ across all CL levels (High τ: 40.6%, 33.3%, 20.8%; Low τ: 12.5%, 15.6%, 13.5%). As shown in Figure 5.4, the bars representing High τ (green) are consistently higher than those representing Low τ (orange) at all CL levels.

The effect of CL differed between the two experiments. In Experiment 1, similar to the 3.5D condition, the selection rate tended to decrease as CL increased. Bonferroni test results showed a significant difference only between CL 70% and 90% (70 vs. 90: $p = .033$; 70 vs. 80: $p = 1.000$; 80 vs. 90: $p = .198$). In contrast, in Experiment 2, the selection rates for the Low τ condition were generally low across CL levels, at 70% (12.5%), 80% (15.6%), and 90% (13.5%), and no statistically significant differences were found between the CL levels (70 vs. 90: $p = .160$; 80 vs. 90: $p = .400$). As shown in Figure 5.4, while the decline in selection rate associated with CL remains distinct in Experiment 1, the selection rate in the Low τ condition of Experiment 2 remains nearly constant from CL 70% to 90%, showing no noticeable change based on CL levels. The Bonferroni test results are presented in Table 5.6.

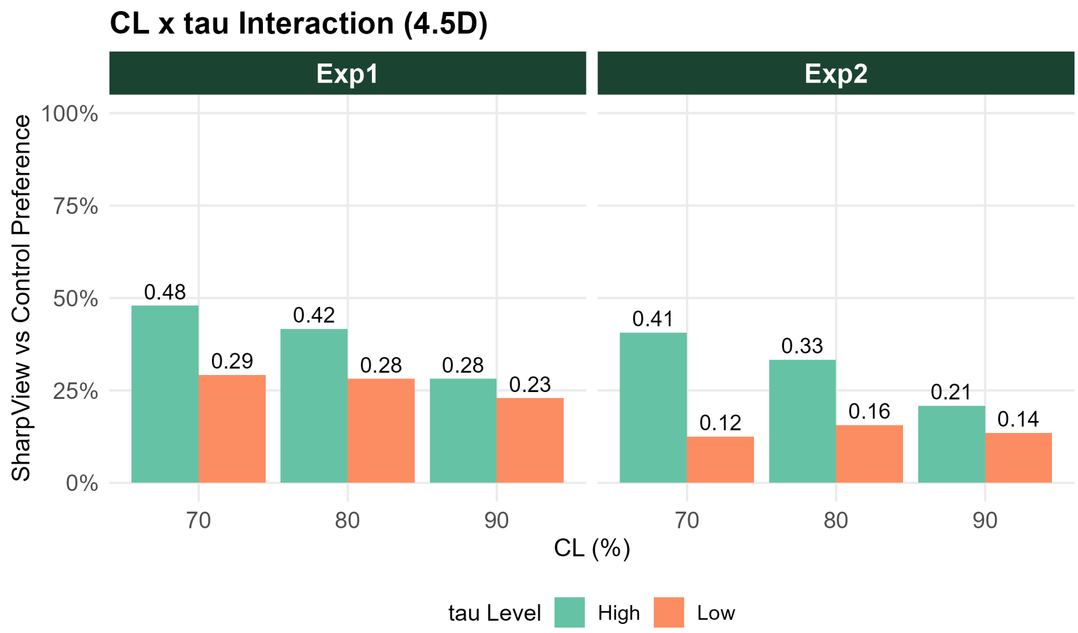


Figure 5.4: SharpView preference rate by CL and τ under the 4.5D aberration condition, shown separately for Experiment 1 and Experiment 2.

Table 5.6: Bonferroni pairwise comparison results for CL under the 4.5D aberration condition, Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
70% vs 80%	1.000	1.000
70% vs 90%	.033*	.160
80% vs 90%	.198	.400

5.4.3 Overall

In both the 3.5D and 4.5D conditions, τ consistently showed a significant effect across both experiments (3.5D: Experiment 1 $F_{1,15} = 17.01, p < .001$; Experiment 2 $F_{1,15} = 22.13, p < .001$; 4.5D: Experiment 1 $F_{1,15} = 5.55, p = .033$; Experiment 2 $F_{1,15} = 23.35, p < .001$). This can be interpreted as follows: as explained in Section 3.8, a Low θ value (corresponding to a High τ) induces stronger correction during the TV optimization process, thereby producing a more pronounced ringing-based correction effect along character edges.

In contrast, the effect of CL varied depending on the aberration and the specific experiment. At 3.5D, the CL effect was significant in both experiments, whereas at 4.5D, it was significant only in Experiment 1. Notably, a consistent trend of reduced selection rates was observed across all conditions when the CL was raised to an excessively high level of 90%. This can be attributed to the fact that higher CL values intensify ringing. At the 90% level, the ringing becomes so pronounced that it actually impairs legibility. In other words, while CL enhances the correction effect up to a certain point, it produces adverse effects beyond that threshold; based on these experimental results, the 70% to 80% range appears to be optimal for SharpView correction. The lack of a significant CL effect under the 4.5D condition in Experiment 2 appears to be due to the fact that the selection rate for Low τ was already low (12.5–15.6%). Consequently, adjusting the CL alone did not result in a noticeable change in the selection rate. In the Incorrect Accommodation scenario (Experiment 2), blur persists regardless of ocular accommodation. Thus, under conditions of high aberration, such as 4.5D, the resulting blur may have been too severe to be overcome solely by adjusting the CL.

In summary, H2 was partially supported. The expectation that a High τ increases the SharpView selection rate was consistently confirmed across conditions 3.5D and 4.5D, as well as in Experiments 1 and 2. However, the expectation that a high CL would yield the greatest effect was not supported. Instead, higher selection rates were observed when the CL was lower (70% or 80%). These results suggest that when designing SharpView parameters, it is important to set an appropriate upper limit rather than simply pursuing the strongest possible correction (i.e., a High CL).

5.5 H3: Effect of Letter Shape on SharpView Preference

5.5.1 3.5D Aberration Condition

In the 3.5D condition of Experiment 1, the $\tau\times$ Letter interaction was statistically significant ($F_{2,30} = 10.55, p < .001$). The same interaction was also significant in Experiment 2 ($F_{2,30} = 5.39, p = .010$).

Under the High τ condition, an examination of selection rates by character revealed that A and E showed markedly higher selection rates than O. In Experiment 1, the selection rates for the High τ condition were 45.8% for A, 46.9% for E, and 27.1% for O. The same pattern emerged in Experiment 2 (A: 39.6%, E: 40.6%, O: 17.7%). Bonferroni post-hoc tests indicated that the differences between A and O, and between E and O, were significant in both experiments (Experiment 1: A vs. O, $p = .041$; E vs. O, $p = .032$; Experiment 2: A vs. O, $p = .0085$; E vs. O, $p = .0054$). Figure 5.5 also shows that the bars representing A and E under the High τ condition are noticeably higher than the corresponding bar for O.

In contrast, this pattern disappeared under the Low τ condition. In Experiment 1, the selection rates for the Low τ condition were 25.0% for A, 13.5% for E, and 29.2% for O. Notably, the rate for O was actually higher than those for A and E. Bonferroni test results showed no significant differences between A and E, A and O, or E and O (A vs. E: $p = .333$; A vs. O: $p = 1.000$; E vs. O: $p = .052$). Similarly, in Experiment 2, the selection rates for the Low τ condition were 13.5% for A, 11.5% for E, and 11.5% for O, showing almost no difference among the three letters and no significant differences in any pairwise comparisons ($p = 1.000$). In other words, under the Low τ condition, the selection rates for A and E dropped to levels as low as those of O, effectively eliminating the differences attributable to letter shape.

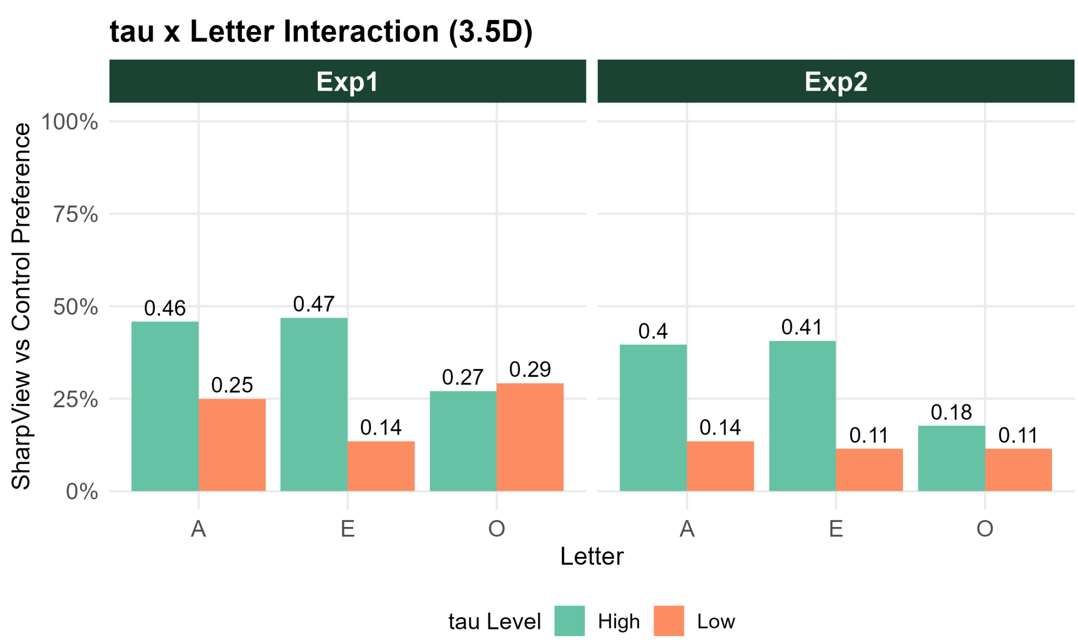


Figure 5.5: SharpView preference rate by τ and Letter under the 3.5D aberration condition, shown separately for Experiment 1 and Experiment 2.

Table 5.7: Bonferroni pairwise comparison results for τ x Letter under the 3.5D aberration condition, Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
High τ: A vs O	.041*	.0085*
High τ: E vs O	.028*	.0054*
Low τ: A vs E	.240	1.000
Low τ: A vs O	1.000	1.000
Low τ: E vs O	.052	1.000

5.5.2 4.5D Aberration Condition

In the 4.5D condition of Experiment 1, unlike the 3.5D condition, the $\tau\times$ Letter interaction was not statistically significant ($F_{2,30} = 1.76, p = .189$). In contrast, in Experiment 2, this interaction remained significant ($F_{2,30} = 6.57, p = .004$).

In Experiment 1, the character selection rates at High τ were 43.8% for A, 39.6% for E, and 34.4% for O. The rate for O increased compared to the 3.5D condition (27.1%), narrowing the gap with A and E. Bonferroni test results also showed no significant differences between A and O or between E and O at High τ (A vs. O: $p = .67$; E vs. O: $p = 1.00$). Figure 5.6 also confirms that the heights of the three bars (A, E, O) for Experiment 1 are closer to one another than those in the 3.5D graph.

In Experiment 2, a pattern similar to that of the 3.5D condition was observed. At high τ, selection rates were 45.8% for A, 30.2% for E, and 18.8% for O. The rate for A was markedly higher than that for O. Bonferroni test results showed a significant difference between A and O (A vs. O, $p = .00091$), whereas the difference between E and O was not significant (E vs. O, $p = .358$).

Under the Low τ condition, there were no significant differences between the characters in either experiment. In Experiment 1, the selection rates for the Low τ condition were 31.2% for A, 18.8% for E, and 30.2% for O. Bonferroni tests revealed no significant differences in any pairwise comparisons (A vs. E: $p = .17$; A vs. O: $p = 1.00$; E vs. O: $p = .25$). Similarly, in Experiment 2, selection rates for the Low τ condition remained low, at 10.4% for A, 17.7% for E, and 13.5% for O, and showed no significant differences (A vs. E: $p = .55$; A vs. O: $p = 1.00$; E vs. O: $p = 1.00$).

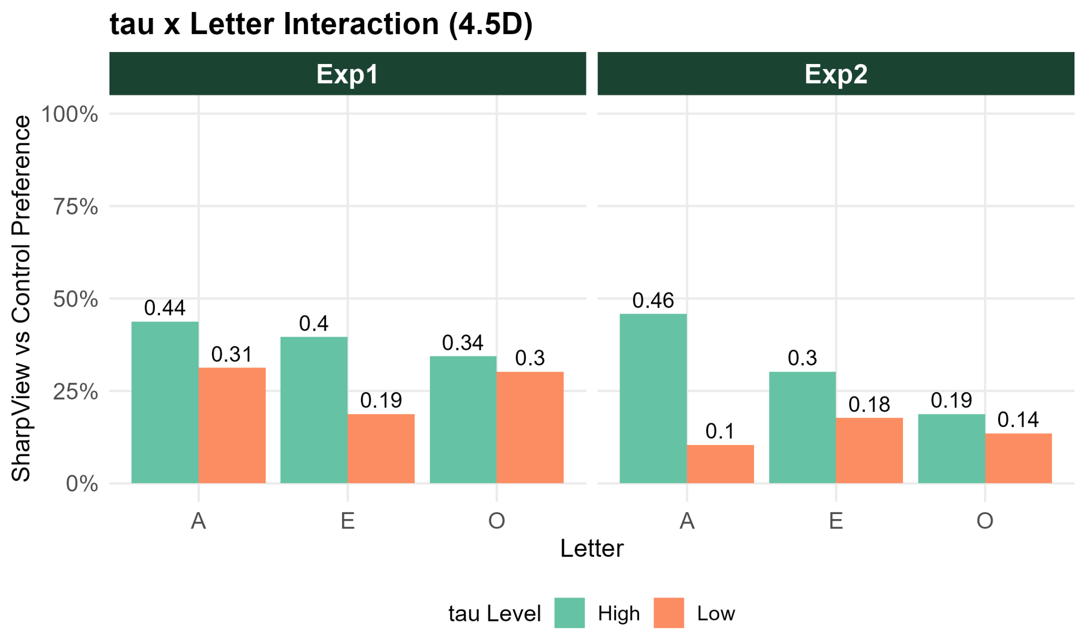


Figure 5.6: SharpView preference rate by τ and Letter under the 4.5D aberration condition, shown separately for Experiment 1 and Experiment 2.

Table 5.8: Bonferroni pairwise comparison results for τ x Letter under the 4.5D aberration condition, Experiment 1 and Experiment 2.

	Experiment 1	Experiment 2
High τ: A vs O	.67	.00091*
High τ: E vs O	1.00	.358
Low τ: A vs E	.17	.55
Low τ: A vs O	1.00	1.00
Low τ: E vs O	.25	1.00

5.5.3 Overall

The patterns of the $\tau\times$ Letter interaction differed between the 3.5D and 4.5D conditions. Under the High τ condition, the tendency for letters A and E to show higher selection rates than O was consistently observed in both experiments conducted at 3.5D and in Experiment 2 at 4.5D. However, in Experiment 1 at 4.5D, this interaction was not statistically significant. The selection rate for O was higher than in the 3.5D condition, thereby narrowing the gap between O and the letters A and E. The precise reason for this result remains unclear, and further research is required.

The reason O showed a lower selection rate compared to other characters is likely linked to SharpView's correction mechanism. SharpView generates a correction effect by inducing ringing along character edges. For characters composed of straight lines, such as A and E, this ringing appears distinctly along the edges, directly contributing to improved legibility. In contrast, since O consists entirely of curves, the ringing effect is dispersed around the character's perimeter, potentially weakening the edge-enhancement effect. However, the current study's results do not clarify whether this outcome stems from inherent limitations of the curved structure itself or from the specific parameter combination used not being optimized for O. Further research is required to determine the exact cause.

It is also noteworthy that differences between characters virtually disappeared under the Low τ condition. This suggests that the pattern posited by H3, specifically that O exhibits the lowest values, manifests only when the correction intensity is sufficiently high (High τ). In other words, the variation in SharpView's effectiveness across character shapes appears to be linked to the intensity of the correction.

In summary, H3 was partially supported. Under the High τ condition, the prediction that A and E would be more effective than O was confirmed in both 3.5D experiments and in Experiment 2 (4.5D), but not in Experiment 1 (4.5D). Under the Low τ condition, the overall correction effect of SharpView itself was negligible, rendering any assessment of differences based on character shape meaningless.

Chapter 6

Limitations and Future Work

This study demonstrated that the SharpView font can enhance legibility under specific parameter combinations in an OST AR environment. However, there are certain limitations, and directions for future research have been proposed accordingly.

Each participant's pupil size was measured once before the experiment and used to generate SharpView images. However, pupil size fluctuates in real-time based on factors such as lighting conditions and gaze direction. Consequently, pre-generated images may not accurately reflect the participant's actual pupil size during the experiment, potentially compromising the accuracy of the SharpView calibration. Future research could address this limitation by integrating real-time pupil size data from Pupil Labs Neon into the application, thereby implementing a system that dynamically provides SharpView images adjusted for real-time changes in pupil size.

Currently, generating SharpView images takes a significant amount of time, making real-time applications difficult. Future research should explore methods to reduce image generation time, such as utilizing GPU parallel processing or developing more efficient algorithms, to ultimately enable the real-time application of SharpView.

This study utilized only three letters: A, E, and O. Future research should expand the scope to include a wider variety of letters and words to more comprehensively validate the effectiveness of SharpView.

The SharpView effect was largely absent for the letter O across both experiments. Although this may be related to its purely curved structure, which could limit the effectiveness of ringing-based edge enhancement, this explanation remains unverified. Future research should investigate whether this is a fundamental limitation of the algorithm for curved letters or an issue of parameter optimization.

This study was conducted using only the Microsoft HoloLens 2. It is necessary to verify whether the validity of SharpView extends beyond this specific device by expanding the experiment to include other commercial OST HMDs with different focal plane distances and optical characteristics.

This study was conducted with 16 participants. Future research needs to recruit a larger number of participants to enhance the generalizability of the findings.

This study modeled the optical characteristics of the eye using a purely mathematical approach based on Zernike polynomials, with pupil diameter and defocus distance as parameters. However, the calculated PSF/OTF values were not independently verified against the participants' actual optical aberrations using instruments such as a wavefront aberrometer. In addition to the spherical aberration modeled in this study, individual participants may have possessed higher-order aberrations, such as coma or astigmatism, that were not accounted for. Consequently, the blur actually experienced by the participants may have differed to some extent from the blur assumed by the algorithm. Future research should compare and validate the modeled PSF against actual optical measurements to confirm whether this correction method is tailored to the specific optical characteristics of individuals rather than relying on an idealized eye model.

Chapter 7

Conclusion

This study evaluated the effectiveness of the SharpView font, which was developed to address reduced text legibility caused by accommodation mismatch in OST AR environments, through experiments involving human subjects. Unlike Arefin et al. [4], who validated the improved SharpView solely through camera-based evaluation, this study conducted experiments using a commercial OST HMD (Microsoft HoloLens 2) under two specific conditions: Incomplete Accommodation and Incorrect Accommodation.

Regarding RQ1, SharpView showed overall selection rates of 32.1% in the Incomplete Accommodation condition (Experiment 1) and 22.6% in the Incorrect Accommodation condition (Experiment 2). In both experiments, participants generally preferred the Control over SharpView. This result aligns with the "mere-exposure effect" discussed in Section 5.3.3. However, for CL values of 70% and 80% combined with the letters A and E, SharpView's selection rate actually exceeded 50%. Examples include the A (59.4%) and E (62.5%) conditions at CL 70% and 3.5D in Experiment 1, and the A (62.5%) condition at CL 70% and 4.5D in Experiment 2. Although these differences did not reach statistical significance after Bonferroni correction, they are noteworthy because SharpView was selected more frequently in these specific combinations, despite the participants' general bias toward the Control driven by familiarity.

Regarding RQ2, the SharpView font was most effective with the combination of High τ and a CL of 70%-80%. High τ consistently increased SharpView selection rates across both aberration conditions and both experiments, whereas a CL of 90% consistently reduced selection rates due to excessive ringing. The effectiveness of SharpView also depended on letter shape. It was effective for letters composed of straight or diagonal lines (A, E) but not for the curved letter (O). However, this pattern was not observed under the 4.5D condition in Experiment 1.

Observations made during the experiment revealed instances where the effect of SharpView appeared momentarily but then vanished. This may be attributed to the failure to reflect real-time changes in pupil diameter. As noted in Chapter 6, this study measured each participant's pupil diameter only once, before the experiment, and used that value to generate the SharpView images. However, since pupil diameter fluctuates in real time based on factors such as lighting and gaze direction, the intended correction effect might not have been fully realized if the pupil diameter during the actual experiment differed from the pre-measured value. The intermittent nature of the effect appears consistent with this possibility. Although this study validated SharpView using only the Arial Bold font, the effect is expected to apply similarly to other fonts, given that SharpView's correction mechanism relies on a principle that generates ringing along letter edges. However, further validation using a variety of fonts is required to confirm this conclusively.

The results of this study do not demonstrate that SharpView, in its current form, definitively outperforms the Control. However, by identifying conditions under which SharpView is ineffective, specifically excessive contrast reduction, insufficient correction strength, and curved letter shapes, this study provides practical guidelines for future parameter adjustments and algorithmic improvements. Such enhancements could bring SharpView closer to consistently and clearly outperforming the Control, thereby contributing to improved text legibility in future OST AR systems used across diverse fields such as healthcare, education, and manufacturing.

Bibliography

[1] Yuta Itoh, Tobias Langlotz, Jonathan Sutton, and Alexander Plopski. Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Computing Surveys*, 54:1–36, 07 2021.

[2] Carlos Montalto Cruz, Ignacio Garcia-Dorado, Daniel Aliaga, Manuel Oliveira, and Feng Meng. A total variation approach for customizing imagery to improve visual acuity. *ACM Transactions on Graphics*, 34:1–16, 05 2015.

[3] Kohei Oshima, Kenneth R Moser, Damien Constantine Rompapas, J. Edward Swan, Sei Ikeda, Goshiro Yamamoto, Takafumi Taketomi, Christian Sandor, and Hirokazu Kato. Sharpview: Improved clarity of defocused content on optical see-through head-mounted displays. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 173–181, 2016.

[4] Mohammed Safayet Arefin, Carlos Montalto, Alexander Plopski, and J. Edward Swan. A sharpview font with enhanced out-of-focus text legibility for augmented reality systems. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 31–41, 2024.

[5] Trey Cook, Nate Phillips, Kristen Massey, Alexander Plopski, and Christian Sandor. User preference for sharpview-enhanced virtual text during non-fixated viewing. pages 1–400, 03 2018.

[6] Microsoft. About HoloLens 2. https://learn.microsoft.com/en-us/hololens/hololens2-hardware, 2023. Accessed: June 9, 2026.

[7] Sara Condino, Marina Carbone, Roberta Piazza, Mauro Ferrari, and Vincenzo Ferrari. Perceptual limits of optical see-through visors for augmented reality guidance of manual tasks. *IEEE Transactions on Biomedical Engineering*, 67(2):411–419, 2020.

[8] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33–33, 03 2008.

[9] Joseph L. Gabbard, Divya Gupta Mehra, and J. Edward Swan. Effects of ar display context switching and focal distance switching on human performance. *IEEE Transactions on Visualization and Computer Graphics*, 25(6):2228–2241, 2019.

[10] F. W. Campbell and G. Westheimer. Dynamics of accommodation responses of the human eye. *The Journal of Physiology*, 151(2):285–295, 1960.

[11] Mohammed Safayet Arefin, Nate Phillips, Alexander Plopski, Joseph L. Gabbard, and J. Edward Swan. The effect of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur on human performance in optical see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2014–2025, 2022.

[12] George Koulieris, Kaan Akşit, Michael Stengel, R. Mantiuk, Katerina Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. *Computer Graphics Forum*, 38:493–519, 06 2019.

[13] Praneeth Chakravarthula, David Dunn, Kaan Akşit, and Henry Fuchs. Focusar: Auto-focus augmented reality eyeglasses for both real world and virtual imagery. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2906–2916, 2018.

[14] Xinxing Xia, Yunqing Guan, Andrei State, Praneeth Chakravarthula, Kishore Rathinavel, Tat-Jen Cham, and Henry Fuchs. 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, 2019.

[15] Andrew Maimone and Henry Fuchs. Computational augmented reality eyeglasses. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 29–38, 2013.

[16] Feng Xu and Dayang Li. Software based visual aberration correction for hmds. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 246–250, 2018.

[17] Pablo Artal. *The Eye as an Optical Instrument*, pages 285–297. Springer International Publishing, Cham, 2016.

[18] G.-M Dai. *Wavefront optics for vision correction*. 01 2008.

[19] Matheus Krueger, Manuel Oliveira, and Airton Kronbauer. Personalized visual simulation and objective validation of low-order aberrations of the human eye. pages 64–71, 10 2016.

[20] Vasudevan Lakshminarayanan and Andre Fleck. Zernike polynomials: A guide. *Journal of Modern Optics - J MOD OPTIC*, 58:1678–1678, 04 2011.

[21] M. Liu, Y. He, and Ye Bin. Image zernike moments shape feature evaluation based on image reconstruction. 10:191–195, 9 2007.

[22] Robert J. Noll. Zernike polynomials and atmospheric turbulence∗. *J. Opt. Soc. Am.*, 66(3):207–211, Mar 1976.

[23] Andrew B. Watson. Computing human optical point spread functions. *Journal of vision*, 15 2:26, 2015.

[24] Junzhong Liang and David Williams. Aberrations and retinal image quality of the normal human eye. *Journal of the Optical Society of America A*, 14:2873–2883, 11 1997.

[25] M. Alonso and A.B. Barreto. Pre-compensation for high-order aberrations of the human eye using on-screen image deconvolution. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, volume 1, pages 556–559 Vol.1, 2003.

[26] John Yellott and John Yellott. Correcting spurious resolution in defocused images - art. no. 64920o. *Proceedings of SPIE - The International Society for Optical Engineering*, 6492, 03 2007.

[27] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[28] J.M. Bioucas-Dias, M.A.T. Figueiredo, and J.P. Oliveira. Total variation-based image deconvolution: a majorization-minimization approach. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 2, pages II–II, 2006.

[29] JosÉ M. Bioucas-Dias and MÁrio A. T. Figueiredo. A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Processing*, 16(12):2992–3004, 2007.

[30] A. Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40:120–145, 2011.

[31] Amir Beck and Marc Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 18(11):2419–2434, 2009.

[32] Carlos Montalto Cruz. *Stability analysis of three inverse problems: The study of the hyperbolic inverse boundary value problem, current density impedance imaging and image based visual correction*. Ph.d. dissertation, Purdue University, 2014.

[33] Mohammed Safayet Arefin. *Augmented reality fonts with enhanced out-of-focus text legibility*. Ph.d. dissertation, Mississippi State University, 2022.

[34] L. Thaler, A.C. Schütz, M.A. Goodale, and K.R. Gegenfurtner. What is the best fixation target? the effect of target shape on stability of fixational eye movements. *Vision Research*, 76:31–42, 2013.

[35] Robert Zajonc. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9:1–27, 06 1968.

[36] Sofie Beier and Kevin Larson. How does typeface familiarity affect reading performance and reader preference? *Information Design Journal*, 20, 10 2013.