

# CHARACTERIZATION OF ACCOMMODATION RESPONSE IN FOVEATED NEAR-EYE HOLOGRAPHIC DISPLAYS

Jani Mäkinen, Erdem Sahin, Atanas Gotchev

Faculty of Information Technology and Communication Sciences, Tampere University, Finland

## ABSTRACT

Foveation is an extremely powerful tool for reducing computational burden in near-eye displays. Applying such methods for displays capable of producing the majority of 3D visual cues, however, requires accurate models and knowledge of the human visual perception to exploit its weaknesses effectively. Spatial resolution is the main concern in regular 2D displays, whereas in the case of 3D displays additional visual cues, such as accommodation, require attention. This paper analyzes such cues for near-eye holographic displays by considering the perceptual properties of the human vision, particularly the visual acuity relative to the gaze direction. We develop a numerical wave optics based reconstruction tool incorporating a perceptual model of the human eye to simulate the natural viewing process with high accuracy. Based on our analysis, general guidelines are found regarding the accuracy of the accommodation cues relative to the gaze of the viewer. More importantly, the developed method can be included into the design process of near-eye holographic displays to computationally optimize their parameters and the foveated rendering techniques.

*Index Terms* — Holography, near-eye displays, foveation, perceptual modeling, wave optics

## 1. INTRODUCTION

Holographic displays, despite their ability to replicate 3D content highly realistically, have not yet seen widespread use due to the extremely demanding computational requirements to achieve full parallax and a wide viewing angle. By restricting the position of the viewer near to the display, as done in near-eye displays [1, 2, 3], the limitations of the human visual system in addition to the restricted position can be exploited to relieve the display requirements, such as viewing angle. Moreover, in such displays the gaze of the viewer can be more easily tracked, thus allowing for foveated rendering methods to further optimize the computations. For this purpose, it is important not only to know the perceptual limits of human vision, such as the spatial resolution and accommodation response, but also to be able to accurately simulate the viewing process of the human eye.

Previous research on characterizing the properties of 3D displays in relation to visual perception mainly considered ray-based light field (LF) displays [4]. In an effort to extend the analysis to holographic stereograms (HS), our previous work [5] employed wave optics based numerical viewing simulations. In these simulations the human eye is considered as a camera with a thin lens for the pupil and a flat uniformly sampled sensor for the retina. However, more accurate analysis and models are needed specifically for near-eye applications. A more precise model of the human visual system (HVS) was utilized in [6], where an importance sampling model was developed based on the spatial resolution decrease away from the visual axis. However, the eye was still approximated by a flat two-plane parametrization. Moreover, the

analysis considered only ray-based optics. As a result, there is an apparent need for accurate analysis and simulation of human vision for wave optics based applications.

In this paper, we analyze the accommodation response of the HVS for a near-eye holographic display. To improve the accuracy of the HVS model, a non-uniform sampling scheme is adopted, based on the density of the human retinal ganglion cell receptive field [7]. Additionally, the simulations include the curvature of the retina. The approach for simulating the viewing process of the human eye utilizes numerical wavefield propagation. We analyze the accommodation response via point spread functions (PSF) as well as the modulation transfer functions (MTF) from the perceived retinal images. The response is evaluated as a function of eccentricity by changing the position of the displayed content relative to the gaze horizontally. Thus, we find the accuracy of the accommodation cues needed to be provided by the hologram in the peripheral vision.

## 2. METHOD

### 2.1. Computational human eye model

Near-eye foveated displays take advantage of the shortcomings of the HVS and rely on models to provide the necessary information regarding e.g. sampling and rendering budget. Such models can also be incorporated into simulation tools designed to mimic the visual process of a human eye. In general, these simulation tools consider the human eye as a camera with a thin lens, where the lens is equivalent to the pupil and the sensor to the retina. This configuration defines three parallel planes: display (or hologram), lens and sensor. Each plane needs to be discretely sampled according their corresponding sampling requirements for numerical simulations, which is typically done in a uniform fashion. Here we will describe how to improve the human eye model, specifically targeting the sampling of the sensor plane.

A key part of an accurate perceptual model of the human vision is the decrease of visual acuity (away from gaze). Here we adopt the model proposed in [7] describing the density of the retinal ganglion cell receptive field as a function of retinal eccentricity  $r = \sqrt{x^2 + y^2}$  and meridian  $m$

$$\rho(r, m) = 2\rho_{cone} \left(1 + \frac{r}{41.03}\right)^{-1} \times \left[ a_m \left(1 + \frac{r}{r_{2,m}}\right)^{-2} + (1 - a_m) \exp\left(\frac{r}{r_{e,m}}\right) \right]. \quad (1)$$

The constants  $\rho_{cone}$ ,  $a_m$ ,  $r_{2,m}$ ,  $r_{e,m}$  fit the model along the four different meridians (temporal, superior, nasal, inferior) [7]. The sampling of the retina (in angle) is then defined from the density as [7]

$$\sigma(x, y) = \frac{1}{r} \sqrt{\frac{2}{\sqrt{3}} \left( \frac{x^2}{\rho(r, 1)} + \frac{y^2}{\rho(r, 2)} \right)}. \quad (2)$$

Thus, the visual acuity drop-off is modeled by non-uniformly sampling the retinal surface in further steps of the analysis.

Another often overlooked property of the HVS in numerical simulations is the curvature of the retina. Though the conventional flat two-plane parametrization is an acceptable approximation, the accuracy of the eye model is improved to some extent by including the curved parametrization. In this paper, we parametrize the retina on a spherical surface with a diameter of  $l$ . As a result, the retina is sampled at different  $z$ -coordinates based on the eccentricity. Specifically, a sample at the eccentricity  $\phi$  on the retina is located at distance  $z_u$  in  $z$ -direction from the pupil:

$$z_u = \frac{l}{2} + \frac{l}{2} \cos \phi. \quad (3)$$

Thus, each sample on the retina has both a horizontal and a depth coordinate.

The final part of the computational model is the pupil, which is modelled as a thin lens of width  $D$ . The 1D transmittance function  $T(s)$  of a thin lens is defined as

$$T(s) = \exp\left(\frac{-j\pi}{\lambda f} s^2\right), \quad (4)$$

where  $\lambda$  is the wavelength of the monochromatic light and  $f$  is the focal length of the lens. By changing  $f$ , the lens can be set to focus at varying distances, thus providing the ability to evaluate the accommodative properties of the human eye in numerical simulations.

## 2.2. Numerical simulation via scalar field diffraction

In addition to the perceptually accurate eye model, the propagation of light from the display to the eye needs to be modeled. A wave optics based approach is adopted here to study the perceived visual properties through numerical simulations. For simplicity, let us consider the equations in 1D. The scalar optical field  $U(\xi; z_0)$  is propagated utilizing plane wave decomposition, i.e. for each spatial frequency component  $f_\xi$  in the field, a plane wave is propagated with its corresponding complex amplitude and sampled at the desired secondary positions  $(x, z)$ . The contributions of each plane wave are combined to obtain the total field  $U(x; z)$  as

$$U(x; z) = \int \mathcal{F}\{U(\xi; z_0)\}(f_\xi) \times \exp[-2\pi j(f_\xi x + f_z z)] df_\xi, \quad (5)$$

where

$$f_z = \sqrt{\frac{1}{\lambda^2} - f_\xi^2}. \quad (6)$$

The fields are discretely sampled in the numerical simulations, thus the integral in Eq. 5 can be replaced with a summation and a discrete Fourier transform is utilized to obtain the complex amplitudes for each plane wave.

The field at the hologram plane is generated based on the content to-be-recorded on the hologram. As our simulation setup models the content as point source(s) of light, the field is obtained as a summation of contributions from each point source  $p$ . Utilizing the Fresnel kernel, the 1D field  $O(x; 0)$  is defined for a single point source as

$$U(x; 0) = a_p \frac{\exp(jkz_p)}{\sqrt{j\lambda z_p}} \exp\left(jk \frac{(x - x_p)^2}{2z_p} + \phi_p\right), \quad (7)$$

where  $a_p$  is the amplitude of the point source,  $\phi_p$  is the relative phase,  $\lambda$  is the wavelength of the monochromatic light and

$k = 2\pi/\lambda$  is the wave number. Due to the viewing conditions of the near-eye display setup, i.e. short viewing distance and stationary position of the eye, spherical illumination for the hologram is beneficial to use; it reduces the size of the viewing zone to extend the field-of-view [8].

## 3. RESULTS AND DISCUSSION

The numerical reconstruction setup for the analysis is presented in Fig. 1. That is, a single point source is placed at coordinates  $(x_p, z_p)$  and a hologram of it is produced. The hologram is then propagated towards the simulated eye at distance  $z_{eye}$ , first passing through the lens with a transmittance function  $T(s)$ . The transmittance function can be modified to focus the eye at different distances. Finally, the intensity values are retrieved from the retina at the positions defined by Eq. 2 and 3. These intensities compose the perceived PSF for the human eye, from which the MTF is obtained as the magnitude of its spectrum. The MTF describes the contrast magnitude across different spatial frequencies.

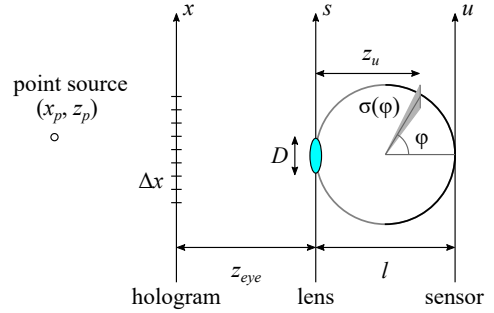


Figure 1. Numerical reconstruction setup for the analysis.

Let us first consider the behaviour of the MTFs as a function of spatial frequency. For this purpose, the presented numerical simulation tool is utilized with the following parameters: the point is set at  $z_p = -50$  cm (i.e. 50 cm behind the hologram) and a Fresnel hologram of it is generated with a sampling step  $\Delta x = 2$   $\mu\text{m}$ , number of samples  $N = 8192$  and wavelength  $\lambda = 534$  nm. The location of the eye and the pupil size are fixed to  $z_{eye} = 2$  cm and  $D = 5$  mm, respectively. The simulated eye is set to focus at the exact distance of the point source, as well as around it at various shifted depths. We define the accommodation shift  $\Delta z$  in diopters relative to the location of the point, i.e.

$$\Delta z = \frac{1}{z_f} - \frac{1}{z_{eye} - z_p}, \quad (8)$$

where  $z_f$  is the distance the eye is set to focus. The focus distance  $z_f$  is related to the focal length of the thin lens in the eye model according to the imaging equation

$$\frac{1}{f} = \frac{1}{z_f} + \frac{1}{l}. \quad (9)$$

Furthermore, the foveated aspect of the model is taken into account by changing the horizontal position of the point. The distance from the gaze is described as an eccentricity angle (see  $\phi$  in Fig. 1). As shown in Fig. 2, the MTF decays towards higher frequencies regardless of the accommodation shift and eccentricity. The relative magnitude between the different MTFs is crucial, as contrast magnitude is one of the factors driving the accommodation response of the eye [4]. Consequently, the likeliest accommodation distance  $\hat{z}_f$  is estimated here as the accommodation shift with the largest MTF magnitude at a given spatial frequency.

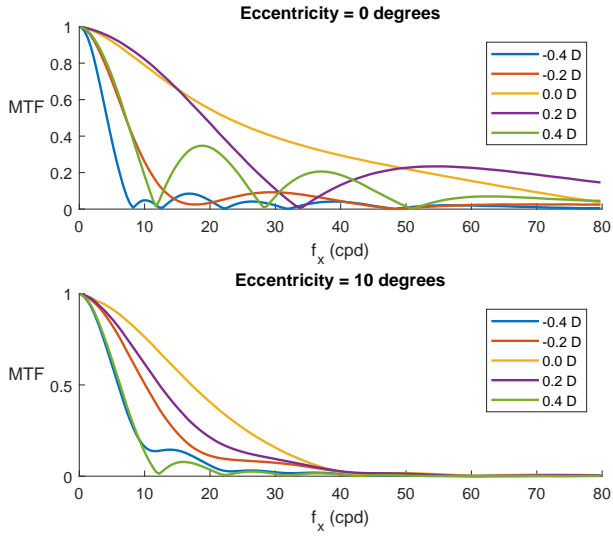


Figure 2. MTFs for different accommodation shifts as a function of spatial frequency. The point is placed 50 cm behind the hologram and horizontally at the center of the gaze (top) or deviated 10 degrees from the gaze (bottom).

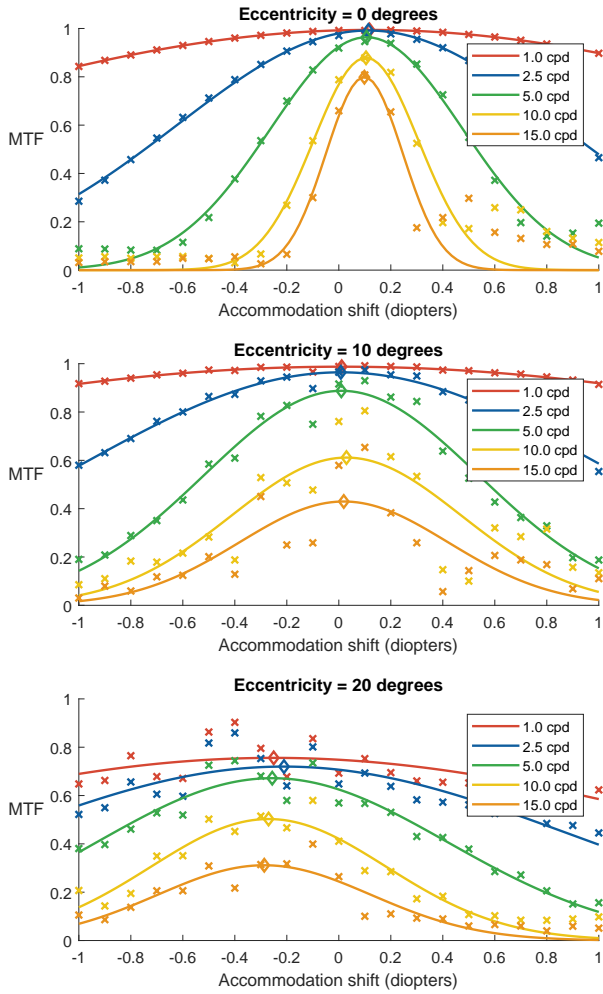


Figure 3. MTF values and a Gaussian fit as a function of accommodation shift. The point is placed 50 cm behind the hologram and horizontally at the center of the gaze (top), deviated 10 degrees (center) or 20 degrees (bottom) from the gaze. The diamond marker denotes the peak of the Gaussian.

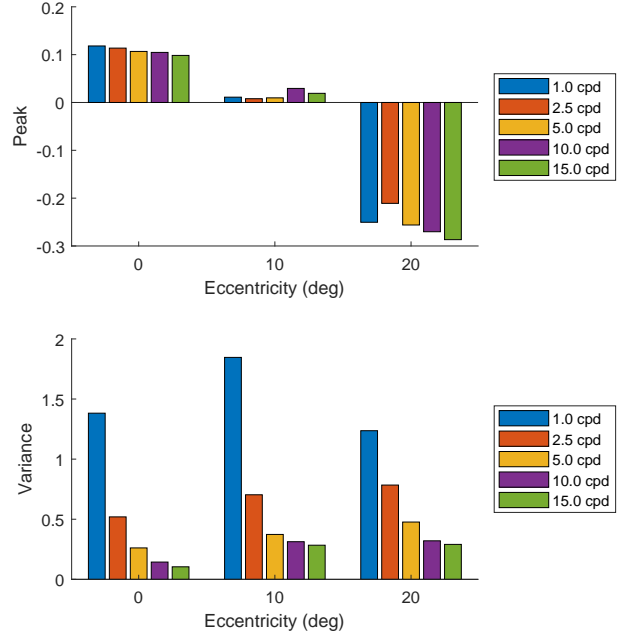


Figure 4. Peak location and variance of the Gaussian functions fitted to the MTF data points at different eccentricity angles.

A representative set of spatial frequencies (for the HVS) is chosen to further illustrate the differences in MTF magnitude between different accommodation shifts. The choice of frequencies is motivated by the neural transfer function, according to which the human eye is most sensitive to frequencies around 10–15 cycles per degree (cpd) [9]. Therefore, the MTF is evaluated at five different spatial frequencies (1, 2.5, 5, 10 and 15 cpd) and at  $\Delta z$  between -1 and +1 diopters (0.1 diopter steps) around the position of the point. As a result, we obtain 21 data points for each spatial frequency. In order to minimize the effect of possible outliers, a unimodal Gaussian is fitted to the data points, from which the maximum value is estimated as  $\hat{z}_f$ . The results shown in Fig. 3 illustrate changes in the contrast magnitude and gradient across different spatial frequencies and point eccentricities. Certain factors are maintained across different eccentricities: the higher spatial frequencies have larger contrast gradient, but lower contrast magnitude. The estimated accommodation distance  $\hat{z}_f$  changes slightly as the distance from the gaze changes, though in general remains within the HVS depth-of-field of 0.3 diopters [10]. Interestingly, at 0 degree retinal eccentricity  $\hat{z}_f$  is estimated to be slightly behind the actual point. Previous studies have observed that such accommodation lag is present even under natural viewing conditions [11, 12].

Finally, let us examine the width of the MTF peak in more detail. The width of the peak (or contrast gradient) is relative to the amount of observed blur: shallow gradient indicates similar levels of blur across a wider range of depth values. The variances of the Gaussian functions fitted to the MTF data points, shown in Fig. 4, indicate that the peak width slightly increases at larger eccentricity angles. Based on the assumption that the accommodation response is driven by the contrast gradient, in addition to the magnitude, the strength of the response trigger can be expected to decrease as the distance from the gaze increases. Our findings are supported by previous research [13], where it was found that peripheral vision can trigger accommodation response, however, at a decreased magnitude as the retinal eccentricity increases. As a consequence, content in the peripheral vision could

be thus described by more ambiguous depth values, which would simplify hologram generation at such regions. It is important to note, though, that other factors may contribute to the accommodation response. Furthermore, it is expected that the accommodation cue is most relevant in the direction of the gaze due to the natural tendency to gaze towards the region-of-interest. i.e. where the focused content is. Nevertheless, other means of accurate reproduction of depth are required for proper focus cues in the peripheral vision. Retinal blur, in particular, contributes to the sense of depth as well as motion estimation for moving targets [14, 15]. Moreover, estimating depth from blur in the peripheral vision can be expected to help drive the accommodation (as the gaze moves) towards the correct distance. Further research on introducing psychophysical limits of human vision to the analysis and model (e.g. blur and contrast discrimination [16]) is left for future work in this topic.

#### 4. CONCLUSIONS

We have presented in this paper preliminary analysis of accommodation response in foveated near-eye holographic displays. The main focus has been on the effects of visual acuity across the retina which is taken into account by the presented computational eye model. Our initial results from the numerical wavefield simulations have suggested that the strength of the accommodation trigger decreases away from the central gaze. Thus, the content recorded on the hologram at such positions can have varying positions in depth and still be accommodated similarly by the human eye. Importantly, this suggests that at the peripheral vision it is feasible to use a fixed set of depth planes for generating holograms, thus permitting the use of pre-calculated patterns and simplifying the calculations significantly. However, further analysis is still required to combine the results of numerical simulations with known psychophysical quantities related to human vision.

#### 5. REFERENCES

- [1] D. Lanman and D. Luebke, "Near-eye light field displays," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 220:1–220:10, Nov. 2013.
- [2] A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 85:1–85:16, July 2017.
- [3] P. Sun, S. Chang, S. Liu, X. Tao, C. Wang, and Z. Zheng, "Holographic near-eye display system based on double-convergence light gerchberg-saxton algorithm," *Opt. Express*, vol. 26, no. 8, pp. 10140–10151, Apr 2018.
- [4] H. Huang and H. Hua, "Systematic characterization and optimization of 3D light field displays," *Opt. Express*, vol. 25, no. 16, pp. 18508–18525, Aug. 2017.
- [5] J. Mäkinen, E. Sahin, and A. Gotchev, "Analysis of accommodation cues in holographic stereograms," in *2018 - 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, June 2018, pp. 1–4.
- [6] Q. Sun, F.-C. Huang, J. Kim, L.-Y. Wei, D. Luebke, and A. Kaufman, "Perceptually-guided foveation for light field displays," *ACM Trans. Graph.*, vol. 36, no. 6, pp. 192:1–192:13, Nov. 2017.
- [7] A. B. Watson, "A formula for human retinal ganglion cell receptive field density as a function of visual field location," *Journal of Vision*, vol. 14, no. 7, pp. 1–17, June 2014.
- [8] C. Haupt, A. Kolodziejczyk, and H. J. Tiziani, "Resolution and intensity distribution of output images reconstructed by sampled computer-generated holograms," *Appl. Opt.*, vol. 34, no. 17, pp. 3077–3086, June 1995.
- [9] D. G. Green and F. W. Campbell, "Effect of focus on the visual response to a sinusoidally modulated spatial stimulus\*," *J. Opt. Soc. Am.*, vol. 55, no. 9, pp. 1154–1157, Sep 1965.
- [10] S. Marcos, E. Moreno, and R. Navarro, "The depth-of-field of the human eye from objective and subjective measurements," *Vision Research*, vol. 39, no. 12, pp. 2039 – 2049, 1999.
- [11] M. Torii, Y. Okada, K. Ukai, J. S. Wolffsohn, and B. Gilmartin, "Dynamic measurement of accommodative responses while viewing stereoscopic images," *Journal of Modern Optics*, vol. 55, no. 4–5, pp. 557–567, 2008.
- [12] H. Mizushima, J. Nakamura, Y. Takaki, and H. Ando, "Super multi-view 3d displays reduce conflict between accommodative and vergence responses," *Journal of the Society for Information Display*, vol. 24, no. 12, pp. 747–756, 2016.
- [13] Y. Gu and G. E. Legge, "Accommodation to stimuli in peripheral vision," *J. Opt. Soc. Am. A*, vol. 4, no. 8, pp. 1681–1687, Aug. 1987.
- [14] D. Vishwanath and E. Blaser, "Retinal blur and the perception of egocentric distance," *Journal of Vision*, vol. 10, no. 10, pp. 1–16, Aug. 2010.
- [15] A. B. Watson and A. J. Ahumada, "Blur clarified: A review and synthesis of blur discrimination," *Journal of Vision*, vol. 11, no. 5, pp. 1–23, Sep. 2011.
- [16] B. Wang and K. J. Ciuffreda, "Foveal blur discrimination of the human eye," *Ophthalmic and Physiological Optics*, vol. 25, no. 1, pp. 45–51, 2005.