Analysis of accommodation cues in holographic stereograms

ANALYSIS OF ACCOMMODATION CUES IN HOLOGRAPHIC STEREOGRAMS

Jani Mäkinen, Erdem Sahin, Atanas Gotchev

Laboratory of Signal Processing, Tampere University of Technology, Tampere, Finland

ABSTRACT

The simplicity of the holographic stereogram (HS) makes it an attractive option in comparison to the more complex coherent computer generated hologram (CGH) methods. The cost of its simplicity is that the HS cannot accurately reconstruct deep scenes due to the lack of correct accommodation cues. The exact nature of the accommodation cues present in HSs, however, has not been investigated. In this paper, we provide analysis of the relation between the hologram sampling properties and the perceived accommodation response. The HS can be considered as a generator of a discrete light field (LF) and can thus be examined by considering the light ray oriented nature of the hologram diffracted light. We further support the analysis by employing a numerical reconstruction tool simulating the viewing process of the human eye. The simulation results demonstrate that HSs can provide accommodation cues depending on the choice of hologram segmentation size. It is further demonstrated that the accommodation response can be enhanced at the expense of loss in perceived spatial resolution.

Index Terms — Holographic stereogram, light field, accommodation

1. INTRODUCTION

In stereoscopic 3D displays, the conflict of visual cues due to their inability to deliver the accommodation (focus) cues results in the so-called vergence-accommodation conflict. Due to this problem, several alternative 3D display techniques have been proposed to resolve this issue, including light field [1, 2] and holographic displays [3, 4]. These methods can provide a wide variety of the visual cues required for 3D perception, including motion parallax, occlusions and accommodation cues. In holographic display methods the coherent accurate representations, such as the phase-added stereogram [5], rely on intrinsic information about the object shape and location, either through a model description (e.g. point cloud or polygon mesh) or depth information (e.g. depth maps). The accurate location allows such methods to provide accurately the focus cues even for scenes with pronounced depth (i.e. deep scenes) [6]. However, the requirement for object information limits the usability of these methods often to synthetic scenes or to a cumbersome capture procedure and equipment for a real scene. Incoherent image-based CGHs provide a simplified solution to this issue at the cost of 3D reconstruction accuracy. In particular, the HS [7] is a widely utilized model due to its use of only images and fast computation. However, it is widely acknowledged that the HS cannot provide correct accommodation cues in the case of deep scenes [8, 9, 10]. The exact properties of the accommodation cues provided by HSs based on the parameters of the hologram, in addition to the range for which the correct accommodation response is achieved, has not been provided yet.

The objective of this paper is to study the accommodation properties of HSs. The characterization of such properties has been analyzed in [11] for 3D ray-based LF displays, which rely on ray optics formalism of light. Here we carry out a similar analysis to examine the effects of hologram parameters on the accommodation response of HSs. The theoretical analysis is supplemented with simulations of a human viewer perceived images, obtained through a wavefield propagation-based pipeline taking into account the properties of the human visual system (HVS). The results of this paper can be utilized as guidelines for determining the parameters of HSs and the recorded scene in accommodation critical applications.

2. HOLOGRAPHIC STEREOGRAM

The HS represents the content recorded on the hologram, i.e. the complex wavefield of the 3D scene, as a collection of planar wave segments, which together form a piecewise approximation of the complete wavefield. Thus, the HS object field (in 1D for simplicity) is defined along the hologram plane x as

$$O_{\mathrm{HS}}(x) = \sum_{m=1}^{M} \mathrm{rect}\left(\frac{x - m\Delta x}{\Delta x}\right) \times \sum_{p=1}^{P} \sqrt{L[m,p]}\exp\left(j2\pi f_x^{mp} x\right), \quad (1)$$

where rect is the rectangular function, Δx is the holographic element (hogel) size, M is the total number of hogels, P is the total number of LF samples within a hogel, f_x^{mp} is the spatial frequency value and $L[m,p]$ is the intensity of a discrete LF sample at $m\Delta x$ and $p\Delta u$ (see Fig. 1). The planar wavefronts are emitted to different directions from the hologram plane based on their spatial frequencies f_x, which can be interpreted as diffraction angles θ_x along the x-axis, based on the grating equation

$$\sin\theta_x = \lambda f_x, \quad (2)$$

where λ is the wavelength of the monochromatic light. That is, by interpreting the set of plane wave segments as a set of light rays according to their propagation direction, the HS can be considered to be information-wise equivalent to a discrete LF, where the hologram plane is sampled at the hogel centers and a collection of light rays is emitted in directions specified by the hologram parameters. Each hogel contains several hologram pixels, thus segmenting the hologram. The hologram generation can be accelerated by realizing the inner sum in Eq. 1 as an inverse discrete Fourier transform. Thus, fast Fourier transform (FFT) based implementation is very commonly utilized and should, therefore, be considered in the analysis. Although in the FFT-based implementation the spatial frequency grid is fixed, which introduces quantization error, the following analysis is valid for both cases.

The hologram pixel pitch Δ_ξ defines the largest spatial frequency content the hologram is able to produce and thus the max-

978-1-5386-6125-3/18/$31.00

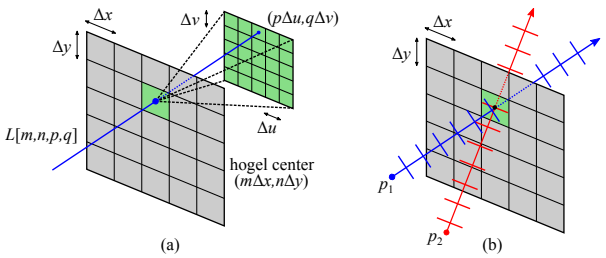

Figure 1: The HS is parametrized by discrete locations on two planes: the hologram plane (x, y) and a parametrization plane (u, v) (a). Plane wave decomposition of the HS (b).

imum diffraction angle θ_{max}, which is defined as

$$\theta_{max} = \arcsin \frac{\lambda}{2\Delta_\xi}. \tag{3}$$

The spatial frequency sampling within a hogel is uniform and is dictated by the size of the hogel Δx. According to Eq. 2, this corresponds to a non-uniform sampling step in the angular domain, especially for large angles. However, the main contribution to the angular sampling Δ_θ is due to the hogel size Δx and can be, therefore, estimated as

$$\Delta_\theta \approx \arcsin\left(\frac{\lambda}{\Delta x}\right). \tag{4}$$

The angular sampling is vital in terms of the accommodation response of the human eye viewing the hologram, as it determines how accurately a hogel can direct the plane waves in the desired direction. Based on the given angular resolution, in one dimension, we define the total number of planar wavefront segments (correspondingly rays) that can be accurately directed inside the eye pupil from a given hogel as

$$N_\theta = \frac{D}{z_{eye} \tan \Delta_\theta}, \tag{5}$$

where D is the pupil diameter and z_{eye} is the distance between the pupil and the hologram. On the other hand the hogel size determines the perceived spatial resolution, which is usually chosen according to the properties of the HVS. When the human eye is assumed to be a diffraction limited imaging system, the minimum resolvable distance between two points at distance z_{eye} is defined by the Rayleigh criterion as [12]

$$\Delta_x^{HVS} = 1.22\frac{\lambda z_{eye}}{D}. \tag{6}$$

That is, for a given eye pupil size D and intended viewing distance z_{eye}, Δ_x^{HVS} is the upper limit for the hogel size which will still ensure that the perceived spatial resolution is maximized. Please note that such choice of hogel size results in $N_\theta \approx 1$ according to Eq. 5. Thus, due to the trade-off between the angular and spatial resolutions, one should sacrifice from the spatial resolution to improve the angular resolution. As will be shown in Sec. 3, this is indeed necessary to enhance the accommodation response and thus maintain correct focus cues.

Another factor to consider in the accommodation analysis is the number of planar wavefront segments emitted from multiple different hogels, for a given 3D point, and intercepted by the eye. The number of such wavefront segments N_h is defined for point at depth z_p away from the hologram plane as

$$N_h = \left|\frac{z_p}{z_{eye} - z_p}\right| \frac{D}{\Delta x}. \tag{7}$$

If the pupil size D and viewing distance z_{eye} are fixed, N_h increases as the point is moved further away from the hologram plane. Increasing hogel size, on the other hand, reduces N_h contrary to N_θ. Please note that N_h corresponds to the parameter defined as the ray density (number of rays incident in the eye pupil) in the context of ray-based LF displays [11].

3. RESULTS AND DISCUSSION

The theoretical analysis is validated by generating HSs with varying parameters and simulating the images perceived by a human eye viewing the hologram at a certain location, focusing at different depths of the scene. These view simulations are implemented by utilizing numerical Fresnel propagation [13] and simulating the HVS as a sensor (retina) behind a thin circular lens (pupil). The eye can be set to focus at different depths by altering the focal length of the lens, thus modifying its transmittance function $T(s, t)$. The image intensity values $I(u, v)$ are retrieved from the wavefield at the sensor plane as

$$I(u, v) = |\mathcal{F}_l\{T(s, t)\mathcal{F}_d\{O_{\mathrm{HS}}(x, y)\}\}|^2, \tag{8}$$

where $\mathcal{F}_z$ is the Fresnel propagation operation by distance z, l is the sensor-retina distance of the simulated eye and d is its distance from the hologram plane.

We simplify the simulations to a 1D case, which is a valid assumption since the Fresnel kernel is separable. A single point source of light is placed in the recorded scene and the corresponding (1D) HS object field is generated. The process is repeated for several different depths of the point (w.r.t. the hologram plane). The simulated eye is set to focus at different depths to obtain a set of perceived point spread functions (PSFs). These can be utilized to evaluate the perceived sharpness of the point as a function of focus distance. Furthermore, by studying the modulation transfer function (MTF), which is defined as the magnitude of the Fourier transform of the PSF, at various spatial frequency ranges, the scene depth at which the eye is most likely going to focus can be determined for each different point depth. Namely, the focus distance which maximizes the MTF in the given range is estimated as the distance where the human eye would accommodate [11]. Please note that in this paper we do not consider the gradient of the MTF while assessing the accommodation distance, though a narrower MTF peak can be expected to provide a stronger accommodation trigger than a wide and shallow peak. The estimate is then compared to the correct depth of the point source and the shift in accommodation is evaluated, which can be compared to the depth-of-field of the HVS to determine the correctness of the provided accommodation cue.

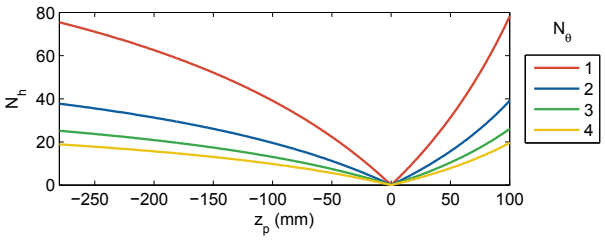

Figure 2: Values of N_h as a function of point distance from the hologram plane for four different N_θ values.

Let us utilize the following experimental setup. A point is placed at a distance z_p away from hologram plane, from 100 mm in front to 280 mm behind the hologram at 10 mm intervals. The

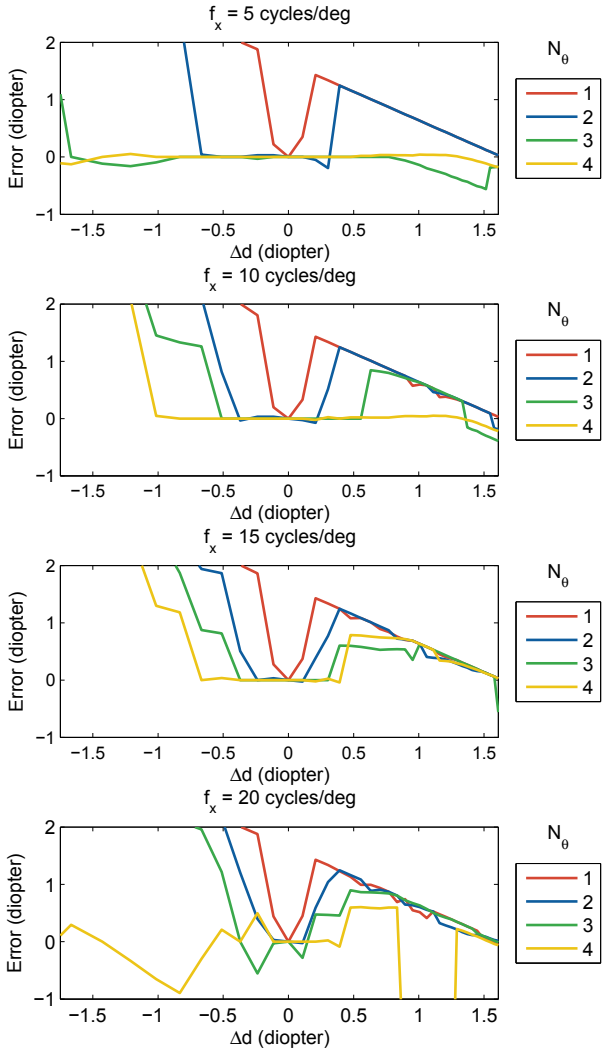

Figure 3: Accommodation shift for regular HSs. Four different spatial frequency values and four different hogel sizes (corresponding to four different N_θ values).

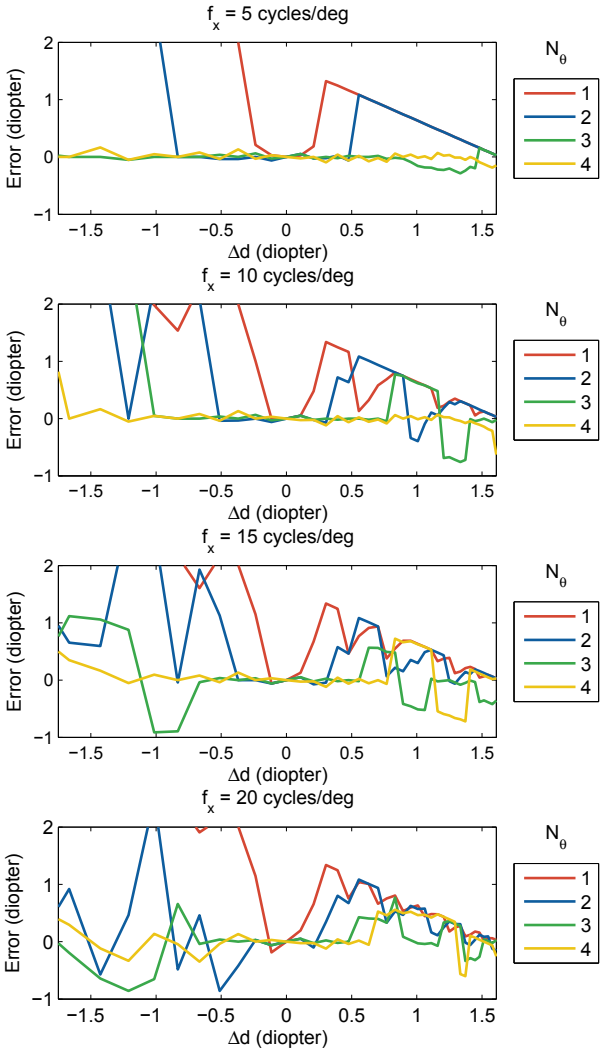

Figure 4: Accommodation shift for FFT HSs. Four different spatial frequency values and four different hogel sizes (corresponding to four different N_θ values).

z-axis is defined to be positive in front of the hologram plane (i.e. between the viewer and the hologram). The point distance with respect to the hologram plane Δd is determined (in diopters) as

$$\Delta d = \frac{1}{z_{eye}} - \frac{1}{z_{eye} - z_p}. \tag{9}$$

Thus, the range of Δd values for the chosen point distances is approximately between -1.6 and 1.6 diopters, where 0 diopters corresponds to the depth of the hologram plane. The simulated eye (D = 5 mm) is placed at distance z_{eye} = 300 mm and is focused at a distance z_f at an oversampled grid with respect to the point locations (by a factor of 4). For each z_p and z_f, the perceived PSF is obtained through the simulation tool. We evaluate the MTF at spatial frequencies 5, 10, 15 and 20 cycles per degree (cpd), as all of these values are reproducible by the chosen HS configurations and are also relevant in terms of the HVS contrast sensitivity [14]. For a given spatial frequency above, we actually evaluate the corresponding response by integrating the set of MTF values in the range of [-2.5, 2.5] cpd. As a result, the MTF values are assessed as a function of z_p and z_f for a discrete set of spatial frequencies. The value of z_f maximizing the MTF is evaluated, and thus, for each z_p an estimate of the likely HVS focus distance $\hat{z}_f$ is obtained. The simulation is replicated for different hogel sizes, corresponding to four different N_θ values from 1 to 4. For the specified point depth range [-280, 100] mm with respect to hologram plane, the corresponding ranges of N_h values for these four cases of N_θ are given in Fig. 2.

The experimental results for regular HSs are presented in Fig. 3. As can be seen from the accommodation shift values, increasing the hogel size preserves the correct accommodation cues further away from the hologram plane. On the other hand, the smallest hogel size (corresponding to slightly denser sampling than Δ_x^{HVS}) has large accommodation shift already at very small depth values, indicating the lack of correct accommodation cues immediately as the objects are moved away from the hologram plane. Furthermore, the general trend regarding the relation between hogel size and accommodation shift is present also in FFT-based HS, as seen in Fig. 4. If the shift in accommodation exceeds a threshold of 0.3 diopters, the accommodation cue is outside the HVS depth-of-field and can be considered to be incorrect [15]. The results for the low spatial frequencies (5 and 10 cpd) suggest that even two views within the pupil extent can provide correct accommodation cues, though at a very limited range (approximately between -0.6 . . . 0.4

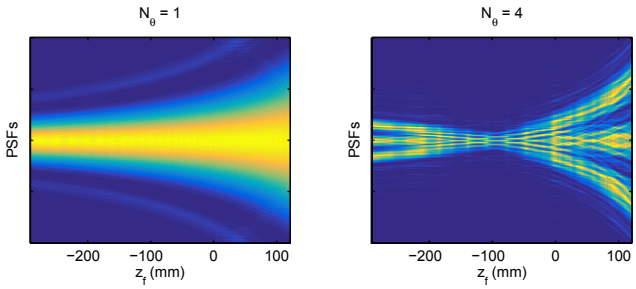

Figure 5: Comparison of perceived PSFs given two different hogel sizes corresponding to N_θ = 1 and N_θ = 4 at different values of z_f for a point at z_p = -100 mm.

diopters away from hologram plane). Significant improvement in terms of accommodation is gained with hogel sizes corresponding to N_θ = 3 and N_θ = 4, which are able to maintain correct cues within (almost) the complete evaluated range of -1.6 . . . 1.6 diopters away from hologram plane. It should be noted, though, that in these cases the perceived spatial resolution is reduced by a factor of 0.4 or 0.3 relative to the Rayleigh diffraction limit, effectively increasing the pixel size by 2.5 and 3.3 times in comparison to what the HVS could resolve. On the other hand, simulation results also indicate that at this range of hogel sizes, the angular resolution is the dominant factor in determining the accommodation response. Thus, increasing N_h has no similar positive effect on the accommodation response as in the case of ray-based LF displays [11], where the angular resolution of lenses (which correspond to hogels in HSs) is quite high compared to HSs.

The noise in the two larger spatial frequency results makes them less conclusive, yet a similar ordering between hogel sizes and general trend as in the lower frequency results can be seen, albeit the accommodation shift exceeds the acceptable threshold already at lower point depth values. Interestingly, in the case of N_θ = 1 and N_θ = 2, the accommodation shifts start to drop for point depths after around 0.4 diopters. Fig. 5 further illustrates different PSFs (vertical slices) for different focused depths given in the horizontal axis. In this figure it can be seen more clearly that for low angular resolution, sharper PSFs are observed at further (more negative) focal depths, which results in failure of producing correct accommodation cues. Further analysis of this effect is required. It is, however, outside the scope of this paper and as such left as future work. For higher N_θ, on the other hand, the behaviour of PSFs justifies the availability of focus cues.

4. CONCLUSIONS

In this paper we have presented analysis of the accuracy of accommodation cues in HSs. Specifically, we have examined the effects of hogel size by utilizing the ray-based interpretation of the hologram and characterizing the number of wavefront segments (corresponding to light rays) incident the viewer's pupil extent. The numerical simulations employing Fresnel propagation to evaluate the perceived accommodation shift have verified the analysis assumptions, thus indicating the presence of correct accommodation cues at a limited range based on the chosen hogel size. Critically, increasing the number of wavefront segments seen from a single hogel by the eye improves such cues. Furthermore, when compared to the minimum resolvable distance between two points of the human eye, a trade-off between accommodation and perceived spatial resolution is seen. That is, if the perceived spatial resolution is maximized by choosing the hogel size according to the Rayleigh criterion, the depth range for correct accommodation cues is extremely limited, and extending this range requires sacrificing spatial resolution. Importantly, we have demonstrated that correct accommodation cues can be provided by HSs. Moreover, HS displays that can provide a limited accommodation response to the viewer can be designed, thus permitting their use in applications where accurate 3D perception particularly regarding accommodation is required, without needlessly increasing the complexity of the model by including model information.

5. REFERENCES

[1] Y. Takaki, "High-density directional display for generating natural three-dimensional images," *Proceedings of the IEEE*, vol. 94, no. 3, pp. 654–663, 2006.

[2] D. Teng, Z. Pang, Y. Zhang, D. Wu, J. Wang, L. Liu, and B. Wang, "Improved spatiotemporal-multiplexing supermultiview display based on planar aligned OLED microdisplays," *Opt. Express*, vol. 23, no. 17, pp. 21549–21564, Aug 2015.

[3] F. Yaraş, H. Kang, and L. Onural, "Real-time phase-only color holographic video display system using LED illumination," *Applied Optics*, vol. 48, no. 34, pp. H48–H53, 2009.

[4] H. Zhang, N. Collings, J. Chen, B. A. Crossland, D. Chu, and J. Xie, "Full parallax three-dimensional display with occlusion effect using computer generated hologram," *Optical Engineering*, vol. 50, no. 7, pp. 074003–074003–5, 2011.

[5] M. Yamaguchi, H. Hoshino, T. Honda, and N. Ohyama, "Phase-added stereogram: calculation of hologram using computer graphics technique," in *Proc. SPIE*, 1993, vol. 1914, pp. 25–31.

[6] Q. Y. J. Smithwick, J. Barabas, D. E. Smalley, and V. M. Bove, Jr., "Interactive holographic stereograms with accommodation cues," in *Proc. SPIE*, 2010, vol. 7619, p. 761903.

[7] J. T. McCrickerd and N. George, "Holographic stereogram from sequential component photographs," *Applied Physics Letters*, vol. 12, no. 1, pp. 10–12, 1968.

[8] M. W. Halle, "Holographic stereograms as discrete imaging systems," in *Proc. SPIE*, 1994, vol. 2176, pp. 73–84.

[9] P. St. Hilaire, "Modulation transfer function and optimum sampling of holographic stereograms," *Applied Optics*, vol. 33, no. 5, pp. 768–774, Feb. 1994.

[10] K. Wakunami and M. Yamaguchi, "Calculation for computer generated hologram using ray-sampling plane," *Opt. Express*, vol. 19, no. 10, pp. 9086–9101, May 2011.

[11] 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.

[12] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 2nd edition, 1996.

[13] D. P. Kelly, "Numerical calculation of the Fresnel transform," *Journal of the Optical Society of America A*, vol. 31, no. 4, pp. 755–764, Apr. 2014.

[14] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *The Journal of Physiology*, vol. 197, no. 3, pp. 551–566, 1968.

[15] 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.