https://doi.org/10.2352/EI.2023.35.2.SDA-391

Observed optical resolution of light field display: Empirical and simulation results

Yuta Miyanishi, Qi Zhang, Erdem Sahin, and Atanas Gotchev; Tampere University; Tampere, Finland

Abstract

Light field (LF) displays are a promising 3D display technology to mitigate the vergence-accommodation mismatch. In this study, we empirically evaluated the optical resolution of a near-eye LF display test bed by capturing rendered test images and compared it to simulation results based on a previously developed computational model. The LF display prototype employs a time-multiplexing technique and achieves a high angular resolution of 6×6 viewpoints in the eyebox of a 2.8-mm square. The test image was rendered at various depths ranging 0–3 diopters, and the displayed images were captured by a camera for analysis of the optical resolution the display achieved at varying focusing depths. Both the simulation and measurement results indicated that the display correctly provides the focusing effects to the camera, although errors up to 0.5 diopters were found in the measurement. The measured responses were much more limited than the simulated responses, and on- and off-axis aberrations did not fully explain that difference, suggesting large effects of potential errors in the optical alignment and the LF image processing pipeline. An additional simulation on a hypothetical model indicated that larger viewpoint subapertures may be important for providing better optical resolution with LF displays.

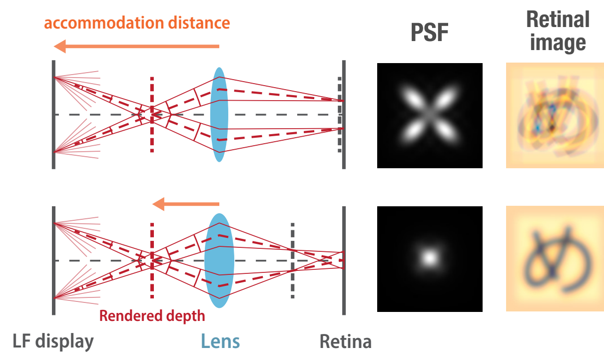

Figure 1. *Basic mechanism of reproducing focus information by LF display. In both the upper and lower diagrams, the LF display is rendering a point at a depth closer than the display surface with four rays (only two are shown). In the upper diagram, the eye is focusing (accommodates) at the display surface. In the lower diagram, the eye is focusing at the rendered depth. PSF refers to the point spread function of the rendered point, and Retinal image indicates examples of a letter's optical image on the retina.*

Introduction

Conventional displays including stereoscopic displays cannot reproduce the focus information of a 3D scene, which means the image an observer's eye receives gets sharpest when and only when the display surface is focused (accommodated) at. Such lack of the correct focus information gives rise to the well-known vergence-accommodation mismatch, which causes visual discomfort in viewers and disturbs performances in some perceptual tasks [1, 2].

Light field (LF) displays are promising solution to the vergence-accommodation mismatch by reproducing the natural focusing effects in a displayed 3D scene. *Light field* is a concept of representing light from a scene as rays, each of which are specified by its position and direction. LF displays are meant to reproduce a light field by controlling the intensity and color of the light depending not only on its emanating position but also on the emanating angle. Importantly, an LF display may provide the natural focusing effects to a viewer if it achieves a high angular resolution so that two or more beams, which conceptually correspond to rays, are projected into the viewer's pupil.

Fig. 1 illustrates the mechanism of providing the natural focusing effects with an LF display. When the eye accommodates at the display (upper diagram), each beam generates an optically focused image on the retina. However, the individual point images are placed with lateral shifts because of the beams' entering angles into the pupil. The final retinal image of the point is a superposition of the individual point images, which appears like the PSF (point spread function) shown to the right of the diagram. On the other hand, if the eye accommodates at the rendered depth (lower diagram), the individual point images are optically defocused, but they are perfectly overlapped on the retina, which results in a compact point image after the superposition. The latter is expected to be recognized "more focused" or "having better image quality" than the former. Because this mechanism involves optical defocusing and the superposition of individual images, the observed resolution that LF displays achieve is limited not only by pixel resolution but also by optical restriction.

To design practical LF displays, various parameters have to be determined so that the display satisfies viewers' demand on the reproduced focusing effects and the observed resolution. The reproduced focusing effects have been examined in numerous studies [3, 4, 5, 6, 7], but there has been much less effort on studying the quality of observed images. This study was aimed at evaluating the observed optical resolution of an LF display quantitatively and systematically. For that purpose, we measured the spatial frequency responses (SFRs) in an actual LF display test bed for varying image depths and focusing depths, and the measured SFRs were compared to simulated SFRs based on a model of the display. The comparison indicated technical and practical limitation in the actual LF display and its possible causes, and a key factor for improving the observed optical resolution was suggested.

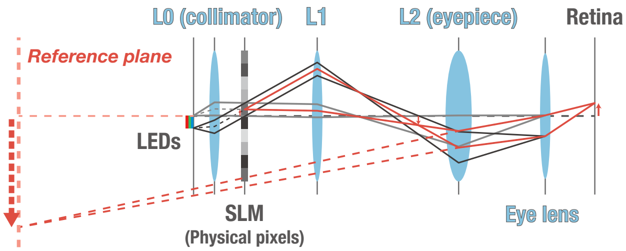

Figure 2. *An optically equivalent model of the LF display test bed. Light beam(s) for a single viewpoint and a single pixel is shown. Black and gray lines indicate marginal light from the light source responsible for the viewpoint. Red lines indicate the light for imaging the pixel on the retina.*

LF display test bed

We measured and simulated SFRs on an actual LF display test bed (made by CREAL SA, Switzerland) and its optical model. Fig. 2 shows a simplified but optically equivalent model of the display, which includes the display's optical structure and a viewer's eye. The spatial light modulator (SLM), which works as physical pixels, is backlit by LEDs, each of which corresponds to one of the viewpoints. For each viewpoint, the image of the viewpoint's corresponding LED (light source) in the eye lens plane works as an aperture that images the pixels on the retina. We call this aperture as a *viewpoint subaperture* or simply a *subaperture*. The physical pixels are imaged at approximately 1 m away from the eye ("Reference plane" in the figure).

The display generates 32 viewpoints in an eyebox of a 2.8-mm square. The viewpoints are on a 6×6 grid except its four corners. Each viewpoint subaperture is a circle of 0.3-mm diameter. The LF frame rate of the display, i.e., the number of LF frames rendered in every second, is 30 Hz. In this display, LF rendering is based on time multiplexing; the 32 images for the viewpoints are displayed on the physical pixel turn by turn during each LF frame, and the light sources corresponding to the viewpoints are synchronously turned on and off to backlight the pixels. This time-multiplex technique enables the display to avoid the trade-off between the angular resolution and the observed pixel resolution [8]. Thus, the observed pixel resolution is equal to the physical pixel resolution, which is 912×1140 (horizontal × vertical). The field of view that the pixels subtend is 46°×29° (horizontal × vertical).

LF display modeling and image simulation

In this study, the LF display test bed was modeled and the observed images on a camera sensor were simulated based on the wave-optics-based framework that we have recently proposed [7]. Fig. 3 shows the model and the process of simulating and evaluating observed images on the retina of rendered 3D points. Since we compared the simulated images to measured images that were captured with a camera, we modified the model so that the eye lens and retina represented the camera's lens and sensor. Specifically, we adjusted the lens-to-retina distance to that of the camera we used in the measurement, and the lens was assumed to be aberration free based on the assumption that only near on-axis light is responsible for imaging the rendered LF on the sensor and thus the aberrations of the camera lens were negligible in this case.

In the model, an LF display is represented by a conceptual plane, which is defined as the *optical reference plane*, from which directed light beams emanate. The width of beams is primarily determined by the angular resolution of the display, and each beam forms an elemental point image on the retina. The effective aperture for the beam in the plane of the eye lens is generally defined as the viewpoint subaperture. The retinal image of the rendered point (Retinal PSF in Fig. 3) is a superposition of the elemental point images, which vary with the focusing depth of the eye lens for a given depth of the rendered point. The retinal PSF was analyzed in its frequency domain (OTF: optical transfer function), specifically in the one-dimensional horizontal slice of its modulus (MTF: modulation transfer function), which is called also as an SFR in the context of measurement of imaging systems.

This model only includes two critical parameters of the LF display. One is the optical distance to the reference plane from the eye lens, which was set to 0.99 m, and the other is the size and positions of the viewpoint subapertures, which were set as described as in the previous section.

Measurement of SFRs

The measurement of SFR was based on the edge SFR method specified in ISO12233 [9, 10]. We prepared the viewpoint images for LF frames which contain full-contrast edges slanted by 5° at the depths of 0.1, 1.0, 2.0, and 3.0 diopters with Blender. The viewpoint images of the LF frames were confirmed that they had correct pixel disparities between viewpoints. When showing each LF frame on the display, the display driver accepts these image files, processes them, and send the frame to the display.

A camera (Grasshopper3 USB3 GS3-U3-51S5C-C, Teledyne FLIR LLC) equipped with a focus-tunable lens (ELM-25-2.8-18-C, Optotune AG) was used to capture the displayed LF frames. For computation of spatial frequency in *cycles per degree* (cpd), we captured a real visual target (TE100 test chart) with the camera, and the parameter of *pixels per degree* was calculated from the camera-to-target distance, the target's size, and the number of the pixels that the target filled in a captured image. The measurement SFRs in cpd were obtained by multiplying SFRs in *cycles per pixel* that the edge SFR method provides by the pixels-per-degree value.

Results

We obtained SFRs of the LF display from the measurement and the simulation. The SFRs were measured and simulated for the image depths of 0.1, 1.0, 2.0, and 3.0 diopters with varying focusing depths of the camera, the range of which covered the image depth and the optical reference plane depth.

Fig. 4 shows the obtained SFRs. In both the measured and simulated responses and for all rendered image depths, the best responses were observed when the camera was focusing at around the image depths (the red arrows in the figure) rather than the depth of the display's optical reference plane (the green arrows). For better visibility of focusing effects in the SFRs, cut-off frequencies for the gain of 0.1 are plotted in Fig. 5. The peak of each curve of the cut-off frequencies indicates the focusing depth at which the best response was obtained. In the curves from the simulation, the peaks are always exactly at the corresponding image depths, which means the best responses were predicted to be obtained when the camera would be focusing exactly at the image depths. In the measured responses, the peak in each curve appears

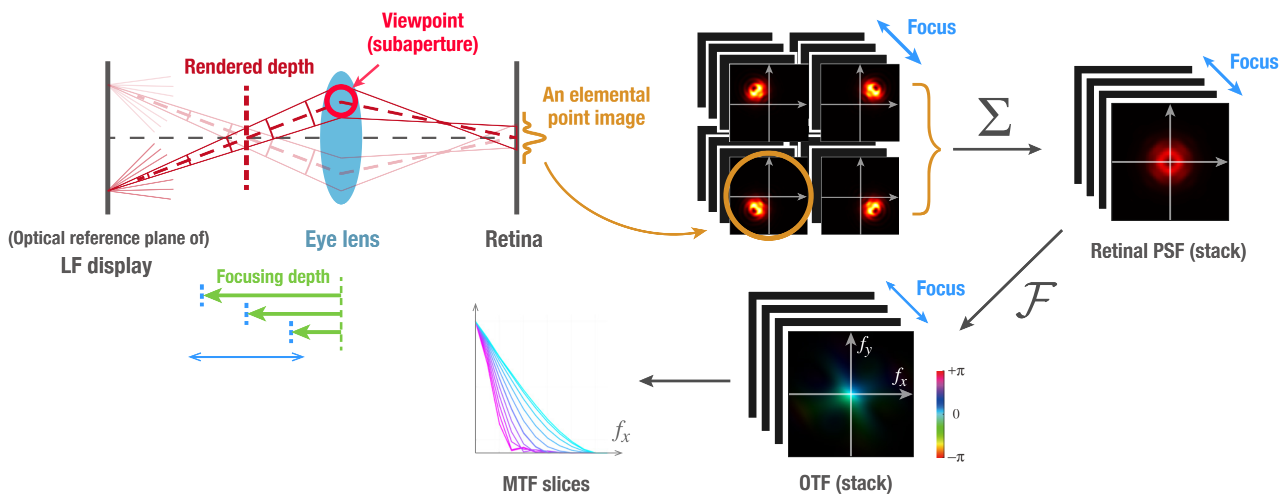

Figure 3. The model of an LF display and the process of simulation and evaluation of rendered retinal images. An LF display that renders a point with four beams is modeled, but only two beams are shown in the diagram for simplicity.

to be around the image depth, although they tend to be closer to the viewer by up to about 0.5 diopters especially for the images rendered at 0.1 and 1.0 diopter.

The simulation results mean that the display was predicted to provide "perfect" focusing effects to the camera in the simulation, i.e., the best SFR would be obtained when the camera is focusing exactly at the rendered image depth. On the other hand, the measured SFRs showed that the display reproduces the focusing effects on SFRs about the rendered images so that the best SFRs were observed when the camera was focusing at around the image depths, although there were the errors between the best focusing depths and the image depths.

The measured responses were almost always worse than the simulated responses for all rendered image depths as it is clearly visible in the differences between the 3D surfaces and the colored curves in Fig. 4 and between the solid and dashed curves in Fig. 5. The peaks of the cut-off frequencies in the measured responses were often less than half of these in the simulated responses. This means that the optical performance of the actual display was found much more limited than predicted from the simulation.

Discussion

Simulated and measured responses

The measurement indicated that the optical performance of the display is significantly lower than the simulated performance. The simulation model only included the optical distance to the reference plane as well as the profile of viewpoint subapertures, and there are several possible factors that explain the difference between the measurement and simulation results. One is optical aberrations in the display's optics, which were neglected in the simulation.

To test whether the optical aberrations are the main reason of the limited optical performance, we simulated SFRs for individual viewpoints on the display's full model in Zemax, which consisted of the models of the actual lenses in the display and thus represents the optical aberrations in them. The results are shown in Fig. 6. Two representative SFRs for a single viewpoint were simulated on the aberrated full model and plotted with the solid and dashed red lines. The curve labeled "Aberrated: center (Zemax)" represents the response for the rays propagating through a viewpoint in the center of the pupil, which reflects almost only on-axis aberrations. The other curve labeled "Aberrated: edge (Zemax)" represents the response for rays responsible for a viewpoint at the edge of the 6×6 grid, which may suffer more from aberrations because of off-axis propagation through the lenses. For reference, SFRs on an ideal (non-aberrated) model were simulated by the method referred previously in this paper (Ideal (wavefront)) and in Zemax (Ideal: center (Zemax)).

The simulation results showed very small differences in the responses between the on-axis and the off-axis aberrated models and also between the ideal model and the aberrated model. This means that including the on- and off-axis optical aberrations scarcely affected the simulated SFRs; therefore, the optical aberrations in the display do not fully explain the difference observed in the measured and simulated SFRs. However, the effects of including the optical aberrations were evaluated only for single viewpoints. Because the SFR does not preserve all information of a point spread function (point image), the effects of the optical aberrations on the single-viewpoint SFRs cannot strictly be extrapolated to final retinal images, which are superposition of the point images.

Other factors may explain the difference between the measured and the simulated responses. One is potential errors in the optical alignment in the display's optics, which we cannot directly and quantitatively measure. Artifacts in the image processing engine may also have affected the measured optical performance. For example, imperfection in antialiasing or synchronization failure in time multiplexing should damage the final image quality, although we again cannot quantify them.

Limitation in optical performance

We simulated and measured SFRs in the LF display, and the measured responses were almost always much more limited

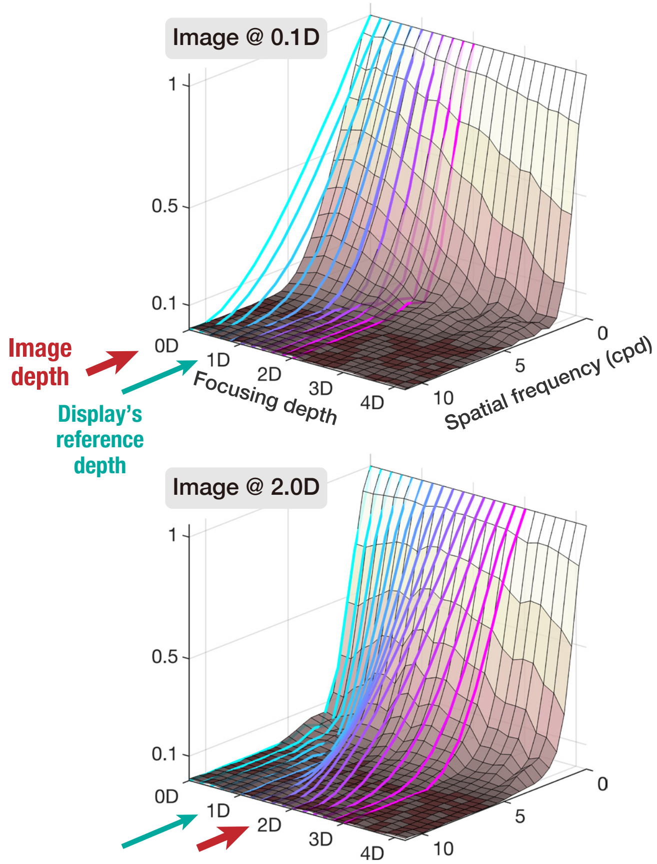

Figure 4. *Measured and simulated spatial frequency responses (SFRs) about images rendered at 0.1, 1.0, 2.0, and 3.0 diopters. 3D surfaces and colored curves indicate the measured and simulated SFRs, respectively. The SFRs, which are functions of spatial frequency, are visually stacked over the camera's focusing depth, and the image depth and the display's optical reference plane depth are respectively shown by red and green arrows.*

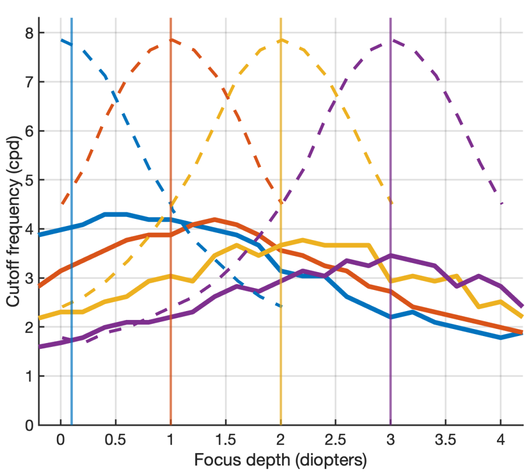

Figure 5. *Cut-off frequencies for the gain of 0.1 in the SFRs as functions of camera's focusing depth. Blue, red, yellow, and purple curves are the cut-off frequencies for the image depths of 0.1, 1.0, 2.0, and 3.0 diopters, respectively; the image depths are also shown as vertical lines with corresponding colors. Solid and dashed curves correspondingly indicate the measured and the simulated results.*

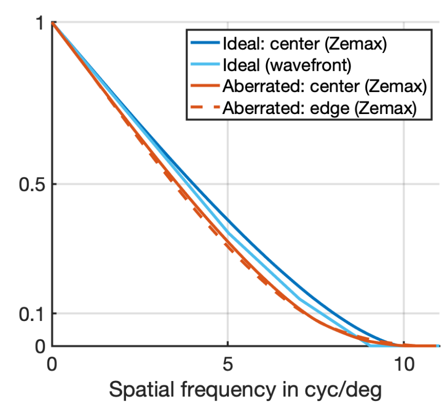

Figure 6. *Simulated SFRs for a single viewpoint of the display. The curves colored with dark and light blue represent the SFRs in the ideal models calculated in Zemax and in the proposed method, respectively. The red solid and dashed lines show the SFRs in the display's full model, which includes optical aberrations in the lenses. The solid and dashed curve respectively indicate the SFR for a viewpoint at the center of the pupil and that for a viewpoint at the edge of the eyebox, which is assumed to suffer more from off-axis aberrations.*

than the simulated responses. Even at the best focused states in the simulation, the responses indicated quite limited optical performance of the display. For example, even the best-focused response in the simulated SFRs (see Fig. 4) reaches zero around the spatial frequency of 10 cpd, which means that displayed visual features finer than 10 cpd would not be delivered to the viewers.

In general, the optical resolution that an optical system achieves is directly linked to compactness of point images in it. A point image's compactness is primarily determined from the system's aperture size when the other factors (focal length, wavelength, and optical aberrations) are fixed. Therefore, the elemental point images (see Fig. 3) are greatly affected from the size of the subapertures. We tested if the limited optical resolution simulated (and observed) for the display is mainly because of the display's small subapertures.

Fig. 7 shows the simulated SFRs in the original display model and in the hypothetical display model with expanded subapertures, in which the subapertures were enlarged so that they touched each other in the eye lens plane. The SFRs simulated for the expanded subapertures are clearly better than the original SFRs at the best-focused states, yet the defocused SFRs (for example, responses to the image at 0.1 diopters when the camera's focusing depth was 2.0 diopters) are equally limited for the original display model and the expanded subaperture model.

The effects of expanding the subaperture size is more visible in cut-off frequency. Fig. 8 shows the cut-off frequencies in the same manner as in Fig. 5, but only these of the simulated SFRs on the original display model and the expanded subaperture model are plotted. The cut-off frequencies both for the original and the expanded subaperture model indicate their peaks at each corresponding for all rendered image depths, which means that expanding the subaperture size did not affect the focusing effects that the display was predicted to provide. The peaks of the cut-off frequencies for the expanded subaperture model were approximately twice more than these for the original display model. This indicates that the optical resolution in the display and the camera at the best-focused states was predicted to improve significantly as expanding the subaperture size.

A simulation study on a general LF display model suggested that few viewpoints, 2×2 viewpoints in a 3-mm pupil for example, may be enough just to provide the focusing effects on human viewers' eyes [7]. Considering that, larger subapertures may be more important than having more viewpoints in designing practical LF displays to ensure satisfactory optical resolution provided by the displays.

Conclusion

In the current study, we measured spatial frequency responses in an actual LF display test bed and compared them to simulated responses in an optical model of the display. The simulation predicted that the display would provide perfect focusing effects to the camera, and the measured responses also indicated that the display did provide the focusing effects, although errors up to 0.5 diopters were observed.

The simulated SFRs at the best-focused states indicated a severely limited optical resolution that the display would provide, and the measured responses were even more limited than them. On- and off-axis aberrations did not fully explain the difference between the simulated and measured responses, thus the degra-

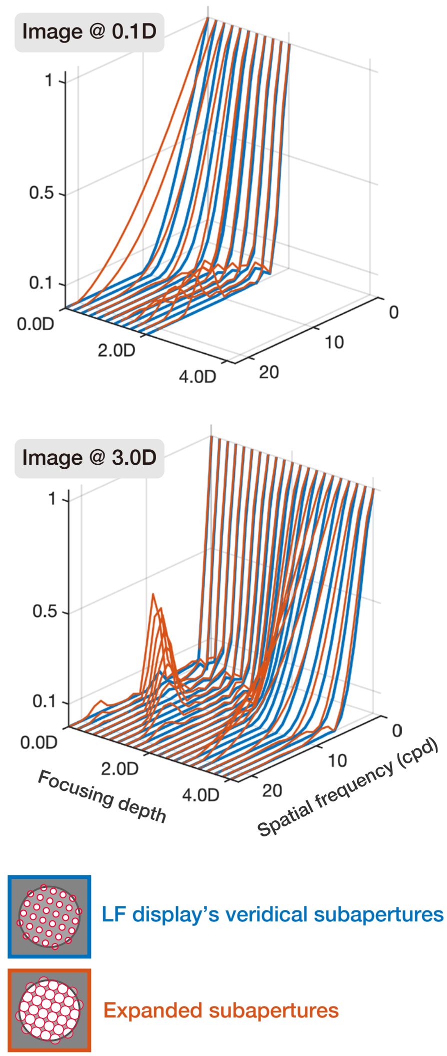

Figure 7. *Simulated SFRs for the LF display model (blue curves) and for a hypothetical display model with expanded subapertures (red curves). As examples, the responses for images rendered at 0.1 and 3.0 diopters are plotted in the same manner as in Fig. 4.*

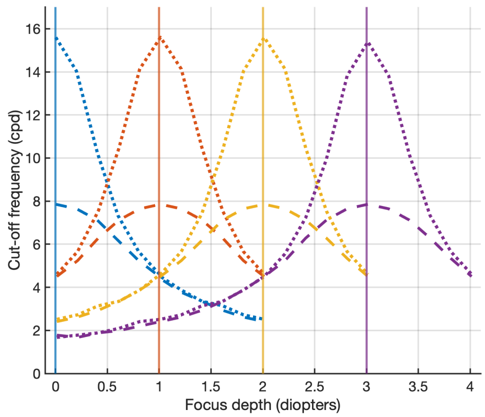

Figure 8. *Cut-off frequencies for the gain of 0.1 in the SFRs simulated in the original display model (dashed lines) and the hypothetical display model with expanded subapertures (dotted lines).*

dation might be due to other reasons such as misalignment in the optics and errors in the LF image processing pipeline, yet we are not able to quantify these possible factors.

An additional simulation on a hypothetical display model with expanded subapertures suggested that the primary factor of the limited optical resolution may be strong diffraction caused by small subapertures. Therefore, large subapertures may be important in designing practical LF displays to ensure satisfactory optical resolution, and perhaps it should be more prioritized than having more viewpoints.

Further empirical studies are needed to extend the current findings. Human viewers' accommodation responses to LF images should be measured to test whether "few viewpoints" are enough for mitigating the vergence-accommodation mismatch, and the observed optical resolution should also be empirically evaluated under effects of human eyes' aberrations and the characteristics of visual functions, such as the contrast sensitivity.

Acknowledgments

Our colleague Ugur Gudelek developed and provided the LF image generation script used in the current study. We thank Dr. Jonathan Masson and Dr. Alexander Kvasov from CREAL SA for their generous decision on sharing technical details to us and for active and friendly discussions. The project is supported by the Academy of Finland research project "Modeling and Visualization of Perceivable Light Fields", decision number 325530. The author Qi Zhang receives funding from the European Union's Horizon 2020 Research and Innovation Programme AERO-TRAIN under grant agreement No. 953454.

References

[1] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.*, vol. 8, no. 3, pp. 1–30, 2008.

[2] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations," *Displays*, vol. 29, pp. 106–116, 2008.

[3] Y. Miyanishi, E. Sahin, and A. Gotchev, "Providing focus cues by light field displays: A review," European Light Field Imaging Workshop, Jun. 2019.

[4] Y. Kajiki, H. Yoshikawa, and T. Honda, "Hologramlike video images by 45-view stereoscopic display," in *SPIE Electronic Imaging*, S. S. Fisher, J. O. Merritt, and M. T. Bolas, Eds., vol. 3012. International Society for Optics and Photonics, May 1997, p. 154.

[5] T. Sudo, H. Morishima, T. Osaka, and N. Taniguchi, "3D display using intersection of light beams," in *SPIE Electronic Imaging*, J. O. Merritt, S. A. Benton, A. J. Woods, and M. T. Bolas, Eds., vol. 3957. San Jose, CA, USA: International Society for Optics and Photonics, May 2000, p. 215.

[6] H. Huang and H. Hua, "Systematic characterization and optimization of 3D light field displays," *Opt. Express*, vol. 25, no. 16, p. 18508, Aug. 2017.

[7] Y. Miyanishi, E. Sahin, and A. Gotchev, "Optical modelling of accommodative light field display system and prediction of human eye responses," *Opt. Express*, vol. 30, no. 21, Sep. 2022.

[8] T. Sluka, "Near-eye sequential light-field projector with correct monocular depth cues," Sep. 26 2019, US Patent App. 16/348,768.

[9] "ISO 12233:2017 Photography — Electronic still picture imaging — Resolution and spatial frequency responses," 2017.

[10] P. Burns, "Burns Digital Imaging," http://burnsdigitalimaging.com, Accessed 2023-01-29.

Author Biography

Yuta Miyanishi received his B.E. degree in computer science and M.S. degree in human centered science and biomedical engineering from Tokyo Institute of Technology in 2016 and 2018, and he is currently pursuing his doctoral degree at Tampere University in Finland. After his study on stereopsis in a psychophysical approach in Tokyo Tech, he is now involved in theoretical and empirical research on light field displays and human observers.

Qi Zhang received his B.E. degree in information engineering and M.S. degree in optical engineering from Zhejiang University in 2011 and 2014, and he is now a Ph.D. student at Tampere University in Finland. His research focuses on the implementation of 3D displays.

Erdem Sahin received the Ph.D. degree from the Electrical and Electronics Engineering Department, Bilkent University in 2013. In 2014, he joined the 3D Media Group in Faculty of Information Technology and Communication Sciences at Tampere University, as Marie Curie experienced Researcher, where he has been Senior Research Fellow since 2019. His current research interests include development of computational light field and holographic imaging algorithms and methods for next-generation 3D cameras and displays.

Atanas Gotchev received the M.Sc. degrees in radio and television engineering (1990) and applied mathematics (1992), the Ph.D. degree in telecommunications from the Technical University of Sofia (1996), and the D.Sc.(Tech.) degree in information technologies from the Tampere University of Technology (2003). He is currently Professor of Signal Processing with the Tampere University. His recent work concentrates on developing methods for multi-sensor 3D scene capture, and light field imaging and display.