International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2022)
J.-M. Normand and H. Uchiyama (Editors)

A Rendering Method of Microdisplay Image to Expand Pupil Movable Region without Artifacts for Lenslet Array Near-Eye Displays

Bi Ye[1], Yuichiro Fujimoto[1], Taishi Sawabe[1], Masayuki Kanbara[1], Geert Lugtenberg[1] and Hirokazu Kato[1]

[1]Graduate School of Science and Technology, Nara Institute of Science and Technology, Japan

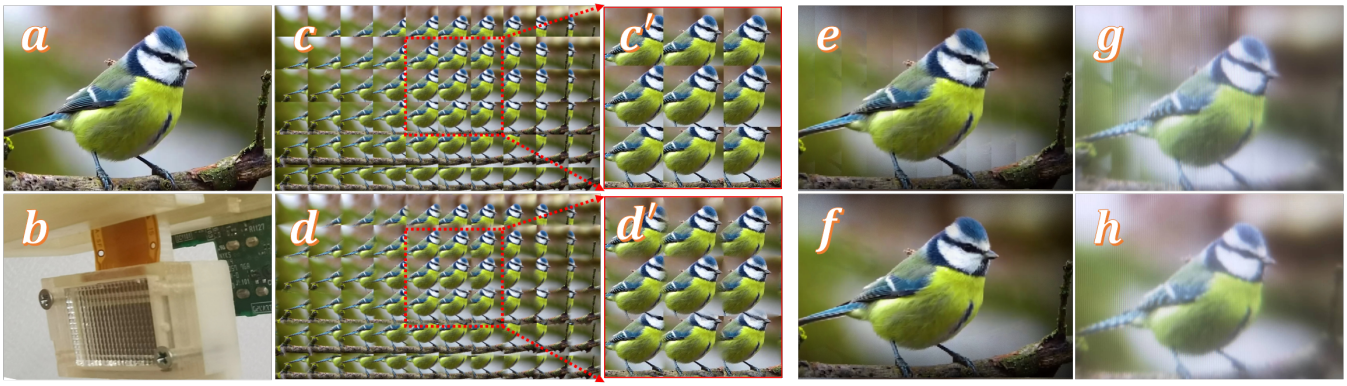

Figure 1: *(a) The input image. (b) The optical part of the prototype. The prototype comprises a microdisplay in front of a lenslet array (LA). We used a 3D-printed glasses frame. The microdisplay size was 15.36 × 8.64 mm, and the resolution was 1,280 × 720. The driver comes from a Sony HMZ-T1 personal media viewer whose magnifying eyepieces were removed. (c) Microdisplay image generated by the traditional rendering method. The edges of each elemental image are clearly seen. (d) Microdisplay image generated by rendering method we newly proposed. (c') and (d') are the enlarged images to better contrast the difference in image details. The edges of each elemental image can clearly be seen in c'. The edges of each part cannot be distinguished in d'. (e) The simulated retina image with the traditional rendering method with transition distance (Td) is 1.2 mm. (f) The simulated retina image with our rendering method with Td is 1.2 mm. (g) The observed image was taken with a Sony RX1 camera with an aperture of 8 mm and a focal length of 35 mm and Td is 1.2 mm under the condition of the traditional rendering method. (h) The observed image with the same parameters as the camera is in the condition of our rendering method.*

Abstract

Near-eye displays (NEDs) with lenslet array (LA) are a technological advancement that generates a virtual image in the observer's field of view (FOV). Although this technology is useful for designing lightweight NEDs, undesirable artifacts (i.e., cross-talk) occur when the user's pupil becomes larger than the pupil practical movable region (PPMR) or moves out of the PPMR. We proposed a rendering method for microdisplay images that takes pupil size into account and included the idea of pupil margin in the ray tracing process. Ray lights emitted by one microdisplay pixel (MP) enter the pupil and pupil margin area after passing through a number of lenses. Each lens at the MP corresponds to one virtual pixel (VP) on the virtual image plane. The weight of each VP is the intersection area between the ray light column and the pupil and pupil margin divided by the sum of intersecting spaces between all the ray light columns generated by the MP and the pupil and pupil margin. The value of each MP is determined by the number of VPs and the related weight. Through retina image simulation studies, we confirmed that the proposed rendering approach substantially enlarges PPMR to accommodate large pupil diameters and wide transition distances while reducing eye relief to an optimal (sunglasses) distance.

CCS Concepts

• ***Computing methodologies*** → *Microdisplay image processing; Rendering method; Ray tracing;*

DOI: 10.2312/egve.20221284

1. Introduction

Head-mounted display (HMD) is a critical key component for Virtual Reality (VR) technologies. The main platform device for VR is near-eye displays (NEDs). VR NEDs are tasked with producing the immersive environment for the user [Sut68, XGS*19], thereby providing a new user experience for various applications [PL22]. For traditional VR devices that use a single lens [Viv, Ocu], it is impractical to manufacture a focal length that is significantly less than the width of the lens, making the required space between the display and the lens difficult to compress [LL13, YFU*22], which results in the relatively huge size of HMDs. As wearable devices, bulk lenses and mirrors (which are typically used in NED optics) impose limitations on the achievable weight and form factor. It is difficult to realize wearable and aesthetically plausible devices comfortably [PL22, KS13, JBM*17]. The development trend of VR is toward miniaturization [RSAA20, LL13, BJCL21].

Lanman and Luebke [LL13] demonstrated a NED device that uses a microdisplay in front of a lenslet array (LA). It provides a means to achieve a thin and lightweight structure. Another study, carried out by Bang [BJCL21], presented a thin and flat VR prototype using a Fresnel LA with a polarization-based optical folding technique. Although the LA technique can produce a miniaturized VR device, the eye box (i.e., the area where users observe the full extent of the virtual image) is not large enough [LL13, BJCL21, RSAA20, HKL*08, JYB*21]. When users turn their eyes to observe the peripheral content of the screen, pupils would move out of the eye box and cross-talk appears resulting in degraded image quality. If the cross-talk in the image is tiny, the brain cannot perceive this kind of noise [BS07]. When eyes continue to move out of the eye box, it makes cross-talk dark enough to be perceived. Therefore, in the process of designing an optical system, pupil size and human visual perception should be considered [YFU*22, CBC*20].

In our previous research [YFU*22], to quantify the effect of human visual perception on LA-NED cross-talk, the peak signal-to-noise ratio (PSNR) was adopted. PSNR ranges from infinity to zero and indicates the quality of images or videos from best to worst [Sal08]. The pixels in the target image or video are considered signals, the pixels in cross-talk are considered noise. When the noise in images or videos is sufficiently large, the observed image quality decreases. Typical values of PSNR for lossy images and compressed videos are 30–50 dB [CLN20]. Different types of noise correspond to different PSNR thresholds [Wel99, LJY*19, MWH*13, LJM*16, KWNM20]. The relationship between pupil size, human visual perception and eye box has been discussed. We first defined a new metric Pupil practical movable region (PPMR) and verified it to be more suitable than the eye box for evaluating the area in which no cross-talk can be perceived. In a prototype that considers pupil size and human visual perception, when a pupil moves out of the PPMR, cross-talk is perceived [YFU*22].

We want to achieve high design flexibility for NED (e.g., to design a sunglasses-like prototype), in which case eye relief should be short. When eye relief gets shorter, PPMR becomes smaller [YFU*22]. Users most often perceive cross-talk while rotating their eyes out of the PPMR to focus on the peripheral content of the screen. Cross-talk that occurs when a pupil leaves the PPMR can be eliminated by creating microdisplay images based on eye gaze. However, if the eye relief keeps getting shorter while eliminating PPMR, the prototype cannot be used even based on eye tracking technology. Therefore, expanding PPMR is required for a prototype with LA. Due to the limitation of the optical structure and rendering method, it is hard to reduce eye relief to be short enough to fit the same distance (approximately 10 mm [Wil15, Bro91]) as sunglasses [Yan18, CHB19]. The rendering method proposed in this paper makes PPMR exists even if eye relief gets 10 mm.

2. Related work

2.1. Techniques for extending the eye box in NEDs

The eye box is a crucial parameter in NEDs. The pupil should be located within this area [LL13, RSAA20]. In LA-NEDs, cross-talk occurs if a pupil is outside the eye box, making the image look doubled or ghosted. The eye box in an LA-NED or holographic NED is small, compared to a normal HMD [Viv, Ocu], owing to its optics. To overcome the small eye box, a novel retinal projection NED was proposed that can independently activate two separate groups of viewpoints using the polarization-multiplexing technique. The system is constructed by polarization-dependent elements and multiplexed combiners to provide an extended eye box with a compact form factor [YCML20]. Another method to expand the eye-box of the retinal projection display [JYB*21] constructs independent viewpoints that can implement retinal projection with a wide depth of field. A NED configuration with an extended eye box was proposed [KP18] that used the transparent optical element instead of the bulk concave mirror, to facilitate optical see-through views. It also enables the multiplexing of multiple concave mirrors with shifted focal spots such that the eye box can be extended without double image problems. Another NED with an enlarged eye box due to a 2D beam deflector system was proposed [LZZ*20]. Their prototype provides a simple way to enlarge the eye box due to the advantages of the adopted elements. Nonetheless, all these methods have a narrow FOV. There is a problem in which the eye-box is still limited because it is focused at one point. This may cause the image to be blank, or cross-talk to occur if the eye deviates from the eye box by displacement due to eye rotation.

An approach by [YCML20] increased the number of viewpoints by simultaneously recording multiple angles of the lens on the holographic optical element. In this case, there is a problem with each viewpoint image overlapping or blanking while the user changes the viewpoint. Also, the eye tracker is essential, and the images should be updated according to the viewpoint. The major technique that can expand the pseudo-eye box in the whole FOV is using eye-gaze tracking for dynamic image switching. A novel holographic "Retinal 3D" prototype [JBM*17] used this technique by providing a “dynamic eye-box”, which can be a breakthrough that overcomes the drawbacks of retinal projection-type displays. For the Foveated prototype [KJS*19] with a wide FOV peripheral display, the eye box limitation whose nodal point is translated to follow the viewer’s pupil during eye movements using a travelling optical element. The dynamic position is driven by gaze tracking. However, all methods have difficulty obtaining an ideal eye relief of 10 mm.

2.2. Rendering method of microdisplay image in NEDs

Integral imaging was first proposed by Lippmann [Gab08], which made an outstanding contribution to digital imaging technology. Reconstructing a scene of integral imaging systems is used in VR. To reconstruct a scene, elemental images are displayed on a display device and the rays pass through a LA to reproduce the scene in space. Integral imaging has been widely researched for various applications due to its many advantages such as continuous view angles and virtual display [WXHJ15, YCYW18].

To improve the depth of field, which is capable of expanding the accommodation range for HMD, an algorithm for optimizing the structure parameters of the hybrid computational NED is proposed [ZMX*19]. Another rendering for a NED presents imagery with a near-accurate per-pixel focus across an extended volume allowing the viewer to accommodate freely across this entire depth range [RWBF18]. To solve the problem of viewing distances larger than the optimal viewing distances, a dual-view integral imaging 3D display using polarized glasses is proposed [WLD*18]. Two kinds of elemental images, which are captured from two different scenes, are alternately arranged on the display panel. Two different images are reconstructed in the viewing zone. To expand the eye box in holographic displays, two novel methods are proposed. One is a practical pupil-shifting rendering method of eye-box expansion [JBLL18]. Another one is computational imaging of optical see-through holographic near-eye display for extending eye box [CCPG19]. Our rendering method in this paper considers pupil size and introduces the concept of pupil margin, which is different from the methods used in LA-NEDs. We made a prototype using this rendering method that can tolerate a large pupil size and wide transition distance (Td) in a robust way.

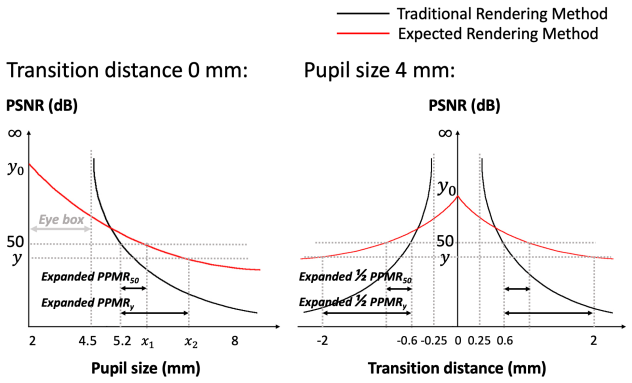

Figure 2: *The basic thinking of the rendering method for microdisplay image considering pupil size and human visual perception. The black curve is the PSNR of the traditional rendering method used in work by Lanman and Luebke [LL13] with an eye box size of 4.5 mm. The PSNR of the expected rendering method is represented by the red curve. (Left) Condition in the traditional rendering method as the black curve shown where a pupil is located at the center of the prototype; the Td is 0 mm. No cross-talk can be perceived when the pupil size is smaller than 5.2 mm. Cross-talk will be perceived if the pupil continues to grow beyond 5.2 mm. (Right) In the black curve, a pupil size of 4 mm and varied Tds. When Td is less than ±0.6mm, there is no cross-talk. If the Td continues to enlarge from ±0.6mm, the cross-talk will be perceived. y is a threshold for the expected rendering method. PSNR greater than this value indicates a high-quality image without perceptible cross-talk. If y remains 50, the expanded $PPMR_{50}$; otherwise the expanded area is $PPMR_y$. Section 3.2 verifies the particular value of y using human visual perception.*

3. The principle of enlarging PPMR for LA-NEDs

In this section, we introduce the basic thinking and the principle of the rendering method proposed in this paper. An analysis of the design parameter is followed.

3.1. Basic thinking

In the traditional rendering method, pupil size is not considered in the process of generating a microdisplay image. The elemental image allocation is employed in the traditional rendering method for LA-NEDs as in Lanman and Luebke's paper [LL13]. Each microdisplay pixel is correlated to a lens. In practice, considering a pupil size range from 2 mm to 8 mm in a prototype whose eye box is 4.5 mm [MFVHVdS18], the PSNR curve (black curve) of retina images is displayed on the left in Fig. 2. When the pupil size is smaller than the eye box size of 4.5 mm, the PSNR value is infinite [Sal08], indicating that there is no cross-talk. When the pupil size exceeds the 4.5 mm eye box, the lights from some pixels not only pass through one lens into the pupil but also other lenses, making cross-talk appear which causes PSNR to hugely decrease. Due to visual perception limitations, cross-talk still cannot be perceived when a PSNR value is larger than the threshold of 50 for the traditional rendering method, which we have already demonstrated in a previous study [YFU*22]. But when the pupil continues to enlarge making PSNR value smaller than the threshold, cross-talk will be perceived.

In the condition of the pupil size is a constant value about 4 mm [WG03] as seen on the right side of Fig. 2. Users should move their gaze to focus on the peripheral content of the display. The various Tds correspond to different PSNRs. When Td is less than ±0.6*mm* for a prototype with an eye box of 4.5 mm, the PSNR value of the retina image is larger than 50, indicating that cross-talk cannot be perceived. If the pupil continues to move far away from the prototype center, cross-talk will be perceived.

When a pupil dilates or moves out of the eye box, the retina image quality rapidly descends, resulting in the cross-talk being perceived in the traditional method. It verifies that the robustness of the prototype [Wei06, OHKS04], which refers to the ability of a system to resist change without adapting its initial stable configuration of the traditional rendering method, is low. So we want to improve the system robustness by modifying the rendering process as the red curve shows in Fig. 2, with the prototype being able to accommodate large pupils and Td.

Pupil size is considered in the process of generating microdisplay images. If the pupil is beyond the eye box, lights from some pixels enter the pupil not only through one lens but different lenses

and cause cross-talk to appear. These pixels are an essential component of an entire image, and cross-talk cannot be eliminated by simply deleting these pixels. If those pixels are deleted it would cause the intensity of image brightness not to be uniform, which will have a negative impact on image quality. Instead, when a pupil enlarges or moves out of the eye box, each pixel value should be computed based on the lights going into the pupil through several lenses. The rendering method based on this thinking results in a difference between the perceived image and the target image. But as long as the PSNR value is greater than a threshold y in Fig. 2, the brain can not perceive the difference in image quality, and the image still is considered a high-quality image. The expected PSNR curve of the expected rendering method is the red line shown in Fig. 2. When the threshold y equals 50, the expanded $PPMR_{50}$ is $(x_1 - 5.2)$. If cross-talk caused by the expected algorithm is not sharp as the traditional one, which means cross-talk is not expected to be perceived. It causes the threshold y to be smaller than 50. The expanded $PPMR_y$ is $(x_2 - 5.2)$. The threshold y of PSNR should be verified based on human visual perception.

3.2. Principle of the rendering method

To enlarge PPMR for LA-NEDs, a completely new rendering method is proposed considering pupil size and pupil margin as shown in Fig. 3. Different from the transitional microdisplay image rendering method, the lights of pixels not only go into the pupil and pupil margin through one lens but also through others, causing cross-talk to appear. Each lens should be considered for a pixel on the microdisplay. Ray lights emitted by one microdisplay pixel (MP) enter the pupil and pupil margin area after passing through a number of lenses. Each lens at the MP corresponds to one virtual pixel (VP) on the virtual image plane. The weight of each VP is the intersection area between the ray light column and the pupil and pupil margin divided by the sum of intersecting spaces between all the ray light columns generated by the MP and the pupil and pupil margin. The value of each MP is determined by the number of VPs and the related weight. As shown in Fig. 3, the lights of a pixel A through number z lenses go into the pupil and pupil margin. We quantify the intensity of the lights from pixel A entering the pupil and pupil margin after passing through different lenses by the area where the lights intersect with the pupil and pupil margin. The intersection area between each light beam and pupil and pupil margin is S_n as shown in Fig. 3. So the value a of pixel A is related to pixel V on the virtual image. For the eye in Fig. 3, the lights of pixel A go into the pupil and pupil margin through lenses 7, 8, 9 and 10. The lights of pixels A through lens 6 are not going into the pupil, so the weight of pixel V6 at pixel A is zero. The value a of pixel A is related to virtual pixels V7, V8, V9 and V10. In the process of generating a microdisplay image, if the virtual pixel V corresponding to microdisplay pixel A is out of range of the virtual image, the corresponding lens n will not be considered. Pixel V has a weight W_n for pixel A, and the equation of weight is as follows:

$$W_n = \frac{S_n}{\sum_{i=1}^{z} S_i} \tag{1}$$

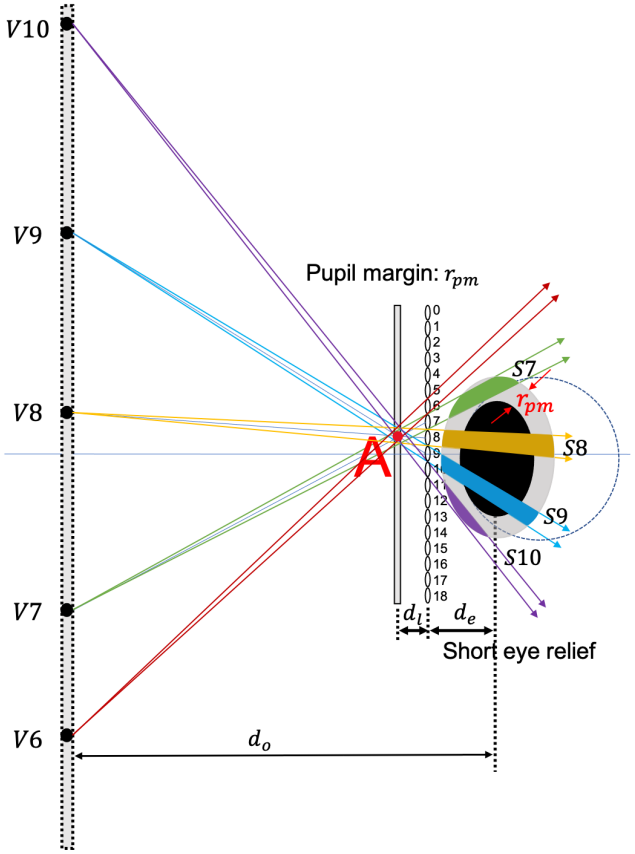

Figure 3: *The theory of the rendering method for microdisplay images proposed in this paper. Take pixel A on a microdisplay, whose light emits into the pupil plane through each lens. Cross-talk appears when the human eye advances beyond $PPMR_y$. "Pupil margin" r_{pm} is a crucial parameter in the system, which represents the maximum boundary of the area that a pupil can enlarge or move to without causing perceptible cross-talk. The number of ray lights through each lens goes into the pupil and the pupil margin should be counted. Among all the rays entering the pupil and pupil margin, calculate the weight ratio of each light column to the total light. For the eye, four light columns V_7, V_8, V_9 and V_{10} on the virtual plane go into the pupil and pupil margin. The areas S_7, S_8, S_9 and S_{10} are the intersections between the lights column and pupil and pupil margin. The lights of pixels A through lens 6 are not going into the pupil, so the weight of virtual pixel V6 at pixel A is zero. The weight of each column is calculated as introduced in equation 1. The final value a of pixel A is shown in equation 2.*

The value of a pixel A on a microdisplay equals:

$$a = \sum_{i=1}^{z} v_i W_i = \frac{\sum_{i=1}^{z} v_i S_i}{\sum_{i=1}^{z} S_i} \tag{2}$$

3.3. Result comparison

The comparison of rendering methods between different contents of images shown in Fig. 4, Fig. 6 and Fig. 1(e,f). The background

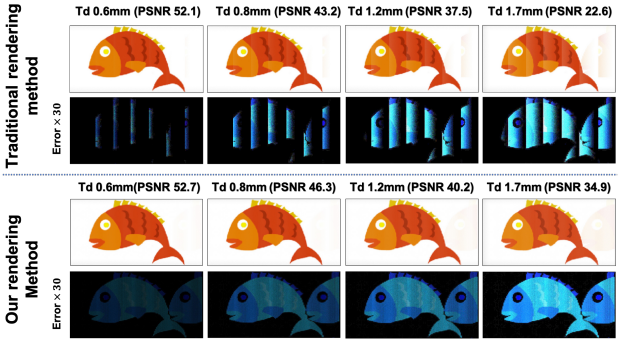

Figure 4: *The relationship between Td and PSNR values of the traditional rendering method used in Lanman and Luebke's paper [LL13] and the rendering method proposed in this paper under the condition of eye relief 15 mm and pupil size 4 mm in the retina simulation process. The first and third rows are the retina images with different rendering methods and Tds. The second and fourth rows are the amplified approximation errors between their upper corresponding retina images and target retina images times 30. In the simulated retina images of the traditional rendering method, the cross-talk is clear to be seen when Td is larger than 0.8 mm. In the simulated retina images of our rendering method, when Td is smaller than 1.2 mm, cross-talk can almost not be seen. Comparing the two kinds of retina images between two rendering methods, the PSNR of the former is a little bit larger than the latter, but the cross-talk in the former is perceived much earlier. So the two algorithms use different thresholds based on human visual perception. The PSNR thresholds of retina images in the traditional rendering method and our rendering method are 50 and 40 respectively.*

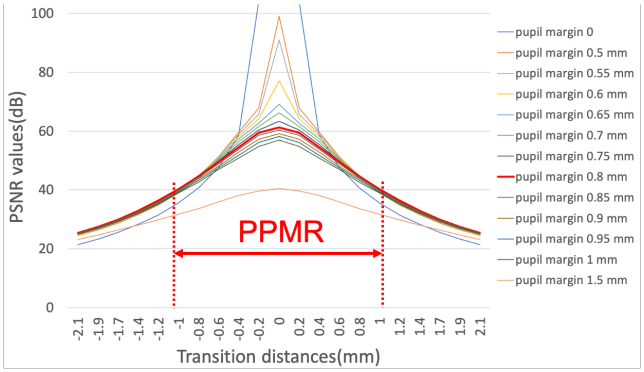

Figure 5: *PSNR curves with different pupil margins when eye relief is 15 mm. The $PPMR_{40}$ is largest when the pupil margin is 0.8 mm.*

of the fish image in Fig. 4 is uniformly white. When the pupil Td is 1.2 mm of the traditional rendering method for the fish content image in Fig. 4, cross-talk is clearly perceived. But in the rendering method proposed in this paper for the fish content image, when the Td is 1.2 mm almost cannot be perceived. The PSNR value of the Td 1.2 mm retina image is 41.4. To more easily see the difference, the approximation error images are scaled 30 times as shown in Fig. 4. As the pupil moves away from the center of the prototype, the intensity of the cross-talk-producing pixels is weakened in our rendering method. Thus, the intensity of the cross-talk generated by the traditional rendering method is stronger than our rendering method in the same Td condition. Different image contents verified that the cross-talk is easier to perceive in the traditional rendering method in the same Td condition. The cross-talk in a uniform white background content of images of both rendering methods is easier to perceive than the complex contents of images. So we conservatively set the threshold y in Fig. 2 to 50 and 40 for the traditional rendering method and our rendering method based on human visual perception respectively.

3.4. Design parameter selection

In this section, the best pupil margin size is chosen based on the threshold y of 40, which we conservatively set in section 3.3.

Pupil margin is a crucial parameter on the microdisplay image. Different pupil margins will generate different microdisplay images. The relationship between the results of our and traditional rendering methods is shown in Equation 3. When the double size of pupil margin r_{pm} is less than the size of the eye box W_e minus pupil size W_p, our rendering results MD_{our} and traditional rendering result $MD_{Traditional}$ are the same. In the condition of the eye relief is 15 mm and W_e is 4.5 mm for a pupil size W_p of 4 mm, the relationship between PPMR and pupil margin is shown in Fig. 5. The PPMR is at its highest value of $\pm 1.3mm$ when the pupil margin is 0.8 mm. In the traditional rendering method, PPMR does not exist when eye relief is 10 mm which causes the prototype cannot to be used in such short eye relief [LL13, BJCL21]. In the same condition of eye relief is 10 mm, high-quality images exist using our rendering method, and the PPMR40 value is $\pm 0.4mm$.

$$\begin{cases} MD_{our} = MD_{Traditional} & W_e - W_p \leq 2r_{pm} \\ MD_{our} \neq MD_{Traditional} & W_e - W_p > 2r_{pm} \end{cases} \quad (3)$$

In the model, we consider the case where the scene brightness changes pupil size physiologically. When the pupil size change from 3 mm to 5 mm, the intensity of the image used in Fig. 4 is 190, 190 and 192. So changes in image mean intensity are not significantly affected by pupil changes.

4. Implementation

Based on the considerations described in Section 3, we implemented a prototype of monocular LA-NED. In this section, the prototype is described in terms of hardware and software.

4.1. Hardware implementation

The monocular prototype shown in Fig. 1 (a) consists of a driver board, a display, and an LA. The glasses frame was printed using a 3D printer (Agilista-3200). The driver board comes from a Sony HMZ-T1 personal media viewer [Sona,Sonb] with magnifying eyepieces removed. The display in the prototype is a Sony ECX332A, which offers $1,280 \times 720$ (720p) resolution at 0.7" size. The focal length and size of each lens of the LA was $f = 3.3$ mm and

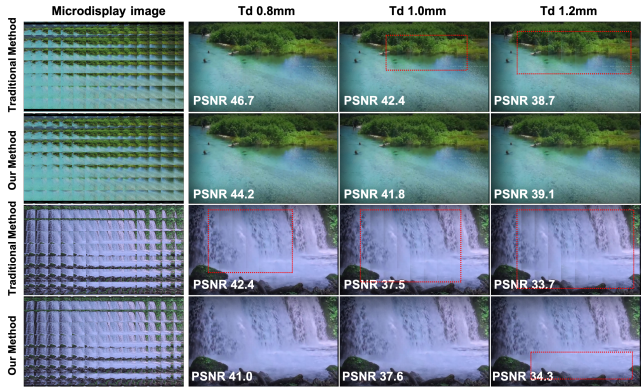

Figure 6: *Comparison of rendering methods between different contents of images. Using the traditional rendering method, when Td is 1.2 mm, cross-talk is perceived, and current PSNRs are 38.7 and 33.7 dB. For our rendering method, even if Td is 1.2 mm, PSNRs are 39.1 and 34.3 dB, cross-talk cannot be perceived.*

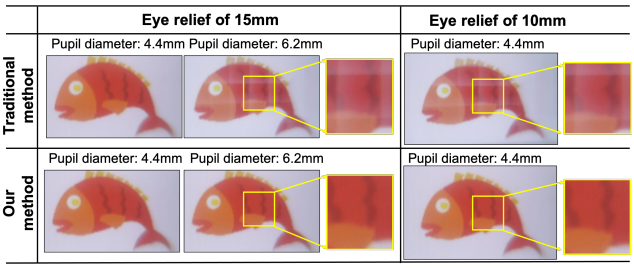

Figure 7: *Display results according to the pupil diameter (aperture size of the camera). The results are captured with a camera lens of f / 8 and 5.6, which can be adjusted mechanically on the camera, and a focal length of 35 mm. White arrows indicate the cross-talk. Such cross-talk exists both horizontally and vertically in the captured images, which means the pupil is out of the PPMR both horizontally and vertically.*

$w_l = 1$ mm. The LA is an N_l array of 23×15. The LA is placed at a distance d_l in front of the microdisplay. The parameters in our prototype are completely the same as Lanman's [LL13] prototype. Although d_l affects the distance from the display to the virtual image plane, we empirically set it so that the user can see a blur-free image through the LA-NED at $d_o = 0.5$ m.

4.2. Software implementation

All programs were implemented in C++ with an OpenGL library on a PC with an Intel Core i7 (Ubuntu, 3.7-GHz CPU, 16-GB RAM, and NVIDIA GEFORCE GTX 1080Ti GPU). The software structure included three parts: calibrations, microdisplay image generation, and retina image simulation part. Calibration results were used in the microdisplay image generation. The simulation part works separately from the main system and was used to visually and quantitatively analyze the images that users see, as shown in Fig. 5 in Section 3.

The angle offset calibration calculates the offset angle between the microdisplay and the LA. An angle offset caused by the manufacturing process is inevitable, and it converges each pixel of the elemental image at different points through the LA. It causes the virtual image to become dark and blurred. The purpose of angle offset calibration is to correct those pixel positions, causing those that belong to one virtual image to overlap. We observed changes in imaging quality by rotating the image on the screen and choosing the sharpest image through visual inspection. The rotation angle corresponding to the sharpest image was the offset angle between the LA and the microdisplay.

When generating microdisplay images, as shown in Fig. 3, each pixel of the microdisplay should be calculated based on the weight of each pixel on the virtual image. According to the principle of single-lens imaging, each pixel value of the microdisplay can be calculated. When the pupil position changes as shown in Fig. 3 from upper to lower, the weight of each pixel should be recalculated, causing each pixel value to change. The regenerated microdisplay images are shown in Fig. 1 (f) and the corresponding simulated retina image is shown in Fig. 1 (h).

5. Display result

In this section, we experimentally confirm that the proposed rendering method shows good performance consistent with the analysis.

5.1. Camera

In order to capture the same view as the VR user, the camera must have the ability to adjust the focus plane, and the entrance pupil must not be located too deep from the lens surface [BJCL21]. Thus, we used a Sony RX1 camera [Sonc], which has f / 8 and 5.6 mm, 35 mm focal length. In the display result experiment, we adjust the focus plane of the camera to 500 mm in front of the lens to capture the image through our prototype.

5.2. Different pupil sizes and eye reliefs

The camera lens is located at the center of the microdisplay. Simulated pupil diameter changes through camera aperture size change [Agu19]. The eye relief was measured to be 15 mm and 10 mm. The observed images show the captured images according to the camera aperture size as shown in Fig. 7. When eye relief is 15 mm, in the traditional rendering method cross-talk can be clearly seen in both horizontal and vertical axes with a pupil diameter of 6.2 mm, which means the pupil size is out of the PPMR both horizontally and vertically. In our rendering method, cross-talk cannot be seen in such pupil size. This experiment verified that our rendering method enlarges PPMR to accommodate a large pupil size.

In the condition of a sunglasses-like prototype with an eye relief of 10 mm, display results are rendered by traditional and our method as shown in Fig. 7. The cross-talk is perceived clearly in the traditional rendering method. But using our rendering method cross-talk disappears, which demonstrates our rendering method enables the prototype to be used at close eye relief.

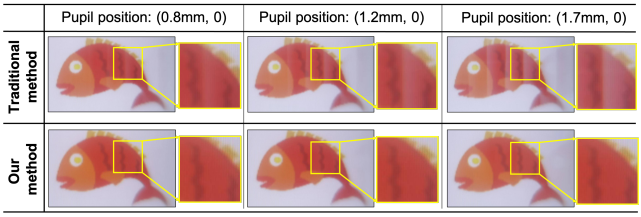

Figure 8: *Display results of pupil transition experiment in two rendering methods captured with the camera lens of f/8 and focal length 35 mm. The camera movement is along the horizontal axis. In the condition of the traditional rendering method and our rendering method, the cross-talk is detected when the pupil transition is larger than 0.5 mm and 2 mm. White arrows indicate the cross-talk. If such cross-talk only exists vertically in the captured images it means the pupil is out of the PPMR vertically.*

5.3. Different Tds

Fig 8 is the result of the pupil transition experiment with an eye relief of 15 mm. As expected from the analysis in Section 3, the cross-talk appears even when the Td is 0.5 mm. In our rendering method, the cross-talk does not appear. The perceived image is slightly distorted when the pupil Td is further than 2 mm. This also verifies that enlarged PPMR using our rendering method can tolerate a large Td in fixed pupil size.

6. Limitation and future work

To address the shortcomings of traditional rendering methods, we propose a rendering algorithm for LA-NEDs. When the human pupil becomes larger or moves further than the PPMR in the traditional method, the image quality degrades rapidly, negatively impacting the user experience. This kind of rendering algorithm is fragile and incapable of adapting to pupil changes. Our algorithm increases the system's robustness by allowing it to accommodate a large pupil and a wide range of pupil movements. But the perceived image will be distorted when Td is significant up to 2 mm as shown in Fig. 8. If the pupil continues to move out of the PPMR, the cross-talk will be perceived. That is, even though our rendering method significantly increased PPMR, the enlarged PPMR could not cover a large FOV prototype. The limitation can be overcome if the pupil position is detected in real-time and the displayed image is changed to correspond to the pupil position [BJCL21, YFU*22]. In the future, we will use eye tracking techniques and dynamically generate microdisplay images to eliminate all cross-talk in a large FOV prototype and for a wide range of fast-moving eyes.

7. Conclusion

We proposed a novel rendering method with strong robustness for lenslet array near-eye displays (LA-NEDs) that can enlarge the pupil practical movable region (PPMR) to accommodate a large pupil size and a wide pupil Td and get a short enough eye relief. Our rendering method counts the number of lights emitted by each pixel on the microdisplay that enters the pupil and pupil margin after passing through each lens of the lenslet array. Each lens at the microdisplay pixel corresponds to one virtual pixel on the virtual plane. The weight of each virtual pixel is determined by calculating the proportion of ray lights that go into the pupil and pupil margin. The value of each pixel on the microdisplay is calculated based on corresponding virtual pixels and weights. The experimental results captured using a camera, whose focus plane and aperture size can be adjusted to simulate a normal pupil size, demonstrated that the proposed rendering algorithm in this paper for LA-NEDs successfully enlarges PPMR to adopt a large pupil size and wide pupil Td, resulting in shortened eye relief such as with sunglasses. We anticipate that our efforts will significantly advance the VR industry.

References

[Agu19] AGUIRRE G. K.: A model of the entrance pupil of the human eye. *Scientific reports 9*, 1 (2019), 1–10. 6

[BJCL21] BANG K., JO Y., CHAE M., LEE B.: Lenslet VR: thin, flat and wide-fov virtual reality display using fresnel lens and lenslet array. *IEEE Transactions on Visualization & Computer Graphics 27*, 05 (2021), 2545–2554. 2, 5, 6, 7

[Bro91] BROOKS C. W.: *Understanding lens surfacing*. Elsevier Health Sciences, 1991. 2

[BS07] BLAKE R., SHIFFRAR M.: Perception of human motion. *Annual review of psychology 58* (2007), 47. 2

[CBC*20] CHOLEWIAK S. A., BAŞGÖZE Z., CAKMAKCI O., HOFFMAN D. M., COOPER E. A.: A perceptual eyebox for near-eye displays. *Optics Express 28*, 25 (2020), 38008–38028. 2

[CCPG19] CHANG C., CUI W., PARK J., GAO L.: Computational holographic maxwellian near-eye display with an expanded eyebox. *Scientific reports 9*, 1 (2019), 1–9. 3

[CHB19] CAKMAKCI O., HOFFMAN D. M., BALRAM N.: 31-4: Invited paper: 3D Eyebox in augmented and virtual reality optics. In *SID Symposium Digest of Technical Papers* (2019), vol. 50, Wiley Online Library, pp. 438–441. 2

[CLN20] CHERVYAKOV N., LYAKHOV P., NAGORNOV N.: Analysis of the quantization noise in discrete wavelet transform filters for 3D medical imaging. *Applied Sciences 10*, 4 (2020), 1223. 2

[Gab08] GABRIEL L.: La photographie intégrale. *Comptes-Rendus, Académie des Sciences 146* (1908), 446–551. 3

[HKL*08] HAHN J., KIM H., LIM Y., PARK G., LEE B.: Wide viewing angle dynamic holographic stereogram with a curved array of spatial light modulators. *Optics express 16*, 16 (2008), 12372–12386. 2

[JBLL18] JANG C., BANG K., LI G., LEE B.: Holographic near-eye display with expanded eye-box. *ACM Transactions on Graphics 37*, 6 (2018), 1–14. 3

[JBM*17] JANG C., BANG K., MOON S., KIM J., LEE S., LEE B.: Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina. *ACM Transactions on Graphics 36*, 6 (2017), 1–13. 2

[JYB*21] JO Y., YOO C., BANG K., LEE B., LEE B.: Eye-box extended retinal projection type near-eye display with multiple independent viewpoints. *Applied Optics 60*, 4 (2021), A268–A276. 2

[KJS*19] KIM J., JEONG Y., STENGEL M., AKSIT K., ALBERT R. A., BOUDAOUD B., GREER T., KIM J., LOPES W., MAJERCIK Z., ET AL.: Foveated AR: dynamically-foveated augmented reality display. *ACM Transactions on Graphics 38*, 4 (2019), 99–1. 2

[KP18] KIM S.-B., PARK J.-H.: Optical see-through maxwellian near-to-eye display with an enlarged eyebox. *Optics letters 43*, 4 (2018), 767–770. 2

[KS13] KRESS B., STARNER T.: A review of head-mounted displays

(HMD) technologies and applications for consumer electronics. *Photonic Applications for Aerospace, Commercial, and Harsh Environments IV 8720* (2013), 62–74. 2

[KWNM20] KUO G., WALLER L., NG R., MAIMONE A.: High resolution étendue expansion for holographic displays. *ACM Transactions on Graphics 39*, 4 (2020), 66–1. 2

[LJM*16] LEE S., JANG C., MOON S., CHO J., LEE B.: Additive light field displays: realization of augmented reality with holographic optical elements. *ACM Transactions on Graphics 35*, 4 (2016), 1–13. 2

[LJY*19] LEE S., JO Y., YOO D., CHO J., LEE B.: Tomographic near-eye displays. *Nature Communications 10*, 1 (2019), 2497. 2

[LL13] LANMAN D., LUEBKE D.: Near-eye light field displays. *ACM Transactions on Graphics 32*, 6 (2013), 1–10. 2, 3, 5, 6

[LZZ*20] LIN T., ZHAN T., ZOU J., FAN F., WU S.-T.: Maxwellian near-eye display with an expanded eyebox. *Optics Express 28*, 26 (2020), 38616–38625. 2

[MFVHVdS18] MATHÔT S., FABIUS J., VAN HEUSDEN E., VAN DER STIGCHEL S.: Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior research methods 50*, 1 (2018), 94–106. 3

[MWH*13] MAIMONE A., WETZSTEIN G., HIRSCH M., LANMAN D., RASKAR R., FUCHS H.: Focus 3D: Compressive accommodation display. *ACM Transactions on Graphics 32*, 5 (2013), 153. 2

[Ocu] OCULUS: https://www.oculus.com/. Accessed: 2022-07-22. 2

[OHKS04] O'MAHONY M., HURLEY N., KUSHMERICK N., SILVESTRE G.: Collaborative recommendation: A robustness analysis. *ACM Transactions on Internet Technology 4*, 4 (2004), 344–377. 3

[PL22] PARK J.-H., LEE B.: Holographic techniques for augmented reality and virtual reality near-eye displays. *Light: Advanced Manufacturing 3*, 1 (2022), 1–14. 2

[RSAA20] RATCLIFF J., SUPIKOV A., ALFARO S., AZUMA R.: ThinVR: Heterogeneous microlens arrays for compact, 180 degree FOV VR near-eye displays. *IEEE transactions on visualization and computer graphics 26*, 5 (2020), 1981–1990. 2

[RWBF18] RATHINAVEL K., WANG H., BLATE A., FUCHS H.: An extended depth-at-field volumetric near-eye augmented reality display. *IEEE transactions on visualization and computer graphics 24*, 11 (2018), 2857–2866. 3

[Sal08] SALOMON D.: Data compression: The complete reference (by d. salomon; 2007) [book review]. *IEEE Signal Processing Magazine 25*, 2 (2008), 147–149. 2, 3

[Sona] SONY HMZ-T1: https://www.sony.jp/hmd/products/HMZ-T1/spec.html. Accessed 2021-08-15. 5

[Sonb] SONY-HMZ-T1-PERSIONAL-3D-VIEWER: https://www.cnet.com/reviews/sony-hmz-t1-personal-3d-viewer-review/. Accessed 2021-08-15. 5

[Sonc] SONY RX1: https://www.sony.jp/cyber-shot/products/DSC-RX1/. Accessed: 2022-07-22. 6

[Sut68] SUTHERLAND I. E.: A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I* (1968), pp. 757–764. 2

[Viv] VIVE: https://www.vive.com/jp/. Accessed: 2022-07-22. 2

[Wei06] WEISBERG M.: Robustness analysis. *Philosophy of science 73*, 5 (2006), 730–742. 3

[Wel99] WELSTEAD S. T.: *Fractal and wavelet image compression techniques*, vol. 40. Spie Press, 1999. 2

[WG03] WITTING M. D., GOYAL D.: Normal pupillary size in fluorescent and bright light. *Annals of emergency medicine 41*, 2 (2003), 247–250. 3

[Wil15] WILSON D.: Vertex distance and pantoscopic angle–a review, 2015. 2

[WLD*18] WU F., LV G.-J., DENG H., ZHAO B.-C., WANG Q.-H.: Dual-view integral imaging three-dimensional display using polarized glasses. *Applied Optics 57*, 6 (2018), 1447–1449. 3

[WXHJ15] WANG J., XIAO X., HUA H., JAVIDI B.: Augmented reality 3D displays with micro integral imaging. *Journal of Display Technology 11*, 11 (2015), 889–893. 3

[XGS*19] XIA X., GUAN Y., STATE A., CHAKRAVARTHULA P., RATHINAVEL K., CHAM T.-J., FUCHS H.: Towards a switchable AR/VR near-eye display with accommodation-vergence and eyeglass prescription support. *IEEE transactions on visualization and computer graphics 25*, 11 (2019), 3114–3124. 2

[Yan18] YANG J.: *Study and optimization of an optical see-through near to eye display system for augmented reality*. PhD thesis, Université de Strasbourg, 2018. 2

[YCML20] YOO C., CHAE M., MOON S., LEE B.: Retinal projection type lightguide-based near-eye display with switchable viewpoints. *Optics Express 28*, 3 (2020), 3116–3135. 2

[YCYW18] YAO C., CHENG D., YANG T., WANG Y.: Design of an optical see-through light-field near-eye display using a discrete lenslet array. *Optics Express 26*, 14 (2018), 18292–18301. 3

[YFU*22] YE B., FUJIMOTO Y., UCHIMINE Y., SAWABE T., KANBARA M., KATO H.: Cross-talk elimination for lenslet array near eye display based on eye-gaze tracking. *Optics Express 30*, 10 (2022), 16196–16216. 2, 3, 7

[ZMX*19] ZHAO J., MA Q., XIA J., WU J., DU B., ZHANG H.: Hybrid computational near-eye light field display. *IEEE Photonics Journal 11*, 1 (2019), 1–10. 3