AdaptiVisor: Assisting Eye Adaptation via Occlusive Optical See-Through Head-Mounted Displays

Yuichi Hiroi
y.hiroi@imlab.ics.keio.ac.jp
Keio University

Yuta Itoh
itoh@imlab.ics.keio.ac.jp
Keio University

Takumi Hamasaki
hamasaki@imlab.ics.keio.ac.jp
Keio University

Maki Sugimoto
sugimoto@imlab.ics.keio.ac.jp
Keio University

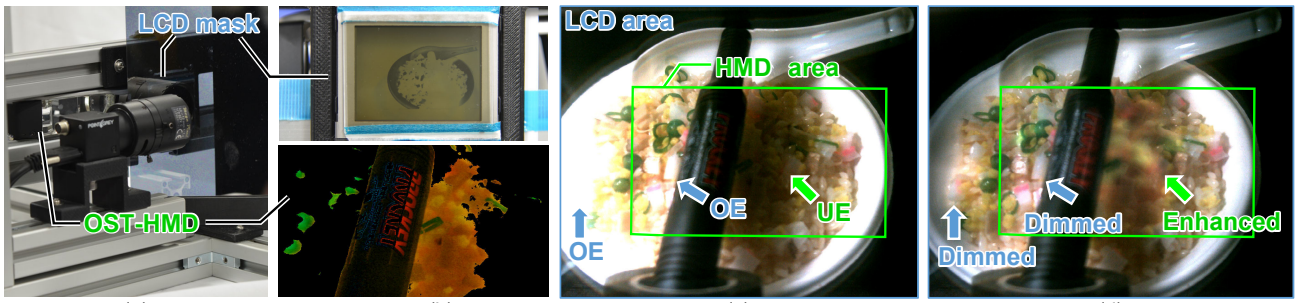

Figure 1: An overview of our AdaptiVisor system: (a) The front view of the system, (b top) the LCD showing an occlusion mask captured directly by a camera with a white background, (b bottom) an input image for the OST-HMD screen to show an additive image, (c) a scene image without adaptation assistance where one can observe over-/underexposure (OE/UE), and (d) with the assistance where the OE is suppressed by the occlusion mask and UE is compensated by the OST-HMD. The scene blur is due to the diffraction of the occlusion mask used in our prototype.

ABSTRACT

Brightness adaptation is a fundamental ability in human visual system, and adjusts various levels of darkness and light. While this ability is continuously used, and it can mostly handle sudden lighting changes in the environment, the adaptation could still take several minutes. Moreover, during the adaptation, the color perception changes as well. This slow reactivity and perception change of the eyes could lead to mistakes for tasks performed in dazzling or temporally high-contrast environments such as when driving into the sun or during a welding process.

We present AdaptiVisor, a vision augmentation system that assists the brightness adaptation of the eye. Our system selectively modulates the intensity of the light coming into the eyes via occlusion-capable Optical See-Through Head-Mounted Displays (OST-HMD). An integrated camera captures highlights and brightness in the environment via high-dynamic range capture, and our display system selectively dims or enhances part of field of views so that the user would not perceive rapid brightness changes. We build a proof-of-concept system to evaluate the feasibility of the adaptation assistance by combining a transmissive LCD panel and an OST-HMD, and test it with a user-perspective, view-point camera. The evaluation shows that the system decreases the overexposed area in a scene to 1/15th, and enhances the color by reducing majorly underexposed area to half. We also include a preliminary user trial and it indicates that the system also works for real eyes for the HMD part and to some extent for the LCD.

KEYWORDS

Vision Augmentation, Augmented Reality, Optical See-through, Head-mounted Displays, Occlusive HMD, Near-eye Displays, Brightness Adaptation, AdaptiVisor

ACM Reference format:
Yuichi Hiroi, Yuta Itoh, Takumi Hamasaki, and Maki Sugimoto. 2017. AdaptiVisor: Assisting Eye Adaptation via Occlusive Optical See-Through Head-Mounted Displays . In *Proceedings of Augmented Human, Mountain View, CA, USA, March 16-18 2017 (AH'17),* 9 pages.
DOI: http://dx.doi.org/10.1145/3041164.3041178

1 INTRODUCTION

Brightness adaptation of the eye is a fundamental ability in human visual system. By sensing the surrounding brightness, our eyes adjust the amount of light passing through the lens by widening the pupil [34]. Moreover, our eyes adjust the sensitivity of the retina by switching between rod and cone cells in the photoreceptors [7].

However, once the switching occurs in the both types, it takes a certain amount of time to return to the state where the cells can react again. In other words, the brightness adaptation of the eye cannot keep up with the rapid change of the brightness [7]. When the environment lighting shifts from bright to dark, the adaptation already takes 20-40 seconds to adapt to the dark environment [3]. Even worse, when the shift is from dark to bright, the adaptation takes more than 30 minutes to completely adapt to the bright environment [9].

This is a pre-print.

In our daily life, there are various situations where such adaptation cannot complete in time, which may lead to decreased task performance, or even serious accidents. For example, when a light from an oncoming vehicle gets into the eyes during night driving, the eye first dazzles since the intense beam light irradiates the eyes that are adapted to a dark environment, and then the bright adaption is quickly occurred. As the result, it would take time for the eyes to adapt back to the dark environments.

To avoid such risks caused by the adaptation, this paper conducts a feasibility study of Vision Augmentation (VA) in brightness adaptation assistance via Optical See-Through Head-Mounted Displays (OST-HMDs). For the study, we build AdaptiVisor, an OST-HMD system that assists the brightness adaptation (Fig. 1). Our approach combines a transmissive LCD panel as an occlusion layer and an OST-HMD as an additive layer to dynamically block or add the light entering into the eyes, which allows controling the brightness of the scene at a viewpoint.

Among related VA systems that cut the amount of environment light reaching to the eyes (See Sec. 2.1), Bhagavathula et al. proposed one of the most similar systems to ours [5]. They developed a low-power, 32×64-pixel CMOS glare sensor, and combined it with an LCD shutter glasses to form a wearable glare reduction glasses. While their system can reduce a glare of a moving point light, it does not block light from different field of view (FoV) areas simultaneously. Furthermore, due to the limited resolution of both the sensor and the LCD, the shape of their LCD occlusion mask is limited to a simple square, thus it cannot smoothly dim bright, complex surfaces. Another problem of their system is that it only switches the transmittance of the LCD between on and off, thus it does not adaptively control the occlusion strength according to the brightness of the scene.

Finally, such wearable dimming glasses practically requires spatial calibration between the display system and user's eyes, i.e. the coordinates of the display layers and of the eyes have to be calibrated each other so that the system can render occlusion masks at the right location to correctly dim bright FoV [27].

Our system uses a scene camera that detects both the change of the brightness and the position of particularly bright or dark areas in the scene. The scene camera is optically aligned to the position of the viewpoint for the OST-HMD by a beam splitter. Combined with a spatially-calibrated transmissive LCD and HMD, our system can selectively control the light entering into the eye up to the pixel resolution of each occlusion and display layers in realtime.

Contribution: Our main contributions include:

- implementing a proof-of-concept OST-HMD system that selectively suppress and enhance the amount of the light entering into the user's vision up to the pixel resolution of the LCD mask and HMD screen in realtime,
- conducting quantitative and qualitative analyses of the system with a user-perspective camera, which shows that the system achieves to decrease the overexposed area in a scene to 1/15th, and enhances the color by reducing majorly underexposed area to half.
- providing a thorough analysis of the current setup including limitations and possible research directions.

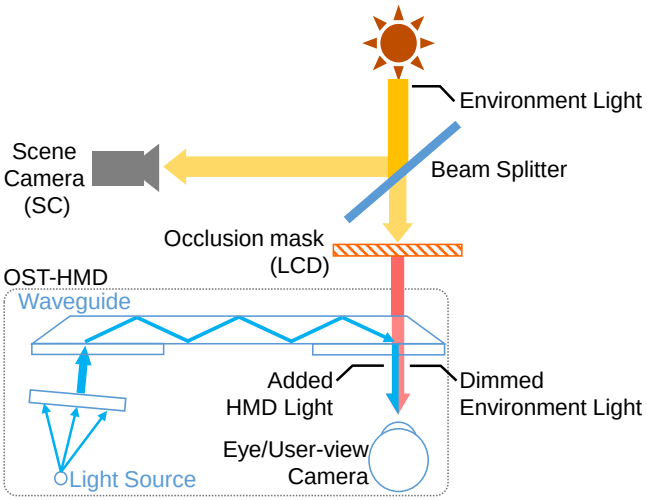

Figure 2: Schematic diagram of our adaptation assistance system with occlusive OST-HMDs.

2 RELATED WORK

2.1 Vision Augmentation

Various VA systems have been proposed to expand and assist people's vision. Itoh et al. proposed a Vision Enhancement concept to correct user's optical defects by an OST-HMD [15]. Their system creates a filter image that is designed to cancel visual aberration, and the image is superimposed on the user's field of view via an OST-HMD.

In the context of VA systems for adjusting the scene brightness, Rekimoto proposed Squama, a programmable window that can dynamically and selectively adjust light-blocking properties using smart window panels that can switch transparency [27]. Their system detects the positions of external light and indoor objects, and can adjust which panel to turn on so that the light does not hit the object.

Ma et al. developed smart sunglasses that dims the light entering into the eyes by modulating the transmittance of the LCD panel [20]. However, the glasses decreases the brightness of the entire scene uniformly, thus cannot selectively control the dimming level. Aiteanu et al. developed a video see-through welding helmet that provides contrast-enhanced images [1]. While such systems can freely manipulate the scene brightness, its video see-through view inevitably loses both the resolution and depth of user's field of view (FoV).

Hara et al. demonstrated a glare removal system for transparent surfaces [10]. Their LCD-based system removes the glare of water droplets on a transparent surface by detecting them with a camera and by tuning the transmittance of the surface on which the droplets stay. While their system can remove such disturbing effects caused by obstacles on the lens, it does not consider controlling the brightness of the light from the scene. Santos et al. developed an occlusion-capable see-through binocular display that uses LCD panels [28]. Although their LCD mask is controllable, the system only creates masks for the position of the virtual contents to be rendered, and does not analyze scene brightness.

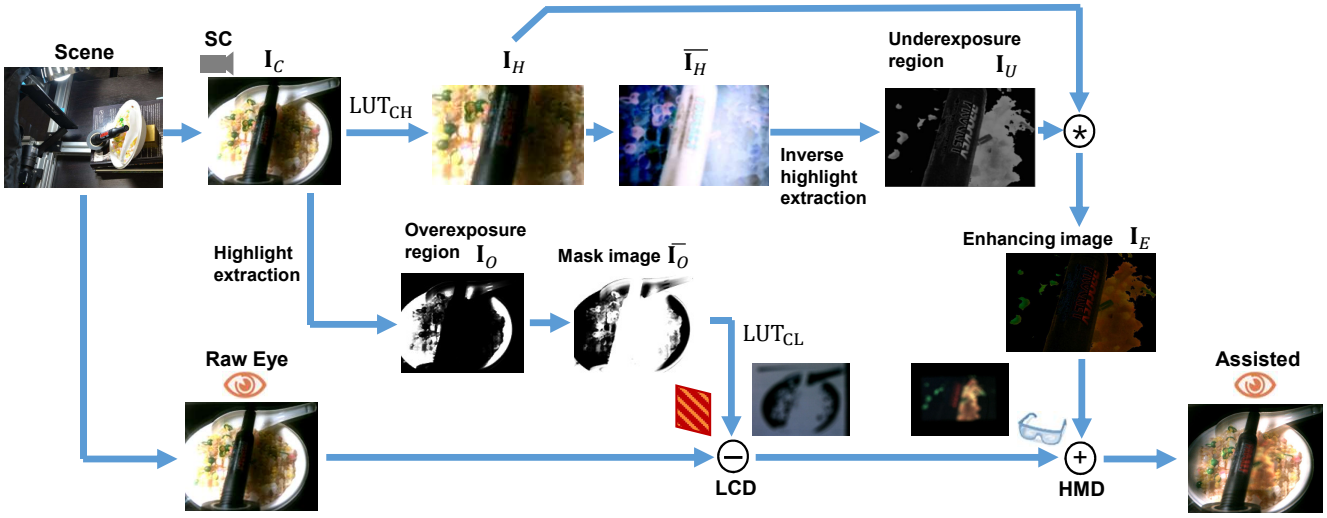

Figure 3: Algorithm flow of the computation of the mask and HMD images.

2.2 Computational Photography

In computational photography, there are rich researches on appearance control. Nayar and Branzoi proposed adaptive dynamic range imaging that adapts the exposure of each pixel on a camera image, based on the brightness of the corresponding scene point [24]. Their system attaches an LCD mask in front of a camera and uses the mask to attenuate light rays for each pixel. Amano et al. constructed a system that dynamically controls the appearance of real objects using a projector-camera system (PCS) [2], realizing various visual-aid effects such as color vision correction, color removal, and contrast enhancement.

Tamburo et al. developed a smart headlight system for automobiles [30]. Their system combines a DLP projector based on a high-resolution digital micro-mirror device with a high-speed camera system. The authors demonstrated that their system with 1 to 2.5 ms latency could illuminate an outdoor scene while avoiding the light to illuminate falling snow flakes that would cause flickering.

2.3 Occlusion-Capable OST-HMDs

There have been works on implementing occlusion capability on OST-HMDs [17]. The common approach is to use a spatial light modulator (SLM) [31]. The approach uses either transmissive SLMs [18] or refractive SLMs [6, 8], and cuts off light rays between the objects and the eyes by selectively controlling display pixels.

These methods, however, require extra optics to guide light rays from the world to pass through the SLM layers. Real FoV of a user thus might be distorted by the optics. Maimone et al. developed an OST-HMD with stacked LCD layers and a shutter [21]. By switching the shutter, their display uses the LCD layers for both displaying contents and occluding background light. Similarly, Maimone et al. proposed Pinlight Display [22]. Their display uses a transparent plate with a dot array that provides the point light sources. Combined with the same time-multiplexing shutter technique, the display can achieve occlusion with larger FoV.

Recently, Yamaguchi and Takagi developed an occlusive OST-HMD using integral imaging display [35]. The display combines three lens-array layers and two LCD panels. The two LCD panels, one for image rendering and another for occlusion, are inserted between each layers separately.

3 METHOD

3.1 System Overview

Figure 2 shows a schematic diagram of our adaptation-assistance system using an OST-HMD and an occlusion mask. Our approach combines a transmissive LCD panel and an OST-HMD to dynamically dim or add the environmental light entering into the eyes, which allows to control the brightness of the scene at a viewpoint. The scene camera (SC) captures the information of the environment light. The user-perspective camera (UC) represents the user's eye for our test to obtain images from the viewpoint. The coordinates of the cameras and the displays are calibrated beforehand, as mentioned in Sec. 4.2.

Firstly, to specify the position in which the light needs to be controlled at the viewpoint, the SC is set at the optically same position as the viewpoint by using a beam splitter as with [19]. Then, the controlled occlusion mask dynamically dims the environment light passing coming through the beam splitter toward the user's eye. The occlusion mask blocks light from a scene area where the highlight is strong at the viewpoint. After that, the OST-HMD layer provides the brightness and the color to the positions that are not clearly visible from the viewpoint. As a result, the amount of the light entering into the user's eye is dynamically controlled.

3.2 Computation of Displaying Images

In our system, the LCD attenuates the overexposure region of the light entering into the eyes and the OST-HMD compensates the underexposure region. By adding or subtracting the light with displays pixel-wise, our system can selectively change the amount

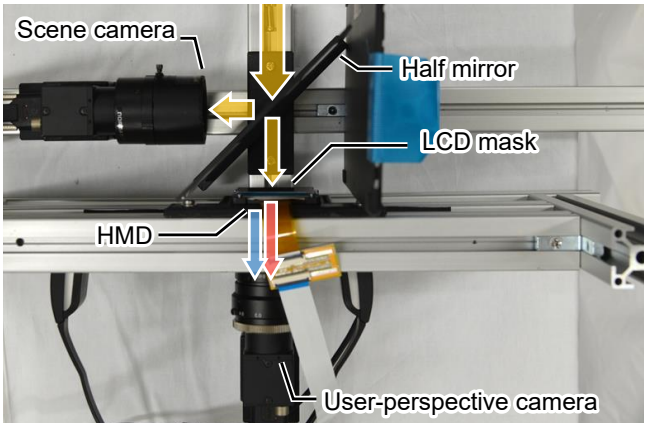

Figure 4: Hardware setup with schematic visualization of light paths in Fig. 2. The image is vertically flipped for the visualization purpose.

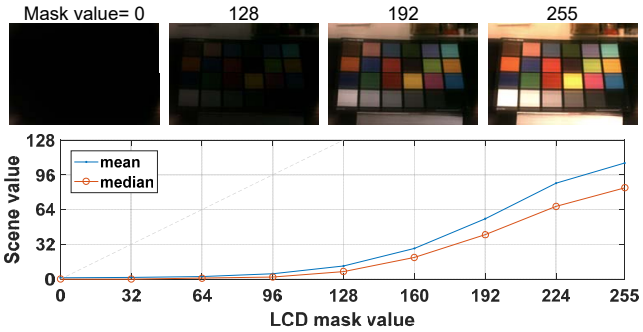

Figure 5: Transmittance profile of the LCD panel used in our setup. Given uniform values on the entire LCD mask from 0 to 255, the transmittance changes nonlinearly when seen by a gamma-corrected camera.

of the light entering the eyes up to the pixel resolution of the displays.

Figure 3 shows an overview of our algorithm flow. Our system has two key computation steps: To compute the mask image displayed on the LCD and the enhancing image displayed on the HMD. $\mathbf{I}$ represents a 2D image. Both the overexposure region and the underexposure region are detected from the SC image $\mathbf{I}_S$.

3.2.1 Compute the mask image. To compute the mask image, we first obtain the overexposure region $\mathbf{I}_O$ from $\mathbf{I}_S$. In our current setup, we apply the real-time highlight removal method proposed by Shen et al [29] to $\mathbf{I}_S$, then we treat the removed highlight image as $\mathbf{I}_O$. Since $\mathbf{I}_O$ shows the pixel-wise intensity of lights on the SC, the mask image $\overline{\mathbf{I}_O}$ is obtained by the negative/positive conversion of $\mathbf{I}_O$ to compensate the intensity. Finally, LCD displays the image $\mathrm{LUT}_{\mathrm{CL}}(\overline{\mathbf{I}_O})$, where $\mathrm{LUT}_{\mathrm{CL}}(\cdot)$ is coordinate transformation by the look-up table (LUT) between the SC and the LCD (described in Sec. 4.2.3).

3.2.2 Compute the enhancing image. To obtain the image displayed on HMD, we first transform $\mathbf{I}_S$ to $\mathbf{I}_H = \mathrm{LUT}_{\mathrm{CH}}(\mathbf{I}_S)$ by the LUT between SC and the HMD (Sec. 4.2.2). Due to the limited FoV of the OST-HMD, the area of captured OST-HMD screen is smaller than the entire camera image. We thus transform the camera image beforehand to reduce the computation area in the image.

We calculate $\overline{\mathbf{I}_H}$ by the negative/positive conversion of $\mathbf{I}_H$. Since the underexposure region such as blocked up shadows in $\mathbf{I}_H$ is converted to (inverse-)highlight in $\overline{\mathbf{I}_H}$, we obtain $\mathbf{I}_U$ as a gray-scale image by applying the same highlight removal to $\overline{\mathbf{I}_H}$.

While the LCD can display only the gray-scale image, the HMD can compensate the RGB color of the environment. For this reason, we convert the gray-scale image $\mathbf{I}_U$ into the color enhancing image $\mathbf{I}_E$ by using the original image $\mathbf{I}_H$. The RGB color $\mathbf{I}_E(x)$ at pixel x is calculated from $\mathbf{I}_H(x) = [I_H^r(x), I_H^g(x), I_H^b(x)]^T$ and the gray-scale value $I_U(x)$:

$$\mathbf{I}_E(x) = \alpha I_U(x)\mathbf{I}_H(x) \tag{1}$$

where α is the enhancing ratio. In Eq.1, $\mathbf{I}_U$ is applied as a mask to $\mathbf{I}_H$ to obtain the image $\mathbf{I}_E$ displayed on the OST-HMD. By this masking process, the black background is kept black and only the region where the color and brightness is faded can be extracted. Through the experiment, we notice that just displaying $\mathbf{I}_E$ (i.e. $\alpha = 1.0$) is not enough to fully compensate the color and brightness. For this reason, we manually set $\alpha = 2.5$ and trimmed the exceeded elements $I_E^{r,g,b}(x) > 255$ to 255. The proper value of α depended on the brightness of HMD, therefore beyond the scope of this paper. We discuss the color and brightness calibration on Sec.6.3.

4 TECHNICAL SETUP

We describe our hardware and software setup and elaborate calibration steps.

4.1 Hardware and Software Setup

Figure 4 shows our hardware setup. Our system combines an Epson Moverio BT-200 (23° field of view) as an OST-HMD with a LCD panel (SONY LCX017, 1024×768 pixel) for the opaque layer. We use an acrylic beam splitter with 30% transmittance and 70% reflectance. Some fixtures were self-made parts using a 3D printer. All devices are connected to the same Windows 10 desktop machine (Intel Core i7-6700K CPU 4.00GHz, 16GB RAM). The LCD panel is connected to the computer via the controller made by bbs Bild- und Lichtssysteme.

Figure 5 shows the transmittance property of the LCD panel measured by a gamma-corrected camera. We displayed uniform values, i.e. uniform occlusion masks, on the LCD while capturing the scene through the panel by the camera. The figure shows the profile of the captured images in terms of the measured brightness. Note that, in this preliminary test, we did not place the beam splitter.

To output video signal from a PC to the BT200, we used a DM484CS DVI-D (HDMI) Interface for Moverio BT-200 from Colorado Video. We set the virtual video signal to 1280×720 pixel for the display, which is larger than the actual display resolution, yet yields the same aspect ratio of 16:9.

To obtain images from a viewpoint through the OST-HMD, we installed a user-perspective camera placed behind the left optical element of the display. For both the user-perspective and the scene camera, we use PointGrey Flea3 FL3-U3-13E4C-C with a 1280×960

resolution. The model uses 1/1.8" CMOS sensor. For the two cameras, we set the frame rate to 30Hz.

For the scene camera, we used a C-mount lens, Kowa LM12HC (focal length 12.5mm). For the user-perspective camera, we used a varifocal C-mount lens, Tamron M12VM412, and set its focal distance to 12mm.

For calibration procedure and image processing, we used OpenCV and a calibration framework, Ubitrack [25], to handle PointGray cameras, and Matlab 2016 to compute calibration data.

4.2 System Calibration

To control the brightness of each light rays, it is crucial to create a pixel-wise mapping from an observed camera pixel in SC to the corresponding display pixel and a user view, i.e. UC in our case. Figure 6 summarizes each calibration steps described in the following.

4.2.1 Calibration between the User-Perspective and the Scene Camera (Fig. 6 A). We first compute a homography between SC and UC. Assuming that the two cameras are optically almost aligned, we decided to approximate the mapping between the two cameras by homography. We let both SC and UC capture a physical checkerboard simultaneously, and then we applied a checker board detection to automatically detect grid corners. Finally, we computed a 3×3 homography matrix out of the 2D point correspondences. Since both cameras used lenses with rather long focal distances, we left the cameras uncalibrated and did not undistort.

4.2.2 Calibration between the HMD Screen and the User-Perspective Camera (Fig. 6 B1). We compute $\mathrm{LUT}_{\mathrm{CH}}(\cdot)$, the LCT from the UC image plane to HMD screen, by showing gray-code patterns on the HMD screen while capturing them by the UC. By applying a standard LUT calculation, we can compute an LUT from the camera to the screen. Note that the direction of the LUT obtained by the above procedure is opposite of what we require, i.e. we need to transfer the pixel position of the screen to that of the UC image.

In our implementation, we applied Gaussian kernel regression between the 2D correspondences to build a continuous mapping function between the image planes. We then built an integer LUT from the screen to the camera.

4.2.3 Calibration between the LCD Screen and the User-Perspective Camera (Fig. 6 B2). In principal, $\mathrm{LUT}_{\mathrm{CL}}(\cdot)$, the LUT from the UC image plane to LCD mask, can be computed in the same manner as the above step. Note that, however, the occlusion mask is out of focus for the UC and the mask itself is just a transparent layer without background illumination. We therefore let the UC focus on the LCD to capture sharp gray-code patterns. We also placed uniform white diffuse illumination at the background to make the patterns visible.

We then let the UC focus to the final background scene after calibration. Thus, the LCD image is not perfectly calibrated against UC in pixel-wise. This misalignment is, however, has less impact than the blur effect induced by the LCD, and both the spatial calibration and the mask blur are beyond the scope of this paper.

Nevertheless, we discuss this issues and possible solutions later in Sec.6.1 and 6.4 to provide future research questions.

In this calibration, gray-code patterns gave smaller number of reliable 2D correspondence pairs compared to Sec. 4.2.2. To properly estimate the continuous mapping function, we set a wider kernel width for the regression thus the LUT. Since the LCD is larger than the camera lens, the camera captured a part of the LCD panel.

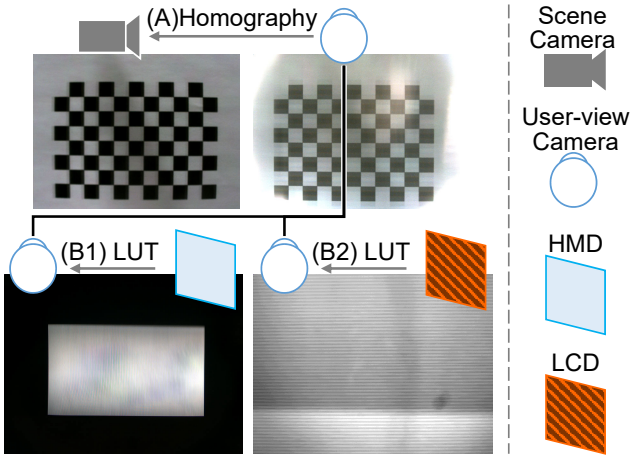

Figure 6: System calibration overview. Combining mapping functions A and B1 or B2 creates desired mappings for mask and HMD image visualization.

5 EXPERIMENTS

We conduct two experiments with our proof-of-concept system. The static capture experiment (Sec. 5.2) quantitatively evaluates the adaptation ability of our system under a controlled environment. The dynamic capture experiment (Sec. 5.3) clarifies issues of the current system under more practical use where the system runs in realtime under a dynamic-lighting condition. We then discuss possible solutions for realizing a practical system in the discussion section. Finally, to see actual impression of the users, we include preliminary feedback from real users by letting them try the system.

5.1 Experiment Setup

For evaluation, we built a workbench that consists of a target object to be captured and an external light source to be switched (Fig. 8). The target is a plastic food model and a black marker pen. The target object is placed at about 30cm away from UC. We then let UC capture the target object while the external lighting is switched on and off. We evaluate different adaptation assistance conditions by choosing if the HMD is used or not and if the occlusion mask by the LCD is used or not. We thus have four conditions: C1 {LCD:off,HMD:off} (i.e. raw view), C2 {LCD:on,HMD:off}, C3 {LCD:off,HMD:on}, and C4 {LCD:on,HMD:on}.

We keep the exposure of UC fixed through the experiment so that the comparison among different conditions becomes fair.

5.2 Static Capture Result

We first evaluate a static condition where we capture the scene with the four conditions offline. Although our system runs realtime, capture and display latency induce a constant delay from the lighting change to the system starts assistance. This affects the visualization quality when the lighting condition changes. To see

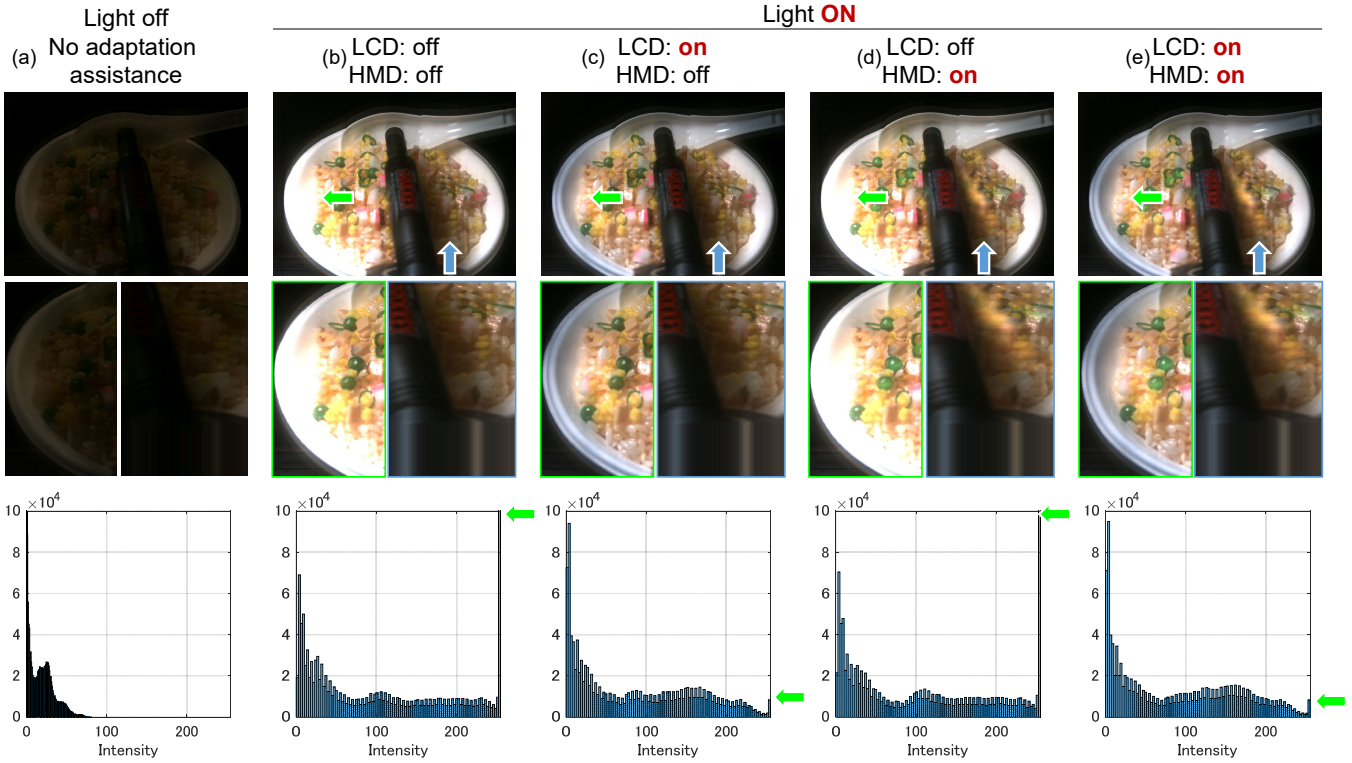

Figure 7: Overview of the static-capture experiment in Sec. 5.2.

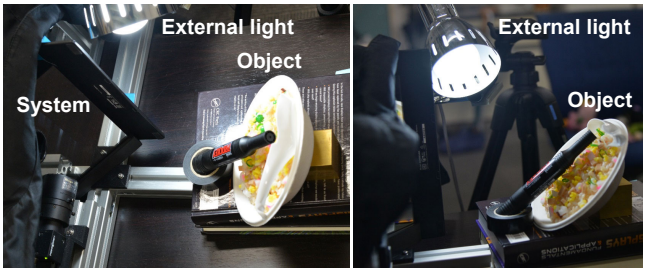

Figure 8: Experiment setup. A target object is placed at about 30cm away from the system, and an extra lighting is set above the object.

the static performance of our proof-of-concept system, we first create in offline an occlusion mask and a brightness compensation image by capturing the static target under the extra lighting, we then capture the view with these mask and image shown on the LCD and the HMD respectively.

Figure 7 shows the summary of the experiment. Without the extra light, the scene appears dark to UC (Fig. 7a). Once the lighting is turned on, the view suffers overexposure and partially darker areas due to shadows (Fig. 7b).

If we only activate the occlusion mask (Fig. 1b top, Fig. 7c), we can see that overexposed area is reduced and can see the detail of the edge of the dish. By comparing the histogram with that of 7b, we can quantitatively confirms that the overexposed area in a scene is shrunk to 1/15th. This clearly shows that the LCD layer successfully dimmed overexposed area selectively.

On the other hand, when we only activate the HMD screen (Fig. 1b bottom, 7d), the histograms do not tell much about the effect. We can, however, qualitatively see that the shadow area, which is at right to the marker pen, is compensated by an additive image overlaid by the HMD. This allows us to see some fake yellow grains brighter. Unfortunately, the overlaid image lost detail due to the different focal lengths of the camera (30cm) and the HMD screen (2m). We discuss this accommodation issue in the discussion section.

Finally, if we combine the LCD and the HMD (7e), we can see both effect from the LCD-only and HMD-only conditions above.

5.3 Dynamic Capture Result

We now evaluate a realtime response of the current proof-of-concept system. In this experiment, we recorded UC image sequences with the four conditions. For each recording sequence, we start without the extra lighting, then we turn on the external lighting, and finally turn off. The lighting is controlled by hand.

In this setup, we defined the overexposure value as that radiance larger than 235 and underexposure lower than 20. This is to see more general tendency of our system.

Figure 9 shows a time-domain analysis of our adaptation assistance system in terms of the over-/under-exposure area ratios. The images in the figure are from the condition C4 {LCD:on, HMD:on}. To align each recording on the timeline in the figure, we detected

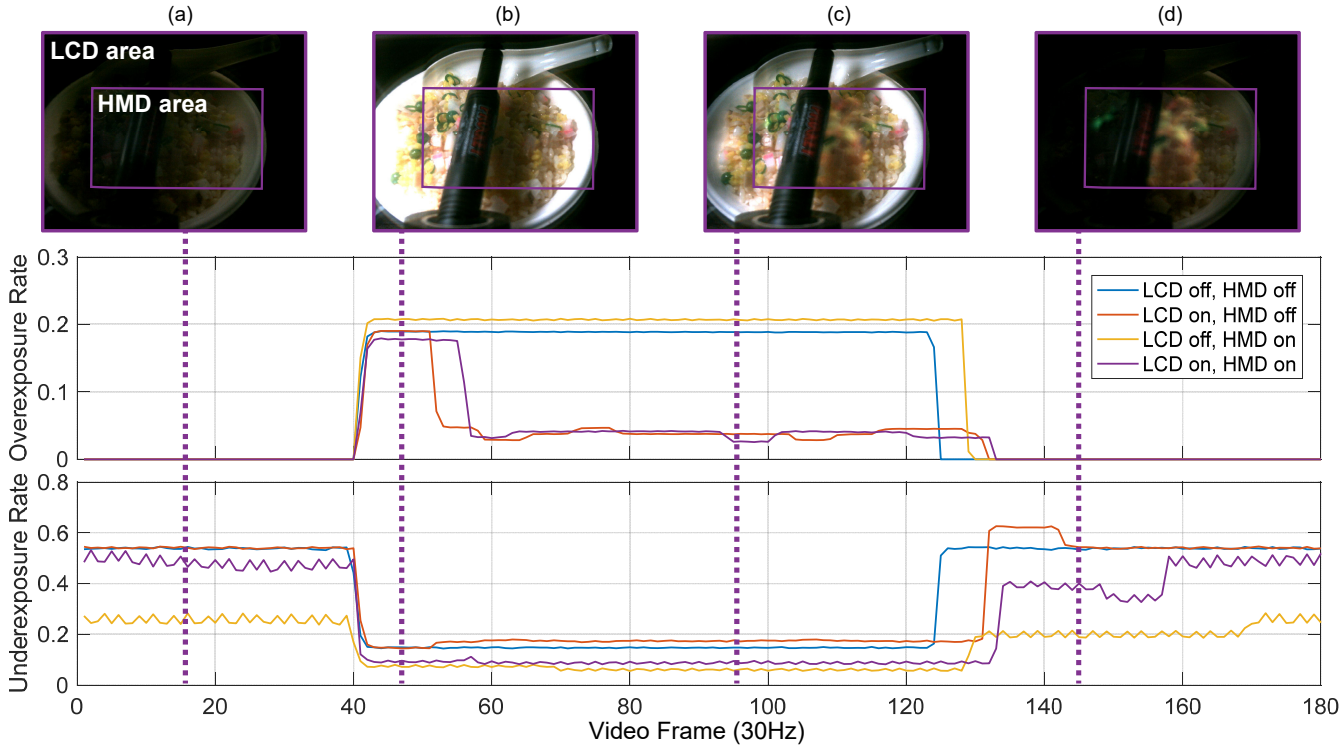

Figure 9: Overview of the realtime experiment in Sec. 5.3.

the timing when the light was turned on by checking a jump of overexposed-pixel count in each sequences.

In the cases with the HMD turned on (i.e. C2 and C4), the graph lines are jittering at the figure of the under-exposure rate in Fig. 9. As well, the enhanced visuals seemed to be jittering in the recorded UC image sequences. The jittering is caused by the difference of refresh rates between HMD and UC, thus the jittering will not be occurred when we see the HMD with the naked eye.

At Fig. 9 (b), the system is not yet able to provide the mask and the compensation image, thus the view is not optimized. This causes a flickering effect since the system will soon start assisting adaptation.

At Fig. 9 (c), our adaptation visualization is shown and we can see the overexposure is suppressed to about 25% in terms of its area size. The underexposure area also decreases about 50% between conditions with the same HMD state, i.e. C1 vs C3, and C2 and C4.

Through this dynamic capture setup, we can see that the proof-of-concept system can handle both over-/under-exposure properly and can assist our brightness adaptation. On the other hand, we observed that a low-latency system is necessary to reduce flicking effects (Fig. 9 (b)) caused by a delay to suppress overexposure. We further discuss this latency issue in the discussion section.

5.4 Preliminary Trial by Real Users

After removing the UC, two real users tried out the proof-of-concept system, and we let them mention the impression of the system. In the feedback, the users first realized that the HMD part clearly enhances darker areas on the object. We further decreased the base brightness setting of the HMD by the requests of the users, and one user mentioned, "The color on the HMD blended seamlessly into the background (dark) object".

For the LCD mask, on the other hand, the users initially could not clearly recognize that the LCD was dimming the environmental light. This is because the mask got stronger blur when seeing with the real eyes than with the UC. The users, however, easily recognized the brightness reduction when the LCD was switched on and off. We speculate that the change of the blur size stems from the difference of that of the aperture of the eye, i.e. pupil, and of the UC. This question follows to our discussion sections (Sec. 6.4).

6 DISCUSSION

In our current system, there are various assumptions and simplifications. In order to encourage the development of a more practical adaptation assistance system in future, we organize these issues and discuss on future research directions below.

6.1 Spatial Calibration with Eyes

Analogous to the spatial calibration of OST-HMD in Augmented Reality (AR) [16, 32], a practical system must track the current 3D position of the eyes with respect to the display system. In our current proof-of-concept implementation, we achieved the pixel-wise mappings between the image layers and the user-perspective camera by LUTs since we can directly get user-perspective images. This is not true for real users.

Fortunately, we can still model the mappings as perspective projections by treating the system as off-axis pinhole cameras where

the layers are image planes and the eye is the camera center. We can thus compute the projection by tracking the eye position. Such automated display calibration with wearable eye trackers has been a hot topic in recent AR research areas [13, 26].

Since the LCD and the HMD's waveguide causes optical distortion, we would also need to calibrate optical aberration to make a predistorted images [14].

6.2 Low-latency rendering

Our current system has a delay roughly about 500msec. This is due to the latency of the screens and the computation loop to calculate $\mathbf{I}_E$ and $\overline{\mathbf{I}_O}$. Although we leave this latency issue to other dedicated works, this is a must-be-solved issue to realize a practical AdaptiVisor system. As we introduced in the related work section, hardware systems optimized for low-latency rendering in the computational photography field [2, 30] are worth exploring.

6.3 Simulating Eye Adaptation

Even if we have a calibrated user-perspective camera, at the end of the day, a practical system is used by a real user. We thus require a way to estimate the current state of the eye's brightness adaptation in realtime so that the system can determine the optimal mask darkness.

Furthermore, the system relies on a scene camera to analyze the scene brightness and thus to determine the mask shapes and darkness. This requires estimating color distortion between the user and the scene camera. In other words, we have to correct the color of an image captured by the scene camera to be consistent with the visual stimulus that the user receives from the scene. The same problem occurs to the HMD screen color, which has been explored in the HMD color calibration context [12, 19]. It is also desirable to use a high dynamic-range camera for the SC so that the system can handle various lighting conditions.

A fundamental difficulty in these issues is that the sensitivity of rod and cone cells of the eyes are different from that of image sensors. We thus would need a perceptually correct color rendering [11] and a way to estimate the current state of the eye.

Another potential problem is the optical property of the transmissive LCDs. Due to their transparent TFT circuits, such LCDs have diffraction and make the see-through images blurred as we can see in figures in our paper. Integral imaging displays might be one solution to reduce this optical degradation while keeping the system compact [35].

6.4 Occlusion Mask Optimization

For realizing a consistent occlusion effect, we have to determine the shape of occlusion patterns so that they only and completely cover undesired light rays to be blocked. Our single-layer occlusion LCD, however, is too close to the user's eyes, and the occlusion patterns normally appear blurred to the eyes that focus at a distant object in a scene. The aperture of the eye, i.e. the pupil, predominantly causes this focus blur. The larger the pupil diameter is, the stronger the blur effect will be.

The blur can also be modeled as a point spread function or its Fourier transform, optical transfer function (OTF). Watson et al. proposed a model of OTF as its mean real function, i.e. radial modulation transfer function for the best-corrected human eye as a function of pupil diameter [33]. Bursky et al. explored vision-realistic rendering by using wavefront data of an actual human [4]. Given such function, we can apply image deconvolution on occlusion masks so that their projected retinal images appear sharp [23].

The same blur effect could occur for OST-HMDs since most existing displays only have a single focus plane. Recent light-field displays can mitigate this issue by displaying volumetric focusible images. Retinal scanning displays can also solve this phenomenon since such display ignores the crystalline lens and can directly project image onto the retina.

Furthermore, as shown in Fig. 5, non-linearity of the transmittance property of the LCD mask is also needed to be calibrated.

6.5 Long-term Use in Daily Life

The proposed system has concerns that the artificial adaptation might have bad effect to the natural adaptation of the eyes. We consider that the system might not degrade the natural adaptation by the photoreceptors, because the system only assists the excessive environment light to fit within the original dynamic range of the eyes. However, as a long-term goal, we would need to evaluate the actual influence of the system on the eyes quantitatively.

7 CONCLUSION AND FUTURE WORK

We presented AdaptiVisor, a proof-of-concept vision augmentation system for brightness adaptation assistance. Our system consists of an occlusive optical see-through HMD and a scene camera which is optically aligned to the viewpoint. For occlusion, we employ a transmissive LCD rigidly attached in front of an OST-HMD.

Our proof-of-concept system with a user-perspective camera as a viewer demonstrates that it enables to selectively and dynamically suppress overexposed areas in the scene by the LCD mask. At the same time, the system can enhance underexposed areas by showing additive image on the screen. We use a highlight detection method to determine which area in the field of view (FoV) to be masked and which area of the FoV to be enhanced by the HMD. For displaying the mask and the image correctly in the user-perspective camera's FoV, we spatially calibrated the LCD, HMD, and the user-perspective camera by using homography and gray-code patterns.

Our evaluations based on the static and dynamic scene capture show that our system was able to suppress the overexposed area to 1/15th compared to raw FoV. It also improved the underexposed area by correctly overlaying the image. Preliminary user trial with two real users indicates that the system also works for real eyes for the HMD part and to some extent for the LCD, and suggests that optimizing occlusion mask shape for a given user's eye is a key issue.

Discussion followed then gave thorough discussion on steps towards practical adaptation assistance systems. The discussion includes eye simulation and occlusion mask optimization.

Acknowledgments

The work was partially supported by JSPS KAKENHI Grant Number JP16H07169 and CREST, JST.

REFERENCES

[1] Dorin Aiteanu, Bernd Hillers, and A Graser. 2003. A step forward in manual welding: demonstration of augmented reality helmet. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*. IEEE, 309–310.

[2] Toshiyuki Amano and Hirokazu Kato. 2010. Appearance control by projector camera feedback for visually impaired. In *CVPRW*. IEEE, 57–63.

[3] Howard Dehaven Baker. 1949. The course of foveal light adaptation measured by the threshold intensity increment. *JOSA* 39, 2 (1949), 172–179.

[4] Brian A Barsky. 2004. Vision-realistic rendering: simulation of the scanned foveal image from wavefront data of human subjects. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*. ACM, 73–81.

[5] Kiriti Bhagavathula, Albert H Titus, and Christopher S Mullin. 2007. An extremely low-power CMOS glare sensor. *IEEE Sensors Journal* 7, 8 (2007), 1145–1151.

[6] Ozan Cakmakci, Yonggang Ha, and Jannick P Rolland. 2004. A compact optical see-through head-worn display with occlusion support. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 16–25.

[7] John E Dowling. 1987. *The retina: an approachable part of the brain*. Harvard University Press.

[8] Chunyu Gao, Yuxiang Lin, and Hong Hua. 2013. Optical see-through head-mounted display with occlusion capability. *Proceedings of SPIE Defense, Security Sensing–Head and helmet-mounted displays XIX: design and applications* 8735 (2013), 87350F.

[9] Charles Haig. 1941. The course of rod dark adaptation as influenced by the intensity and duration of pre-adaptation to light. *The Journal of general physiology* 24, 6 (1941), 735–751.

[10] Takenori Hara, Hideo Saito, and Takeo Kanade. 2009. Removal of glare caused by water droplets. In *Visual Media Production, 2009. CVMP'09. Conference for*. IEEE, 144–151.

[11] Piti Irawan, James A Ferwerda, and Stephen R Marschner. 2005. Perceptually Based Tone Mapping of High Dynamic Range Image Streams.. In *Rendering Techniques*. 231–242.

[12] Yuta Itoh, Maksym Dzitsiuk, Toshiyuki Amano, and Gudrun Klinker. 2015. Semi-Parametric Color Reproduction Method for Optical See-Through Head-Mounted Displays. *IEEE transactions on visualization and computer graphics* 21, 11 (2015), 1269–1278.

[13] Yuta Itoh and Gudrun Klinker. 2014. Interaction-free calibration for optical see-through head-mounted displays based on 3D eye localization. In *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI), Minneapolis, MN, USA, March 29-30, 2014*. 75–82. DOI:http://dx.doi.org/10.1109/3DUI.2014.6798846

[14] Yuta Itoh and Gudrun Klinker. 2015. Light-Field Correction for Spatial Calibration of Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)* 21, 4 (April 2015), 471–480. DOI:http://dx.doi.org/10.1109/TVCG.2015.2391859

[15] Yuta Itoh and Gudrun Klinker. 2015. Vision enhancement: defocus correction via optical see-through head-mounted displays. In *Proceedings of the 6th Augmented Human International Conference*. ACM, 1–8.

[16] Adam L Janin, David W Mizell, and Thomas P Caudell. 1993. Calibration of head-mounted displays for augmented reality applications. In *Virtual Reality Annual International Symposium*. IEEE, 246–255.

[17] Kiyoshi Kiyokawa. 2012. Occlusion Displays. In *Handbook of Visual Display Technology*. Springer, 2251–2257.

[18] Kiyoshi Kiyokawa, Mark Billinghurst, Bruce Campbell, and Eric Woods. 2003. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 133.

[19] Tobias Langlotz, Matthew Cook, and Holger Regenbrecht. 2016. Real-Time Radiometric Compensation for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (2016), 2385–2394.

[20] Chao Ma, Minoru Taya, and Chunye Xu. 2008. Smart sunglasses based on electrochromic polymers. *Polymer Engineering & Science* 48, 11 (2008), 2224–2228.

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

[22] Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke, and Henry Fuchs. 2014. Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources. In *ACM SIGGRAPH 2014 Emerging Technologies*. ACM, 20.

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

[24] Shree K Nayar and Vlad Branzoi. 2003. Adaptive dynamic range imaging: Optical control of pixel exposures over space and time. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 1168–1175.

[25] Joseph Newman, Martin Wagner, Martin Bauer, Asa MacWilliams, Thomas Pintaric, Dagmar Beyer, Daniel Pustka, Franz Strasser, Dieter Schmalstieg, and Gudrun Klinker. 2004. Ubiquitous tracking for augmented reality. In *Mixed and Augmented Reality, 2004. Third IEEE and ACM International Symposium on*. IEEE, 192–201.

[26] Alexander Plopski, Yuta Itoh, Christian Nitschke, Kiyoshi Kiyokawa, Gudrun Klinker, and Haruo Takemura. 2015. Corneal-Imaging Calibration for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)* 21, 4 (April 2015), 481–490. DOI:http://dx.doi.org/10.1109/TVCG.2015.2391857

[27] Jun Rekimoto. 2012. Squama: modular visibility control of walls and windows for programmable physical architectures. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, 168–171.

[28] Pedro Santos, Thomas Gierlinger, Oliver Machui, and André Stork. 2008. The daylight blocking optical stereo see-through HMD. In *Proceedings of the 2008 workshop on Immersive projection technologies/Emerging display technologiges*. ACM, 4.

[29] Hui-Liang Shen and Zhi-Huan Zheng. 2013. Real-time highlight removal using intensity ratio. *Applied optics* 52, 19 (2013), 4483–4493.

[30] Robert Tamburo, Eriko Nurvitadhi, Abhishek Chugh, Mei Chen, Anthony Rowe, Takeo Kanade, and Srinivasa G Narasimhan. 2014. Programmable automotive headlights. In *European Conference on Computer Vision*. Springer, 750–765.

[31] Eric W Tatham. 1999. Technical opinion: getting the best of both real and virtual worlds. *Commun. ACM* 42, 9 (1999), 96–98.

[32] Mihran Tuceryan and Nassir Navab. 2000. Single Point Active Alignment Method (SPAAM) for Optical See-Through HMD Calibration for AR. In *Proceedings of ISAR*. IEEE, 149–158.

[33] Andrew B Watson. 2013. A formula for the mean human optical modulation transfer function as a function of pupil size. *Journal of Vision* 13, 6 (2013), 18–18.

[34] Andrew B Watson and John I Yellott. 2012. A unified formula for light-adapted pupil size. *Journal of vision* 12, 10 (2012), 12–12.

[35] Yuta Yamaguchi and Yasuhiro Takaki. 2016. See-through integral imaging display with background occlusion capability. *Applied optics* 55, 3 (2016), A144–A149.