Real-Time Radiometric Compensation for Optical See-Through Head-Mounted Displays

Tobias Langlotz, *Member, IEEE*, Matthew Cook, and Holger Regenbrecht *Member, IEEE*

Fig. 1. Overview of the results of our system allowing for real-time radiometric compensation for optical-see through head-mounted displays. (Left) Head-mounted display prototype utilizing a beam-splitter to capture the image as seen by the user allowing for the computation of a pixel-precise compensation image. (Middle) Naïve overlay as in standard head-mounted displays showing color artifacts caused by color-blending between the background (here a color chart) and the displayed image (seen through a Epson Moverio BT-100). (Right) Our solution mitigates the effect of color blending by applying a pixel-precise radiometric compensation (seen through the same device but using our radiometric compensation). Both images are cropped to only show the area covered by the display. The small inlay image shows the desired image.

Abstract—Optical see-through head-mounted displays are currently seeing a transition out of research labs towards the consumer-oriented market. However, whilst availability has improved and prices have decreased, the technology has not matured much. Most commercially available optical see-through head mounted displays follow a similar principle and use an optical combiner blending the physical environment with digital information. This approach yields problems as the colors for the overlaid digital information can not be correctly reproduced. The perceived pixel colors are always a result of the displayed pixel color and the color of the current physical environment seen through the head-mounted display. In this paper we present an initial approach for mitigating the effect of color-blending in optical see-through head-mounted displays by introducing a real-time radiometric compensation. Our approach is based on a novel prototype for an optical see-through head-mounted display that allows the capture of the current environment as seen by the user's eye. We present three different algorithms using this prototype to compensate color blending in real-time and with pixel-accuracy. We demonstrate the benefits and performance as well as the results of a user study. We see application for all common Augmented Reality scenarios but also for other areas such as Diminished Reality or supporting color-blind people.

Index Terms—Radiosity, global illumination, constant time

✦

1 INTRODUCTION

Head-Mounted Displays (HMDs) have a long history in research. In fact the first prototype was developed by Ivan Sutherland in 1968 by introducing his prototype of a head-mounted three dimensional display [19]. This prototype, later also known by the name "Sword of Damocles", presented the user computer generated graphics overlaid onto the environment. Sutherland used beam splitters such as half-silvered mirrors in the user's view to reflect the image of small cathode ray tubes displaying the computer generated image. This basic concept is now known as *optical see-through HMDs* (OHMDs). With his Sword of Damocles prototype, Ivan Sutherland did not only invent HMDs but also laid the foundation of what nowadays is known as *Augmented Reality* (AR). Since then AR has evolved and also used other display

hardware, most notably video-see through displays, to achieve the goal of creating mixed reality displays.

While optical-see through HMDs are commonly used in research, it is not until recently that OHMDs have also gained traction in industry and been brought to the consumer market. The main driving factors were announcements of optical-see through devices such as Google Project Glass[1], Epson with the Moverio series[2] and more recently Microsoft Hololens[3], amongst other solutions. With these recent developments, we move closer to a continuous augmentation using optical-see through HMDs as envisioned in the concept of Pervasive Augmented Reality [8].

1.1 Problem Statement

Common to all these optical-see through HMDs is that they follow a similar design as Ivan Sutherland's first prototype. A beam splitter, usually a half-transparent mirror or optical prism, is used as an optical combiner to blend the environment with a computer-generated image. In modern devices the computer-generated image is displayed not using a CRT but a small projector or LCD screen. Consequently, current optical-see through HMDs also share many of the issues of the early prototypes. One is the required eye-display calibration to allow a pre-

- *Tobias Langlotz is with the University of Otago. E-mail: tobias.langlotz@otago.ac.nz.*
- *Matthew Cook is with the University of Otago. E-mail: cooma376@student.otago.ac.nz.*
- *Holger Regenbrecht is with the University of Otago. E-mail: holger.regenbrecht@otago.ac.nz.*

[1] www.google.com/glass
[2] www.epson.com
[3] www.microsoft.com/microsoft-hololens

cise registration of the overlaid graphics [12]. The other big problem is the correct reproduction of colors in optical see-through HMDs [7]. The use of the optical combiner causes the color of each displayed pixel to interact with the color of the environment at the corresponding position [20]. This makes it impossible for the commonly used designs to display black in everyday environments, as well as affecting all other displayed colors [7].

In this paper we introduce an approach that minimizes the effect of blending of the background and the displayed image in a way that the user of the OHMD sees the colors as the application designer intended. Our approach introduces modifications to the design of commonly used optical-see through HMDs by adding a beam splitter, allowing us to virtually place a camera at the position of the user's eye. This allows us to capture an image of the environment as seen by the user and correct the displayed image for each pixel and in real-time. The result is an image that when displayed in the OHMD blends with the environment and better resembles the intended image.

Overall, this paper contributes to the disciplinary knowledge by the provision of a novel hardware prototype for an OHMD allowing to capture the environment as seen by the user. This prototype is fundamental for our later radiometric compensation but also for many other research and application areas. We will present several design iterations of our prototype that are optimised for quality of the radiometric compensation by utilizing hardware with a larger form factor but providing higher quality. We further present three different algorithms for radiometric compensation of OHMDs. These algorithms provide either a global solution, an adaptive perception driven solution to reduce color clipping at costs of changing the color space of the input image, or a combination of both. We present the results that can be achieved with these approaches as well as the results of a user study reporting on the perceived overall quality.

1.2 Limitations

Our new approach for radiometrically correcting OHMDs in real-time has some limitations. Foremost, using our approach we are still not able to produce colors that are darker than the environment. We'll show that we can mitigate the effect by using characteristics of human perception. However, this comes at cost of changing the displayed image content. Furthermore, we developed several prototypes but did not focus on miniaturization yet. The results presented in this paper come from a prototype that is still too large in size to be practically worn. However, all the used hardware components (beam splitter, cameras) can easily be miniaturized and integrated into HMDs without further increasing the HMD's size and we show an early prototype for this. Finally, we did not focus in this work on eye-display calibration which is required to achieve a pixel-precise registration outside of lab environments. However, there are existing approaches which could be integrated into our approach [12].

2 RELATED WORK

Our work combines two research fields: General work on improving OHMDs and work on color blending and correct color reproduction. The latter works are not only driven by OHMDs but also by other fields in AR most notably spatial AR systems. In the following we present the key papers for both research directions.

2.1 Optical See-through Head-Mounted Displays

As stated previously, OHMDs can be traced back to Sutherland's early work using half-silvered mirrors to reflect the image of small cathode ray tubes, blending the physical world with digital data [19]. Since then several design iterations have been introduced in research and commercial products that improve the original design by using different display technologies instead of the initial CRT displays or introduced different optical elements to optically overlay the digital data. A good overview of the various works in that direction is presented by Rolland and Hua [17]. Nowadays, most available OHMDs follow a similar design that blends the environment light with a digital image displayed on an integrated display using a half-mirror (see Figure 2).
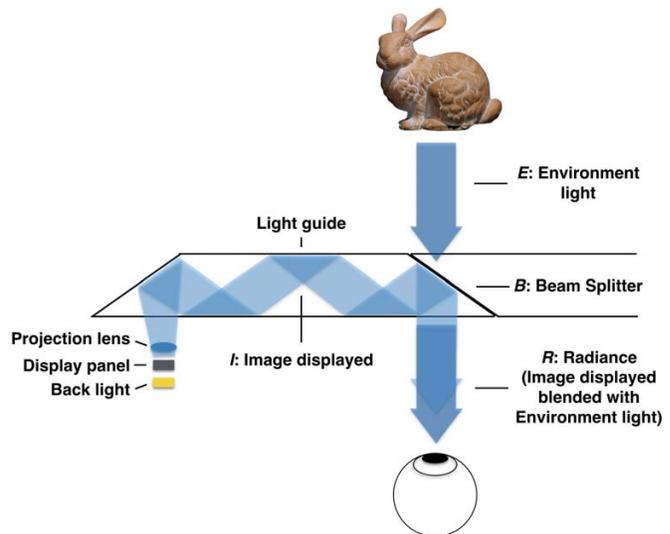

Fig. 2. Illustration of the basic functionality of a Optical see-through Head-Mounted Display.

Despite the long history of OHMDs there are still several limitations that affect the quality of the possible results and consequently the adaption of OHMD technology. One of the limitations is the need for continuous eye-display calibration to guarantee an accurate overlay of the digital data onto the physical world [12, 15]. Besides this spatial coherence, color coherence is an important issue. Due to the optical design it is not possible to correctly reproduce colors as the perceived result is a combination of blending the digital image with the environment light as already described by Gabbard et al. [7]. While there exist solutions for completely blocking the environment light using spatial light modulators and depth cameras for creating correct occlusions [13, 4], correct color reproduction has remained an unsolved challenge.

2.2 Color Reproduction in Augmented Reality

The challenge of correct color reproduction is not unique to OHMDs. In fact, spatial AR systems using projectors to augment the environment face similar problems. Here the projected color blends with the color of the surface projected onto. Bimber et al. were first to demonstrate that the effect can be mitigated by applying a radiometric compensation [1, 2]. They capture the surface information and compute a pixel precise compensation image to neutralize the effect of color blending. Grundhöfer and Bimber later showed an adaptive approach adapting the augmentation to reduce clipping errors caused by the constrained dynamic range of the projector [9]. Both works can be seen as fundamental to the approach described in this paper and will be discussed in more detail later.

Another research direction that addressed color blending is when overlaying transparent images as often found in video see-trough AR approaches. Fukiage et al. showed that the visibility of transparent objects is dependent on color and texture of the background scene [5]. They presented an approach for blending the image overlay with the background image to offer constant visibility but assume always half-transparent overlays.

Despite the work in spatial AR systems and video see-through AR and the fact that many research papers note the issue of color blending in OHMDs (e.g., ARQuake in 2002 [20]), there is only limited work in the fields of both quantifying the issue and working to correct it. Gabbard et al. quantified the interaction between environment light and the image projection within an OHMD [7, 6]. Here they outlined a mathematical description of the optical pathway of blended light before it is viewed by a user and showed an experiment attempting to calculate the effects of inputs to this optical system on the output. However, they

did not focus on correcting the effects resulting from color blending but only demonstrated the results caused by different backgrounds.

Apart from quantifying and demonstrating the effect of color blending there have been also attempts to improve the color reproduction when using OHMDs. Itoh et al. have shown that even against a solid black background the color reproduction of an OHMD can be greatly improved from the out-of-the-box experience by creating and applying a color display profile [11]. Sridharan et al. also created a color profile for correcting OHMD rendering, and additionally they have proposed a method for creating profiles for multiple different background colors and showed that this does improve the perceived colors in an OHMD [18]. However, both approaches did not measure the backgrounds and mainly target the color rendering of the display itself and do not focus on solving the effects of color blending. Furthermore, these techniques are applied globally and cannot solve color blending effects on a per-pixel level.

Two research groups have attempted to apply blending techniques on a per-pixel basis. Hincapi-Ramos et al. presented their Smart-Color system [10] focusing on adapting the 'binned profile' method described by Sridharan et al. [18] to an algorithm that can be run per-pixel in real-time. They succeeded in this goal, but only demonstrated it in a simulated environment that is, no camera was used to collect environment information nor was an OHMD used to display results. Furthermore, the fundamental problem of how to capture the pixel-precise color information of the environment and align it with the display which is needed for actual applications remained unanswered. For this reason it is not possible to determine the extent to which their method corrects blending errors on real hardware and their results are of a more theoretical nature.

An attempt at per-pixel blending correction was undertaken by Weiland et al. [22]. To our best knowledge, this approach is the only one capturing the environment with a camera and uses the camera feed to compute a compensation image for neutralizing the effect of color blending. The primary drawback of their method is that the camera was mounted above the OHMD, and due to the eye and camera not being aligned, the per-pixel correction only worked for background objects at a very specific distance, not on arbitrary background scenes. Their approach also uses a simple subtractive color model and did not deal with the limited dynamic range of the display in the OHMD. Furthermore, they did not evaluate their results or the performance.

Overall, one of the primary unanswered research questions is the pixel precise radiometric compensation in OHMD. While the existing literature shows a number of techniques in isolation (e.g, simulating color blending, display calibration for uniform colored backgrounds) none go so far as to propose or implement a complete system composed of hardware and software for correcting color-blending per-pixel and in real-time. These unknowns provided the initial basis for our research direction. Furthermore, the existing results are usually presented as a visual comparison to the reader and no formal evaluation results were presented.

3 COLOR BLENDING IN OPTICAL SEE-THROUGH HEAD-MOUNTED DISPLAYS

Presently, most OHMDs - such as Google Glass or the Epson Moverio series - are using similar approaches. A small display panel integrated into the HMD's frame shows the image to be overlaid onto the environment (see Figure 2). The displayed image is projected through one or several lenses enlarging the virtual image. Depending on the construction of the OHMD, the virtual image is reflected using a light guide before it hits a half-transparent mirror where it is reflected into the users eye (see Figure 2). The beam splitter (e.g., half-transparent mirror) is not only reflecting the light forming the displayed image but also passes the light from the environment to the user's eye. The user therefore perceives the displayed image blended with the environment light. The blending function is thereby dependent on the used beam splitter and can be considered constant for a particular OHMD. The light blending in OHMD was formalised by Gabbard et al. [7] as

$$L_4 = AR_D(L_3, RF(L_1, B)) \tag{1}$$

Here L_4 is the light that reaches the user's eye and AR_D represents the characteristics of an OHMD display system and is parametrized not only by the characteristics of the used beam-splitter (e.g., how environment light and the displayed image are blended) but also by factors influencing how a digital image is displayed (e.g., display brightness). The displayed image is represented as L_3 while $RF(L_1, B)$ describes the incoming environment light which is dependent on a reflectance functions RF describing how the light L_1 from a light source in the environment interacts with the the background B before hitting the OHMD [7].

A similar notation can be derived from the field of projector-camera systems [2, 3] and is used for the remainder of this work. By taking into account the beam splitter B blending the environment light and the displayed light we can formalise the system as

$$R = t_B E + IF r_B \tag{2}$$

In this equation the perceived radiance R is the blended light as perceived by the user of the OHMD. The incoming environment light E is based on the environmental light source illuminating objects with certain material properties in the user's view but as it is not easily possible for us to know these terms individually so the term can be treated singularly as environment light entering the OHMD and we write here E. The term $t_B E$ is the environment light transmitted through the beam splitter B which is part of the OHMD (see Figure 2). The amount of transmitted light depends on the used beam splitter and and its light-transmissive factor t_B (e.g, 0.5 for half-transparent mirrors). The term I describes the radiance of the displayed image. The form factor F of the device describes the effects of varying image intensities across the entire display surface; for example, the projected brightness of a pixel falls off at the edge of the display due to vignetting. The reflected light depends on the reflective factor r_B of the beam splitter used as part of the OHMD. Similarly to environment light, it is difficult to separate these last two terms such that they are generally considered together as $F r_B$.

From the equation above the cause of the blending problem is simple to isolate: We can measure $F r_B$ and t_B, and the projector intensity I is under our control, but as long as E remains unmeasured and uncompensated this environment light will affect the result perceived by users. One possible, though very simplistic solution is to simply design very dark glasses affecting t_B, such that $t_B E \approx 0$. Of course, blocking all environment light in this way is not at all suitable for AR applications.

As we will show during this work, from equation 2 and the measured $F r_B$ we can infer that if one is also able to determine $t_B E$ we can compute the display image I in a way that it results in the desired visible radiance R. This also requires that we are able to compute this for each pixel. We therefore had to first solve the problem of accurately capturing the radiometric information on the environment E before we are able to compensate the effect of color blending within the OHMDs. Unfortunately, using the camera often integrated in OHMD (e.g., such as Google Glass) does not solve our problem as the perspective is different to the one a user has when wearing the OHMD and looking through the display and would only work for objects in a fixed distance range [22].

4 PROTOTYPE FOR A RADIOMETRIC COMPENSATED OPTICAL SEE-THROUGH HEAD MOUNTED DISPLAY

The key idea of our novel OHMD hardware setup is to virtually place a camera at the same position as the user's eye. This allows us to capture an image of the environment as seen by the user. This is crucial for computing a compensation image that will reduce the effect of color blending. We achieved the goal of placing a virtual camera in the user's eye by using an optical beam splitter B_1 that is placed in front of the optical combiner B_2 used for blending the digital graphics with the environment (see Figure 3, left). By doing so we can capture the environment without the effect of the OHMD's display. Unfortunately, placing the camera via the beam splitter at the position of the users eye does not give us a pixel precise mapping yet. This would require an eye display calibration which is required in any case for precisely
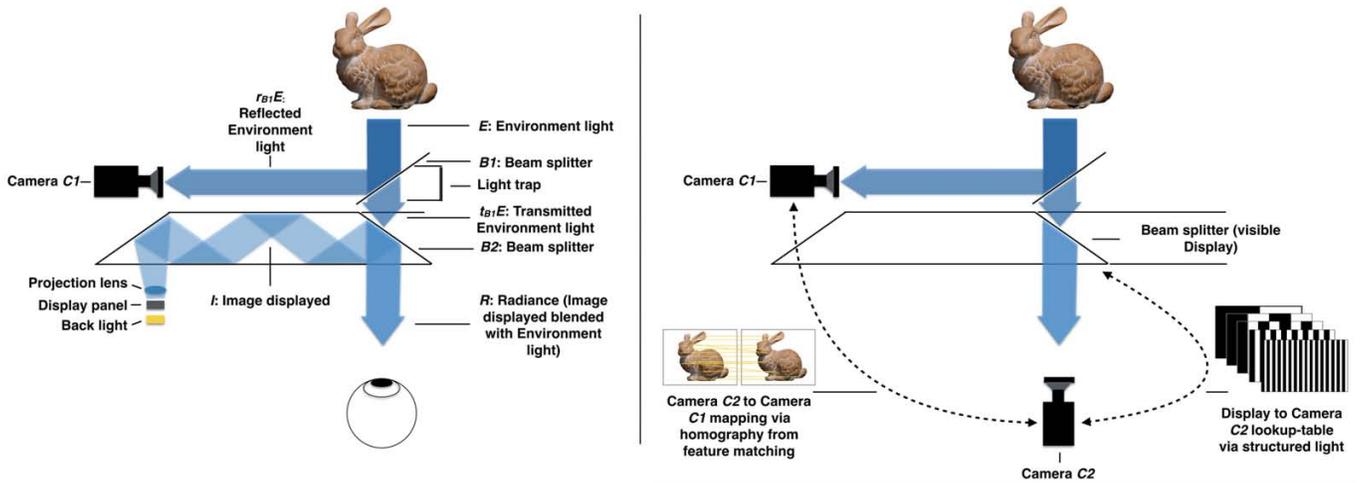
Fig. 3. Schematic overview over our OHMD prototype. (Left) The addition of a beam-splitter allows to partially reflect the incoming environment light towards the camera $C1$. (Right) To allow for a pixel-precise mapping between the camera $C2$ (representing the users eye for our tests) and camera $C1$ we apply a camera display calibration using structured light and a camera-to-camera mapping using a homography.

overlaying graphical content in an OHMD [12, 15] as well as to correlate the display with the camera which can be done in hardware. In our case we replace the users eye with a camera to be able to capture the results of our approach. In the following we describe how we calibrate that overall setup to guarantee a pixel precise mapping.

4.1 Calibration

To be able to precisely compensate a pixel's color we need to create a pixel precise mapping that maps an observed camera pixel in camera $C1$ to the corresponding display pixel so that when perceived by the user (in our case camera $C2$) they exactly align. We achieved this by first calibrating the camera's intrinsic parameters, then creating a pixel-precise lookup-table mapping OHMD display pixels to corresponding pixels in camera $C2$. Finally, we compute a transformation that maps each pixel from camera $C1$ to $C2$ using a homography.

Before calibrating the camera intrinsic, we adjusted both cameras to match their field of view which is relatively easy given that they use the same camera module and lens attached. We then used OpenCV to compute the cameras intrinsic parameters, however, we compute for each camera $C1$ and $C2$ the intrinsic parameter of the camera with the attached lens and in addition the intrinsic parameter of the cameras with all optical elements. These are the attached lens ($C1$ and $C2$), the beam splitter $B1$ ($C1$ and $C2$) and the optical combiner forming the display $B2$ (only $C2$).

Once calibrated, we compute a pixel-precise lookup-table mapping pixels in the display to corresponding pixels in camera $C2$. We created this lookup table using a structured-light approach displaying a gray code in the display which is observed via the camera $C2$ to compute the lookup table based on the captured images. The computation is only needed once ahead of time as the configuration of the two cameras does not change (see Figure 3, right).

We also compute a transformation that maps each pixel from camera $C2$ to $C1$ and vice versa. Both cameras have the same field of view and we manually aligned their position so that they appear to see the environment from the same position via the half-silvered mirror. To achieve pixel-precision, we need to take care of the small rotational offset between both cameras after the manual alignment. We do this by computing a homography between both cameras by matching image features in the captured environment (see Figure 3, right). Again this is only needed once as the configuration remains the same.

4.2 Implementation

During several design iterations, we came up with several prototypes that implement our concept. Our initial prototype is based on an Epson Moverio BT100 OHMD with older PointGrey Firefly cameras with 640x480 pixel resolution (see Figure 4, left). For this as well as the later prototypes we use a 50/50 beam splitter and rods from the Open-Beam project[4]. Some parts were custom made using a 3D printer. This prototype was used to verify our initial assumption but was later upgraded by replacing the cameras with PointGrey Blackfly cameras offering a higher resolution of 1296x1032 pixels as well as offering a larger dynamic range and color gamut. We also replaced the lenses attached to the cameras with zoom lenses allowing us to obtain an as large as possible image of the OHMD display without losing too much resolution to image cropping (see Figure 4, middle). This is the prototype used to produce the results presented in this paper but the higher quality comes at cost of size. We consequently built another prototype to demonstrate the possible option for miniaturization (see Figure 4, right). This prototype uses a smaller beam splitter reflecting 50% of the environment light towards a smaller camera which sits at the top. This however, should only demonstrate the potential for miniaturization and further possible optimizations.

Our implementation for calibrating the displays and cameras uses OpenCV 3.0 to display and compute the lookup-table and the homography. PointGrey Blackfly cameras claim to have a linear response curve so we did not apply further color linearization.

Overall, the prototypes allowed us to capture an image of the incoming environment light E reflected via the newly introduced beam splitter B_1 and to pixel-precisely align it with the displayed image. While we used camera $C2$ to simulate the user's eye, practical solutions would work without this camera but would need to consider an eye-display calibration which is required in any case for precisely overlaying graphical content in an OHMD [12, 15].

5 REAL-TIME RADIOMETRIC COMPENSATION FOR OHMD

Once we are able to map pixels from the camera image $C1$ to the user's eye (represented by camera $C2$) and to the display, we can compute a compensation image compensating the effect of the environment light during the color blending. In the following we present three approaches for computing the compensation image. The first one computes a correction while not considering the dynamic range of the display. Thus this approach can lead to color clipping. We follow
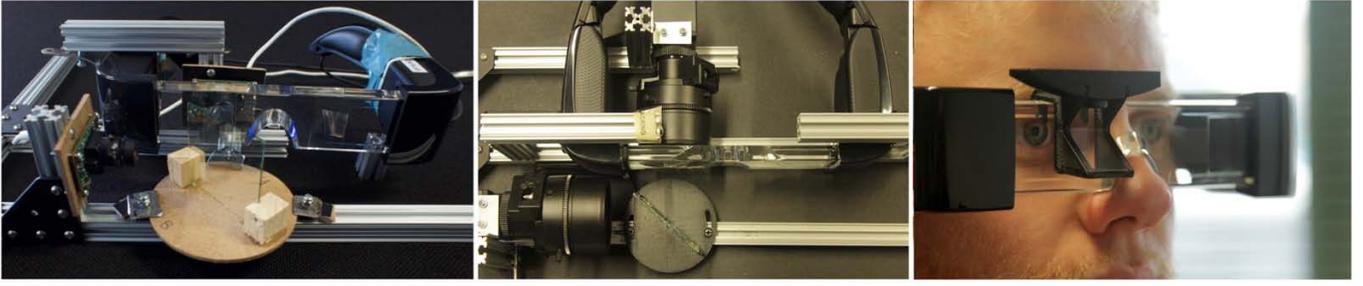
---

[4]www.openbeamusa.com

Fig. 4. Design iterations of our developed prototype. Note that the current setup is monoscopic but can easiy be extended to also cover the other eye. (Left) Front-view of the first prototype using a beam-splitter to capture the radiometric information of the environment. The system is based around an Epson Moverio BT-100 OHMD. (Middle) Top view of the second prototype using the final hardware as used for producing the results in this paper. Note the cameras and lenses offer better quality at the cost of size but allow for better capturing of the results. (Right) Miniaturized version of our proposed prototype design worn by a user. Here the image is not reflected to the side but towards the camera sitting in the top frame.

up with an approach that adapts the overlaid image and consequently the compensation image to reduce clipping artifacts. This adaption, however, leads to usually brighter images. We finally present another approach that also takes into account the dynamic range of the display but tries to not increase the image intensity of the desired image.

5.1 Global radiometric compensation

As stated earlier we use the mathematical notation derived from Bimber et al.'s work on projector-camera systems as a starting point for computing a compensation image [2, 3]. As stated in section 3 the visual result R of a basic OHMD is given by

$$R = t_B E + I F r_B \tag{3}$$

However, given our new prototype described in the previous section this equation is no longer valid as the addition of a beam splitter $B1$ in place affects the visible result R which we need to account for. This beam splitter adds two new parameters one for the transmissive factor t_{B1} and one for the reflective factor r_{B1} of $B1$. Extending equation 3 we can describe the system as

$$R = t_{B2}(t_{B1}E) + I F r_{B2} \tag{4}$$

where t_{B2} is the amount of light transmitted through the beam splitter $B2$. The camera $C1$ only sees the scene through the reflective portion of the beam splitter r_{B1} which leads to

$$C1 = r_{B1}E. \tag{5}$$

We can solve this for E which gives us

$$E = \frac{C1}{r_{B1}}. \tag{6}$$

Substituting equation 6 into equation 3 gives us

$$R = t_{B2}(t_{B1}\frac{C1}{r_{B1}}) + I F r_{B2} \tag{7}$$

We are not interested in computing the expected visual result R but instead want to provide a desired visual output and compute the input image I which delivers the desired output despite the effect of color blending as modeled in equation 7. Consequently solving equation 7 for I gives

$$I = \frac{R - t_{B2}(t_{B1}\frac{C1}{r_{B1}})}{F r_{B2}} \tag{8}$$

which is our equation to compute the compensated input image I. This equation is the driver for the radiometric compensation for color blending in OHMDs. We can measure the combined form factor and display material $F r_{B2}$ ahead of time by displaying a white image on the OHMD over a black background and measuring the per-pixel response

using camera $C2$. The form factor describes imperfections in the display (e.g., vignetting, internal diffraction). An example result can be seen in Figure 5.

Assuming perfect beam-splitters there is no internal reflection as well as the transmissive and reflective terms are valid for the all incoming light. However, we noticed deviations which we try to take care of. Similarly to the other combined terms noted above, the absolute measurement of t_{B1} and r_{B1} describing the characteristics of $B1$ were not required for the implementation of our prototype. Instead, we calculated the ratio $\frac{t_{B1}}{r_{B1}}$ by observing a test scene with both cameras $C2$ and $C1$ via the beam splitter, transmitted and reflected respectively, and used our calibrated homography to collect spatially equivalent point samples of color, followed by a linear regression. The same approach was used to determine characteristics of $B1$ which was modelled as a polynomial function determined by first observing the test scene with $C2$ directly, then again through the OHMD display prism with the display turned off.

Finally, we had to take care of the display characteristics. We assume the display to be linear as we work in linear RGB space but typical display output is non-linear srgb, we adjust for the gamma corrected curve of the final output value.

With these constants measured and the ahead-of-time calibration performed, the result of the global compensation algorithm can be seen in Figure 6. The displayed global compensation image takes into account the background (here a color chart) and tries to neutralize the effect of color blending. One limitation of this approach is the dynamic range of the display integrated in the OHMD. The display can only show values for I that are in the range of $[0, 1]$. However, de-
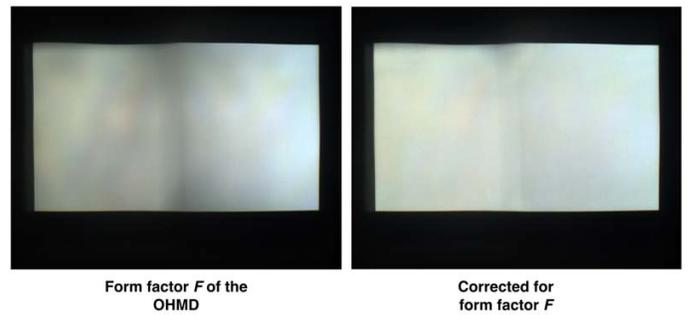

Fig. 5. Example of measuring and correcting the form factor F describing the light falloff in the OHMD caused by optical imperfections in the display: (Left) Displaying a white image in front of a uniform black background gives the form factor $F r_{B2}$. (Right) Compensating only for $F r_{B2}$ result in a more uniform image (here demonstrated again with a white image). Both images are not cropped and recorded through the eye camera $C2$.
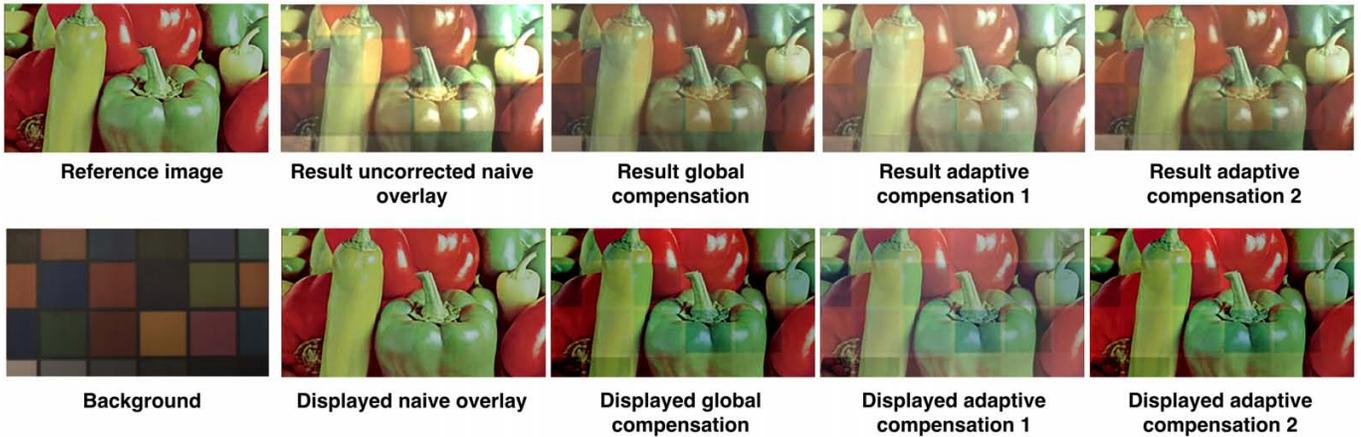
Fig. 6. Overview of the visual results of an uncorrected OHMD (naive overlay displaying the desired image) and our three radiometric compensation algorithms (global compensation, adaptive compensation 1, adaptive compensation 2). One can see that all three compensation algorithms improve the visible results and compensate for the background. However, the adaptive compensation 1 tends to increase the general brightness of the image to avoid clipping), All images are captured through the OHMD by camera $C2$ and are cropped to only cover the display area.

pending on the incoming light from the physical background, desired result, and form factor of the OHMD, it is possible that the algorithm computes an I which is outside this range. One example for this would be a relatively dark desired image R while having a very bright background E resulting in an I which needs to be negative to compensate the bright background. Basically, we would need to darken the background which is not possible with a standard OHMD hardware design. This shows the effective limit of this type of compensation. In the case of a bright background and a dark desired result ($I < 0$) we call this negative clipping. For dark environments, a bright desired result and low form factor F ($I > 1$), we call this positive clipping. The remaining artifacts caused by the background which are still visible in the result image (see Figure 6) are examples for pixels that can't be compensated as the dynamic range of the display is exceeded. In the following we present two alternative approaches called adaptive compensation 1 and adaptive compensation 2 that adapts the contrast and brightness of the overlaid image to reduce the occurrences of clipping errors while using the dynamic range of the integrated display.

5.2 Adaptive radiometric compensation

Both adaptive algorithms still apply the final correction step as in equation 8, but their goal is to modify the desired result R by replacing it with R' which has the characteristic that it reduces the total number of pixels I outside the range $[0,1]$ whilst keeping R' as perceptually similar to R as possible. The inputs to both algorithms are the original desired image R and the background light that would be perceived with the display off, denoted E_D, computed in the same way as equation 7

$$E_D = t_{B2}(t_{B1}\frac{C1}{r_{B1}}) \quad (9)$$

Our first algorithm (labeled adaptive compensation 1 in results) is based on the work in radiometrically correcting projector-camera systems as proposed by Grundhöfer and Bimber [9]. Besides using our additive light model for error computation steps, the primary differences are that we use E_D in place of the projection background material E, and assume that maximum displayable intensity F_M is white for all pixels. In brief (for a more details we refer to [9]), the algorithm works by creating a luminance map L from the Y channel of R in the CIE XYZ color space, using the minimum and maximum values of E_D and the average value of L to perform a global scaling step, then computing the remaining clipping error as Err. The Err map is smoothed with a Gaussian filter, then masked by the threshold map [16] of R, the result of which is used to lighten or darken the globally scaled image locally, giving R'.

The algorithm for adaptive compensation 1 increases the brightness even if clipping does not occur and leads to general brighter results. We therefore developed another algorithm for locally adapting display output (labeled adaptive compensation 2) which does not perform any global compensation, in an attempt to preserve the full dynamic range of the display device in areas that do not have any clipping artifacts. The projected intensity of adaptive compensation 2 is calculated in the same manner as the other compensation methods, once again calculating a new value for R', as in

$$I = \frac{R' - t_{B2}(t_{B1}\frac{C1}{r_{B1}})}{Fr_{B2}} \quad (10)$$

where R' is targeted at maintaining the contrast and chromaticity characteristics of the original desired image D after the blend is applied, sometimes at the expense of direct colour accuracy. First we calculate the uncorrected result R using equation 7, then calculate a threshold map from the blended image. This threshold value T describes the maximum amount of correction that may be applied to any given pixel - as we want some correction possible across the entire image we set a small, empirically determined floor value

$$T = max(0.05, TM(R)) \quad (11)$$

We next find the projection error in RGB color space, which is the amount by which R is out of the $[0,1]$ boundary. For each RGB channel we can determine Err_{RGB} with

$$Err_{RGB} = \begin{cases} R - 1 & \text{if } R > 1 \\ R & \text{if } R < 0 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

The luminance error map Err_Y, representing the Y component in CIE XYZ color space, is calculated by direct multiplication from Err_{RGB}, including negative components. The error is then masked by the threshold map to create Err_M by clamping values on the range of $[-T, T]$. As Err_M is quite sharp at this point (as it is partially derived from a Laplacian pyramid), we smooth it using a boosted Gaussian blurring function with a sigma of $\frac{1}{20}$th of the display width. This blur does not make any pixel darker; it boosts pixels' brightness based on their proximity to any other brighter pixel. The positive and negative values of Err_M are treated in separate passes, then summed to create Err_G.

R' is calculated by modifying the Y channel of D in the CIE xyY color space by the corresponding position in Err_G. An overview of the results achieved with both approaches can be seen in Figure 6.

Fig. 7. Detail of the displayed compensation image computed for each approach. Adaptive compensation 1 tends to compute brighter compensation images to avoid clipping while adaptive compensation 2 uses the dynamic range of the display by saturating the colours compared to the global compensation.
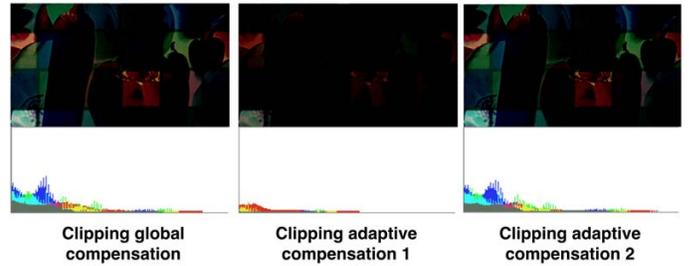

Fig. 8. Images and histogram showing the areas and amount of clipping for each approach and for each colour channel. Adaptive compensation 1 produces the least amount of clipping while adaptive compensation produces less than global compensation but to an lesser extent than adaptive compensation 1.

5.3 Implementation

All three algorithms are implemented using Qt 5.4 and OpenGL 3.3. Our developed application uses QML to display the intermediate and final results, however the radiometric compensation is running in GLSL on a series of framebuffers. We use PointGrey's FlyCapture version 2.8 to access the camera and OpenCV 3.0 for the ahead of time calibration steps. The whole application runs on a Windows 10 x64 desktop computer (Intel Core i7-2600 CPU with 3.40GHz, 8GB RAM, nVidia GeForce GTX750) which was also driving our prototype described earlier which is based around an Epson Moverio BT100 which is driven via a custom HDMI connector instead of the default analog connector to improve the video quality. This custom HDMI connector was taken from a Meta Glass prototype which internally used also a version of the Epson Moverio BT100. All three algorithms were running in real-time, constrained only by the camera frame-rate which was at 25fps.

6 RESULTS

We tested our approach with a combination of several foreground and background combinations. Figure 6 shows some of the results with a test image from a standard Computer Vision image database[5]. Here we can see the reference image which is the desired image used as input for our algorithms. Based on the previously described algorithms we compute the compensations image ("displayed compensation") which is displayed in the display and blends with the background, in this case a color chart, to mitigate the effect of color blending. The final results can be seen in the top row of Figure 6.

Overall, one can see that our three compensation algorithms make visual improvements that better resemble the reference image when compared to the naive overlay. What is also visible is that the adaptive compensation 1 results in a brighter image. When compared to adaptive compensation 1, algorithm 2 performs less overall correction but is more highly targeted to areas with more color blending interference. In areas where both algorithms perform a similar amount of deviation between R and R', adaptive compensation 1 will tend towards white, whereas adaptive compensation 2 will tend towards less bright but more saturated colors. Figure 7 shows some details of the computed calculation image shown in Figure 6 and one can see for example in the lower left corner that the adaptive compensation 1 is brighter while adaptive compensation 2 saturates the colours. This behaviour can also be seen on the details for other results.

Figure 8 shows the regions where clipping occurs together with a histogram. What can be seen is that the adaptive compensation 1 produces less clipping compared to the other approaches. Adaptive compensation 2 produces generally less clipping than the global compensation but the difference is small and concentrates on smaller areas.

6.1 User Study

In addition to the presentation of the visual results of our approaches for radiometric compensation, we conducted a user study to evaluate

[5] http://sipi.usc.edu/database/

the subjectively perceived compensation quality. We decided for a subjective evaluation as in particular the adaptive algorithms change the original image content taking into account the human perception. Consequently, the displayed image can not match the original image as we designed both adaptive algorithms to not reassemble the original image. We consequently decided for a well-established subjective evaluation method similar to the one used in [9]. Instead of letting users wear the head-mounted display, which would require a complex eye-display calibration and would introduce additional, confounding parameters to the study, we decided to show the visual results of different combinations as captured by the eye-camera $C2$ and let the participants rate the visual quality on a standard computer screen.

6.1.1 Study design

We are following a within-subject, repeated measures design showing each participant three different foreground images taken from a standard Computer Vision image database[6] each overlaid onto four different backgrounds: wood texture, outdoor scene, urban scene, and a color chart (Figure 10).

We also showed for each foreground image a reference image which was the foreground image overlaid onto a uniform black background which served as a reference representing the best possible output for the OHMD. We had five conditions which were randomized: Reference image, no compensation (naive overlay), global compensation, adaptive compensation 1, and adaptive compensation 2. Overall, each participant saw 51 images (in three groups of foreground images). The order of the images was randomized so that the three foreground
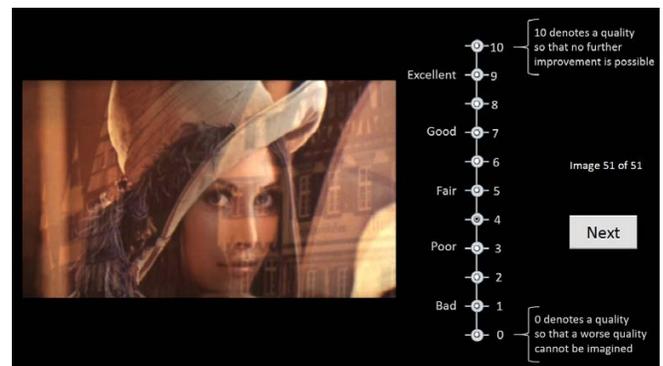
[6] http://sipi.usc.edu/database/

Fig. 9. Rating interface as used in the user study. The image was shown on a non-distracting black background. The rating scale was taken from the original ITU-T recommendation, including the (slightly modified) wording of the anchors (0 and 10). The Next button was activated after a rating had been placed.

Fig. 10. The four background images used in the study representing both, a range of practical environments possibly experienced in actual application scenarios as well as extremes and subtleties of color, contrast, and other image characteristics.
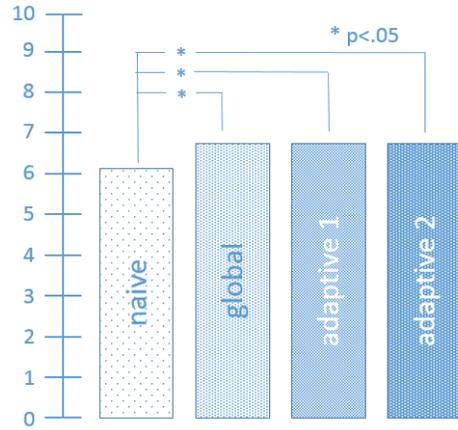

Fig. 11. Means for differential viewer scores for the four user study conditions. Asterisks (*) indicate a statistically significant difference ($p < .05$). There was a significant difference in the perception of the image quality between a naive blending and all of our implementations but not in-between them.

groups and the order of the background images and conditions were differently mixed per participant. We use random.org for the a priori randomization. All (combined) images used for the study were captured through the eye camera $C2$. The camera's brightness was adjusted for each foreground image to match the desired color when shown on a black background (reference image). However, the camera settings were kept the same for each algorithm. The final images were cropped to only show the area covered by the OHMD and had a resolution (after cropping) of 1103x644 pixels. Some examples of images used for the user study can be seen in Figure 12.

We are using a modified ACR-11-HR quality expression method (ITU-T Rec. P.910) as described in [21]. The ACR method is widely used in judging visual picture quality in multimedia applications, including mobile video. Five or eleven point scales can be used, the latter requires a slightly higher assessment time but is having a higher distinguishing power. In addition, ACR does allow to work with a hidden reference image instead of being overt about this. We opted for the 11-point scale together with a hidden reference image (hence ACR-11-HR) as the conservative options. Each image is watched for max. 10 secs in a randomized order (in groups) including the reference images (HREF). The overall (differential viewer) score per image DV(PV S) is then calculated by

$$DV(PVS) = V(PVS) - V(HREF) + 10 \qquad (13)$$

where V(PV S) is the original raw rating on a 11 point (0..10) Likert-like scale (prescribed format by ITU-T) and V(HREF) is the subjective rating of the ("ideal") reference image. Because all images have been randomized in their order the participants did not know which image was HREF and also would not know which image result was produced by which algorithm.

6.1.2 Study Results

Eighteen participants (five female, thirteen male) have been recruited for the study with a mean age of 39 years (range 21-58). All participants had normal or corrected to normal vision with one participant indicating a certain, unspecified form of color blindness. After reading an information sheet and signing an ethics consent form participants followed on screen instructions on a 13" monitor and rated the image quality of each individual picture by clicking a radio button (0..10) as depicted in Figure 9. Each image had to be rated before participants could proceed to the next image. All data has been automatically captured in a comma separated text file. The 51 image ratings per participant have been rescored to their respective differential viewer scores to mitigate effects of individual over- or underestimation of the image quality. The resulting scores have been analysed for normality, skew,

and kurtosis and t-tests with an alpha level of 0.05 have been applied. No outliers or anomalies have been detected.

There was a statistically significant difference between the naive condition (M=6.82, SD=2.26) and each of the other three conditions (global, adaptive 1, adaptive 2) (M=7.48, SD=2.09) but not between the non-naive conditions, See Figure 11.

Those results indicate that our approach for radiometric compensation works well, but that we cannot conclude yet that the rather subtle differences between the algorithms (global and two adaptive versions) will make a perceivable difference for the viewer. Future studies have to show potential differences in users' perceptions.

7 CONCLUSION

We presented a first approach that enables radiometric compensation of OHMDs in real-time and with pixel-accuracy. Existing approaches only focused on creating display profiles improving the color accuracy but ignored the effect of color blending or used an external camera to capture the environment making it not applicable for practical applications as it needs to be geometrically re-calibrated for every new position. In this work we present different design iterations for a novel OHMD prototype virtually placing the camera into the user's eye, which allows us to capture the environment information as seen by the user. This captured environment information is used by our algorithms to compute a compensation image minimizing the effect of color blending between the displayed image and the environment. We presented the results of three different versions of the basic algorithm which all have individual strengths and weaknesses. Visual results as well as results from a user study confirm our hypothesis that we can improve the visual quality of images displayed in OHMDs. While all algorithms produce results which are significantly better than the uncorrected image, the differences between each of them are only marginal when evaluated by the users, warranting further investigation.

We believe this work to be of high relevance to the general field of Augmented Reality but in particular for the development of future OHMDs as well as Heads-up displays suffering from the effect of color blending reducing the visual quality of digital overlays.

One of the biggest challenges of out approach is the limited dynamic range of the camera as well as of the display. The effect can be reduced by using the dark shades that can usually added to the OHMDs (Epson Moverio) or are integral part of the OHMD (Microsoft Hololens). However, cameras with higher dynamic ranges and displays with a higher brightness would allow to increase the range of colors that can be compensated while still being sensitive to small deviations.
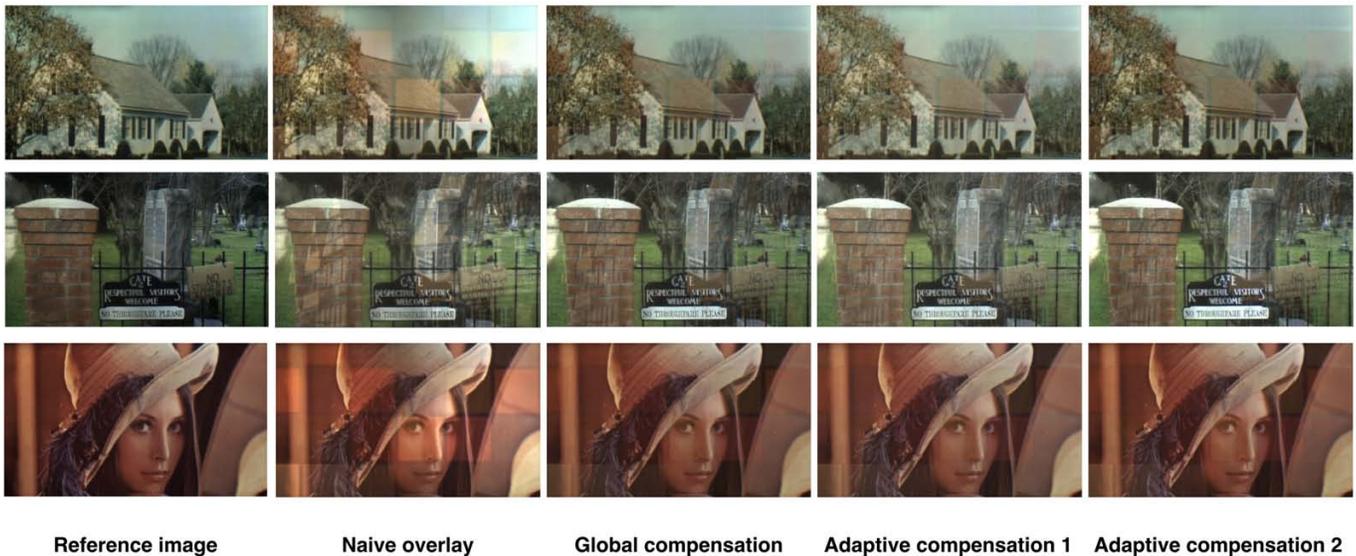
Fig. 12. Some of the result images we used in our user study. All images are captured through the OHMD by camera $C2$ representing the human eye. The reference image shows the image displayed in the OHMD in front of a black background and serves as an ideal result. The naive overlay is with a colored background but without radiometric compensation showing color-blending artifacts. The remaining images show the result of our radiometric compensation using the different presented approaches reducing the effect caused by color-blending.

Future research directions include combining our approach with eye-display calibration to allow the full integration into OHMD and consequently would allow for evaluating our approach while wearing the actual OHMD. Future generations of OHMDs could also replace the beam-splitter and the camera with a transparent image sensor which is directly integrated into the glass of the OHMDs. Early prototypes of these kind of sensors exist [14] but they are still a topic for fundamental research questions and are far from being ready for commercial use. Finally, we see many research directions exploring applications for this kind of OHMDs. In particular smart glasses that augmented the capabilities of the user would benefit from our approach. Instead of trying to achieve a desired color they could change it in order to provide a better perceptual experience for example to raise contrast against a known background environment. This would also open up opportunities for novel, unique AR applications. Examples include color shifting/enhancement for color-blind or hard of sight users, diminished reality applications, and 'user's eye view' recording.

REFERENCES

[1] O. Bimber, F. Coriand, A. Kleppe, E. Bruns, S. Zollmann, and T. Langlotz. Superimposing pictorial artwork with projected imagery. *IEEE Multimedia*, 12(1):16–26, 2005.

[2] O. Bimber, A. Emmerling, and T. Klemmer. Embedded entertainment with smart projectors. *IEEE Computer*, 38(1):48–55, 2005.

[3] O. Bimber, D. Iwai, and G. Wetzstein. The visual computing of projector-camera systems. *Computer Graphics Forum*, pages 2219–2245, 2007.

[4] O. Cakmakci, Y. Ha, and J. P. Rolland. A compact optical see-through head-worn display with occlusion support. In *Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on*, pages 16–25, Nov 2004.

[5] T. Fukiage, T. Oishi, and K. Ikeuchi. Visibility-based blending for real-time applications. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 63–72, Sept 2014.

[6] J. L. Gabbard, J. E. Swan, and A. Zarger. Color blending in outdoor optical see-through ar: The effect of real-world backgrounds on user interface color. *Proceedings - IEEE Virtual Reality 2013 Short Papers and Posters*, pages 157–158, 2013.

[7] J. L. Gabbard, J. Zedlitz, J. E. Swan, and W. W. Winchester. More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces. *Proceedings - IEEE Virtual Reality*, pages 79–86, 2000.

[8] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2016.

[9] A. Grundhofer and O. Bimber. Real-time adaptive radiometric compensation. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):97–108, Jan 2008.

[10] J. D. Hincapié-Ramos, L. Ivanchuk, S. K. Sridharan, and P. Irani. Smartcolor: Real-time color correction and contrast for optical see-through head-mounted displays. *Science and Technology Proceeding - IEEE International Symposium on Mixed and Augmented Reality*, pages 187–194, 2014.

[11] Y. Itoh, T. Amano, and G. Klinker. Semi-parametric color reproduction method for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, pages 1269–1278, 2015.

[12] Y. Itoh and G. Klinker. Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 75–82, March 2014.

[13] K. Kiyokawa, Y. Kurata, and H. Ohno. An optical see-through display for mutual occlusion of real and virtual environments. In *Augmented Reality, 2000. (ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 60–67, 2000.

[14] A. Koppelhuber and O. Bimber. Lumiconsense: A transparent, flexible, and scalable thin-film sensor. *IEEE Computer Graphics and Applications*, 34(5):98–102, Sept 2014.

[15] A. Plopski, Y. Itoh, C. Nitschke, K. Kiyokawa, G. Klinker, and H. Takemura. Corneal-imaging calibration for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)*, 21(4):481–490, April 2015.

[16] M. Ramasubramanian, S. N. Pattanaik, and D. P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 73–82. ACM Press/Addison-Wesley Publishing Co., 1999.

[17] J. Rolland and H. Hua. Head-mounted display systems. In *Encyclopedia of Optical Engineering*, page 2005. Marcel Dekker.

[18] S. K. Sridharan, J. D. Hincapié-Ramos, D. R. Flatla, and P. Irani. Color

correction for optical see-through displays using display color profiles. *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, pages x–x, 2013.

[19] I. E. Sutherland. A head-mounted three dimensional display. *Fall Joint Computer Conference*, pages 757–764, 1968.

[20] B. Thomas, B. Close, J. Donoghue, J. Squires, P. D. Bondi, and W. Piekarski. First person indoor / outdoor augmented reality application: Arquake. *Personal and Ubiquitous Computing*, pages 75–86, 2002.

[21] T. Tominaga, T. Hayashi, J. Okamoto, and A. Takahashi. Performance comparisons of subjective quality assessment methods for mobile video. In *Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on*, pages 82–87, June 2010.

[22] C. Weiland, A. Braun, and W. Heiden. Colorimetric and photometric compensation for optical see-through displays. pages 603–612, 2009.