Color Contrast Enhanced Rendering for Optical See-through Head-mounted Displays

Yunjin Zhang, Rui Wang, Yifan (Evan) Peng, Wei Hua, and Hujun Bao

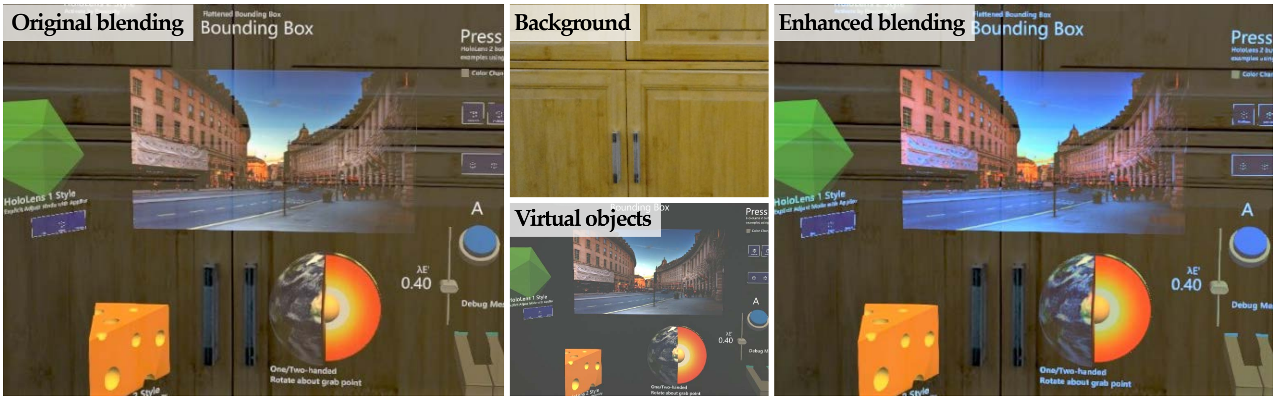


Fig. 1: Representative results of color contrast enhanced rendering for optical see-through head-mounted displays (OST-HMDs). **Left:** Original blending scene perceived from a commercially available OST-HMD in front of a typical background. **Middle Top:** The original background. **Middle Bottom:** The virtual objects in front of a white background. **Right:** Our method improves the visual distinctions between rendered images and the physical environment by performing a constrained color optimization regarding the perception in chromaticity and luminance of the displayed color.

Abstract—Most commercially available optical see-through head-mounted displays (OST-HMDs) utilize optical combiners to simultaneously visualize the physical background and virtual objects. The displayed images perceived by users are a blend of rendered pixels and background colors. Enabling high fidelity color perception in mixed reality (MR) scenarios using OST-HMDs is an important but challenging task. We propose a real-time rendering scheme to enhance the color contrast between virtual objects and the surrounding background for OST-HMDs. Inspired by the discovery of color perception in psychophysics, we first formulate the color contrast enhancement as a constrained optimization problem. We then design an end-to-end algorithm to search the optimal complementary shift in both chromaticity and luminance of the displayed color. This aims at enhancing the contrast between virtual objects and the real background as well as keeping the consistency with the original displayed color. We assess the performance of our approach using a simulated OST-HMD environment and an off-the-shelf OST-HMD. Experimental results from objective evaluations and subjective user studies demonstrate that the proposed approach makes rendered virtual objects more distinguishable from the surrounding background, thereby bringing a better visual experience.

Index Terms—Color blending, color perception, human visual system, mixed reality, real-time rendering, post-processing effect

1 Introduction

In recent years, innovations in optical see-through head-mounted displays (OST-HMDs) have led to the rapid development of mixed reality (MR) technologies. In contrast to virtual reality (VR) HMDs or video see-through HMDs, OST-HMDs mainly allow users to perceive the real, non-rendered environment through the optics, as well as the virtual content through displays. This design scheme drastically relieves the discomfort when wearing common video-based HMDs. However, the optical combiner blends the rendered pixels with the physical background. As such, the virtual objects are unable to be presented independently. This essential property of combiners often causes a color-blending problem, which hinders the ability of clearly observing the virtual content in MR scenarios, especially when the virtual and real scenes show low color contrast.

Existing solutions to the color-blending problem can be divided into two categories, namely, hardware- and software-based solutions. Hardware-based solutions [1], [2] physically adjust each pixel's transparency on displays to avoid the color blending between rendered images and the background. Although certain solutions [3], [4] exploit every effort for miniaturization, their scalability and flexibility are very limited on account of additional hardware. Software-based methods [5] focus on color correction, seeking to

- *Yunjin Zhang, Rui Wang, Wei Hua, and Hujun Bao were in the State Key Laboratory of CAD&CG, Zhejiang University, China.*
Yifan (Evan) Peng was in Electrical Engineering, Stanford University, USA.
Corresponding author: Rui Wang, rwang@cad.zju.edu.cn

mitigate the color blending via subtraction compensation. Researchers have developed high-precision color correction algorithms [6], [7] and accurate colorimetric background estimation [8]. However, for commercial OST-HMDs, such as Google Glass, Microsoft HoloLens, and Magic Leap, using subtraction compensation may decrease the brightness of the virtual content, resulting in low visual distinctions between rendered images and the physical environment.

In certain MR scenarios, we note that color correctness may not be the only goal of displaying virtual objects. Instead, the color contrast of virtual objects to the real background should be sometimes enhanced to allow users to better distinguish the perceived virtual objects from the background. To this point, intuitively increasing the brightness of the virtual objects may lead to a decreased contrast within their surfaces [9]. Therefore, this work proposes a novel real-time color contrast enhancement for OST-HMDs, aiming to improve the distinction between the rendered image and the real background, as well as to consider the consistency between the enhanced display colors to the original displayed colors. In particular, we perform a constrained color optimization in the CIELAB color space to search the optimal displayed color with the maximized color contrast compared to its surrounding background, under constraints regarding the perception in chromaticity and luminance. To further increase the perceptual color contrast, our work utilize one characteristic of the human visual system (HVS) that the color of one region induces the complementary color in the neighboring region [10]. Results show that the proposed method can enhance the perceived distinction between the virtual content and the surrounding background in typical environments.

Particularly, our contributions are as follows:

- We exploit another perspective for color-blending problem, namely enhancing color contrast to improve users' visual experience;
- We develop an end-to-end, real-time rendering algorithm to find optimal colors, improving the distinction between displayed images and physical environments on OST-HMDs;
- We demonstrate the effectiveness of our approach using a simulated environment and a commercial OST-HMD.

2 Related Work

MR devices and corresponding algorithms are becoming prolific research areas. Numerous studies explored how to better present virtual scenes, such as harmonization [11], focal distance adjustment [12], [13], color reproduction [14], [15], color balancing [16], light filtering [17], [18], [19], color correction [6], [7], [20], visibility improvement [9], [21], etc. Moreover, psychophysical experiments are also conducted on color perception problems in MR scenarios, such as brightness matching [22], color matching [23], [24], and white adjusting [25]. In this section, we introduce the most relevant research topics, color correction and visibility improvement, in more detail, and then describe a highly relevant work that the perceived color contrast in the HVS.

Color Correction. Solutions to color blending are commonly categorized into hardware- or software-based approaches. Hardware-based approaches often refer to occlusion support, which avoids color blending physically by

Fig. 2: Examples of the simultaneous color induction. **Left:** Two red patches in the center are colored in the same RGB value $(255, 96, 96)$, yet they appear different colors in perception. **Right:** Two grey patches are colored in the same value $(64, 64, 64)$ but have different perceived colors. More patterns can be found in related studies [37], [38], [39], [40].

using additional relay optics such as transmissive [1] and reflective [26] spatial-light modulators (SLM). By creating occlusion masks, OST-HMDs control environment light into the user's eyes to provide the occlusion effect. Transmissive methods [27], [28] often use a liquid crystal display (LCD) to cut rays that pass through the virtual content at the pixel level. Among reflective approaches, some optical systems [29], [30] utilize a digital micro-mirror device (DMD) for occlusion support, while others [2], [3], [31] use liquid crystal on silicon (LCoS). Certain methods [32], [33] use a high-speed switcher to frequently alternate between the virtual content and the real scene, achieving full or partial occlusion. Some efforts [4], [34] were also made at compactness. However, the flexibility of all the aforementioned solutions is often restricted owing to additional optical components.

By contrast, utilizing existing resources can also handle perception problems on limited hardware [35]. Software-based methods attempt to mitigate color blending by changing the color of display pixels. Typical methods are known as color correction or compensation [5], which start from capturing the background and then accurately map the background to the corresponding virtual content. Thereafter, the background color is subtracted from the rendered color of the virtual scene. Certain works pay more attention to colorimetric estimation [8], [36] or radiometric measurement [7] of the background to obtain background information with higher accuracy, while the subjective distortion of displays [23] is ignored.

However, the aforementioned methods of subtraction compensation result in a decrease in the brightness of displayed images, thereby reducing the visibility of virtual contents. One work [20] manages to increase visibility with high contrast but is demonstrated mainly for the textual content. In general, our goal in the current work is to improve the distinctions between virtual objects and physical backgrounds rather than just keep the perceived image consistent with the input, which is the primary difference between our method and color correction.

Visibility Improvement. Complementary to color correction, few recent works focus on improving the visibility of the virtual content. Fukiage et al. [9] proposed an adaptive blending method on the basis of a subjective metric of visibility. However, on common OST-HMDs, this method increases the brightness but reduces the contrast of virtual objects, leading to a washed-out effect of the texture details. Lee and Kim [21] used tone mapping to enhance the visibility of low gray-level regions under ambient light, but they only demonstrated on grayscale images. The main problem

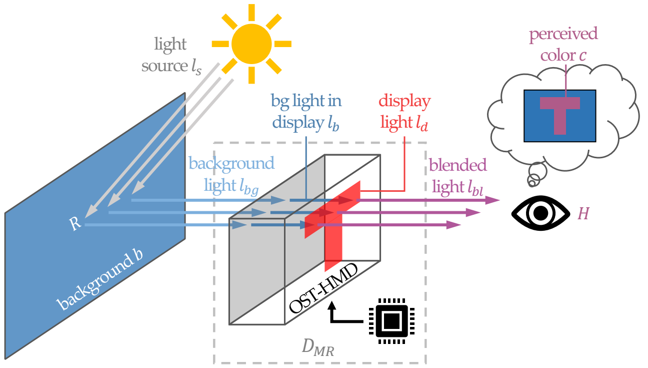


Fig. 3: Illustration of the color-blending problem on OST-HMDs. The light source l_s interacts with the object surface in the background b by the reflectance function R, and then the reflected light (i.e., the background light l_{bg}) enters the OST-HMDs. Subsequently, the entire display system D_{MR} blends the background light in displays l_b with the display light l_d to generate the blended light l_{bl}. Finally, the user's eyes receive l_{bl}, which forms the perceived color c through the operation H of HVS.

of these two approaches is that they only consider the luminance of rendered images, limiting their effectiveness for virtual objects with colorful, intricate textures.

Color Contrast. Color contrast is the perceptual difference between one region and its adjacent region in color. In the HVS, this difference is influenced by the chromaticity and luminance of a test stimulus and its surrounding area when presented simultaneously [10], [41], [42], [43]. This effect is called the simultaneous color induction. For example, a red patch looks redder on a green background than on a red background (see Figure 2). Some studies [37], [42] validate the color induction and related phenomena through various experiments. Researchers also systematically quantify the effects of simultaneous color induction along the dimension of hue [38], [40].

It is generally acknowledged that two contents with different colors displayed simultaneously induce a complementary color shift to each other (the complementarity law [44]). In other words, changes in the color appearance of the induced stimulus are directed away from the appearance of inducing surrounds. However, some recent hypotheses, such as the direction law [39], [44], [45], have challenged the traditional complementarity law, thus indicating that the mechanism of simultaneous color induction is still not understood completely.

Although the elaboration of these hypotheses is not in the scope of this work, they do provide valuable insights for the current study and prompt us to approach color blending with another perspective. In this work, we utilize the complementarity law, and results show that it works well in the experiments.

3 Problem Statement

In this section, we describe the color-blending problem illustrated in Figure 3. We formulate our method, color contrast enhancement, as a constrained optimization of color blending. According to the definition of color blending on OST-HMDs by Gabbard et al. [46], the blending procedure can be formulated as follows:

$$c = H\left(l_{bl}\right) \tag{1a}$$
$$l_{bl} = D_{MR}\left(l_d, l_{bg}\right) \tag{1b}$$
$$l_{bg} = R\left(l_s, b\right), \tag{1c}$$

where c represents the perceived color, and H is the operation of the HVS for the blended light l_{bl} that reaches the user's eyes. D_{MR} is an abstract of the entire OST-HMDs system, which contains multiple parameters, such as lens opacity and display brightness. l_d denotes the display light and l_{bg} refers to the background light that enters the front of the OST-HMDs. Reflectance function R depicts how the light source l_s in the background b interacts with the object surface and finally enters the OST-HMDs. Previous software-based color correction methods devote to accurately measuring and estimating l_{bg} and parameters of D_{MR} (e.g., lens opacity), allowing them to subtract the background light in displays (l_b) from the display light (l_d). As a result, these methods remove the background color from the display pixels to mitigate the color blending effect. Unlike color correction approaches, we are not seeking to handle the distortions [20] or measure hardware parameters of OST-HMDs [7]. Instead, we focus on the HVS's operation H and the perceived color c. Therefore, function $D_{MR}\left(l_d, l_{bg}\right)$ can be approximated as $l_d + l_b$:

$$l_{bl} \approx l_d + l_b. \tag{2}$$

On the basis of the opponent-colors theory proposed by Jameson and Hurvich [41], the perceived color c of a test stimulus can be expressed as follows:

$$\begin{aligned} c = f[&(r-g)_t + (r-g)_i, (y-b)_t + (y-b)_i, \\ &(w-bk)_t + (w-bk)_i], \end{aligned} \tag{3}$$

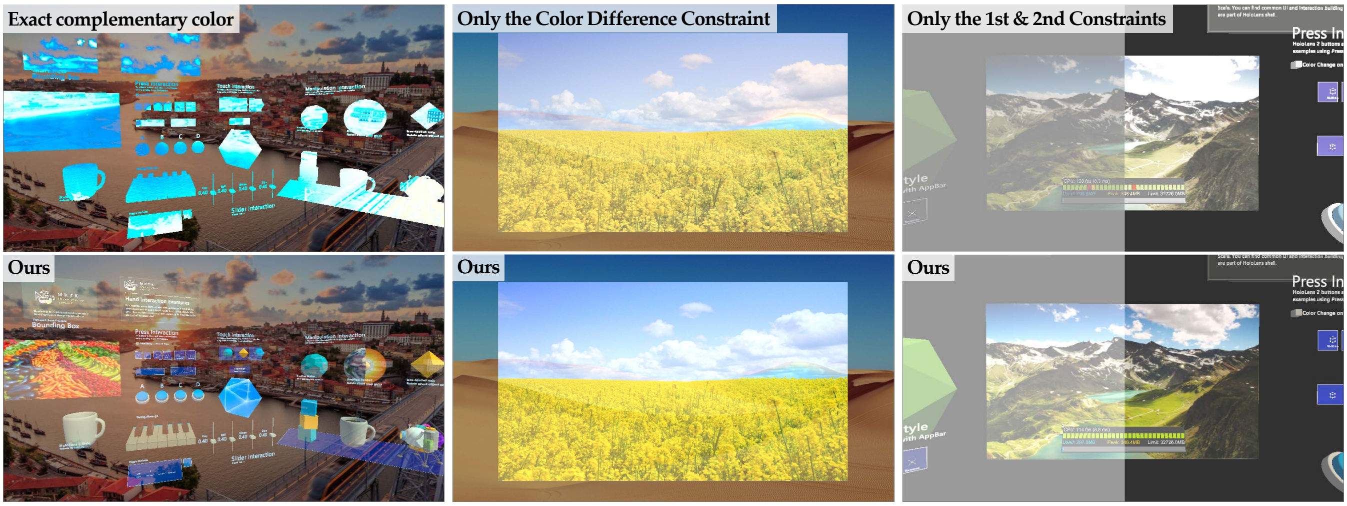
where f processes the input neural responses and outputs the perceived color. $(r-g)_t$, $(y-b)_t$, and $(w-bk)_t$ are responses of three paired and opponent neural systems of the test stimulus. $(r-g)_i$, $(y-b)_i$, and $(w-bk)_i$ denote the responses of corresponding systems from the surrounding area induced by the simultaneous color induction. In converse, this induction also affects the perceived color of the surrounding area.

The color difference ΔE^*_{ab} defined in the CIELAB color space (also referred to as L*a*b* color space) between color x and color y can be calculated as follows:

$$\begin{aligned} \Delta E^*_{ab}(x,y) &= \|x-y\| \\ &= \sqrt{\left(L^*_x - L^*_y\right)^2 + \left(a^*_x - a^*_y\right)^2 + \left(b^*_x - b^*_y\right)^2}, \end{aligned} \tag{4}$$

where L^*, a^*, and b^* are three orthonormal bases of the L*a*b* color space used to describe the luminance and chromaticity of colors.

The key of our approach is to shift l_d for increasing the color difference $\Delta E^*_{ab}(l_d, l_b)$ in Equation (4). Shifting l_d also leads to an increase in the corresponding responses $(r-g)_i$, $(y-b)_i$, and $(w-bk)_i$ in Equation (3), which further increases the perceived color difference between l_d and l_b. In this manner, we improve the distinction between the virtual content and the surrounding background. We've also tested enhancing $\Delta E^*_{ab}(l_d + l_b, l_b)$ and found that $l_d + l_b$ usually has a large luminance component but a relatively



(a) (b) (c)

Fig. 4: Comparisons between our optimization results and rendering results without all/some of the constraints. **(a)** We use a landscape photo as the test background and some objects from the *Hand Interaction Examples*[1]scene as the virtual content. The upper image uses the per-pixel complementary color of the background color, meeting no constraints. The lower image is rendered with our method. This comparison demonstrates the Color Difference Constraint. **(b)** Two landscape photos serve as the test background and virtual content. The upper image is rendered with the optimal colors that meet only the Color Difference Constraint, and the lower image is the optimized result of our method. This comparison demonstrates the Chroma Constraint. **(c)** We use a bright gray and a dark gray as the test background and a landscape photo and other objects as the virtual contents. The upper image is rendered with the optimal colors that follow the first and second constraints but not the third, and the lower image is the optimized result of our algorithm. This comparison demonstrates the Luminance Constraint. Please refer to the supplementary video for real examples captured by the HoloLens.

small chromaticity component. Applying our color contrast enhancement to such a color usually produces brighter colors than that produced from l_d, but decreases the contrast within surfaces of virtual objects. Therefore, we choose to enhance $\Delta E^*_{ab}(l_d, l_b)$ instead of $\Delta E^*_{ab}(l_d + l_b, l_b)$.

When the background light in displays l_b and the display light l_d exhibit the most significant color difference, that is, l_b and l_d are complementary colors, the HVS perceives a maximum color contrast. However, considering only the color contrast may lead to an unintended color altering of displayed objects. The upper image of Figure 4a shows an example of such a case. It is clear that the virtual objects at the bottom left lack texture details and can hardly be recognized. This most straightforward optimization rule can almost only be used for textual content. Therefore, we introduce several constraints when optimizing the optimal displayed color l_{opt}, to maintain its color consistency with the original displayed color as:

$$l_{opt} = \arg\max_{l_{opt}} \Delta E^*_{ab}(l_{opt}, l_b) \quad \text{subject to constraints.} \tag{5}$$

3.1 Constraints of the Optimization

Given the aforementioned issue, the color difference ΔE^*_{ab} between l_b and l_d cannot be unlimited. One should keep the color consistency of the displayed color of virtual objects. On this basis, we introduce the first constraint, named the **Color Difference Constraint**, to restrict the range of l_{opt}:

$$\Delta E^*_{ab}(l_{opt}, l_d) \leq \lambda_E, \tag{6}$$

1. https://github.com/microsoft/MixedRealityToolkit-Unity

where λ_E represents a non-negative color difference threshold. As such, the optimal displayed color l_{opt} and the original displayed color l_d should be kept within a certain range in color.

Further, if the hues of the background color l_b and the display color l_d are similar, l_{opt} would shift along the complementary direction of l_b, resulting in a decrease in chroma of l_d (see Figure 4b for a specific example). We introduce the second constraint to tackle the chroma reduction, named the **Chroma Constraint**:

$$ch_{opt} - ch_d \geq 0, \tag{7}$$

where ch_{opt} and ch_d represent the chroma of l_{opt} and l_d, respectively. Note that although we have restricted the reduction in chroma, there is no boundary to the increment existing with this constraint.

However, in addition to the constraint added on chroma, the luminance can benefit from adding a corresponding constraint. For example, if l_b is a bright, whitish color, the complementary color of l_b is close to a dark grey. In this case, l_{opt} moves toward the direction of l_b's complementary color, resulting in a decrease in luminance (see Figure 4c for an example). Reducing the luminance of displayed color on common OST-HMDs often leads to increased transparency. Therefore, displaying l_{opt} with no constraint on luminance reduces the visibility of virtual objects. Moreover, the experimental results of Fukiage et al. [9] showed that if the visibility of virtual objects in OST-HMDs is enhanced by unconstrainedly increasing the luminance of l_d, then the perceptual contrast within surfaces of these objects decreases. These situations lead to the third constraint, named

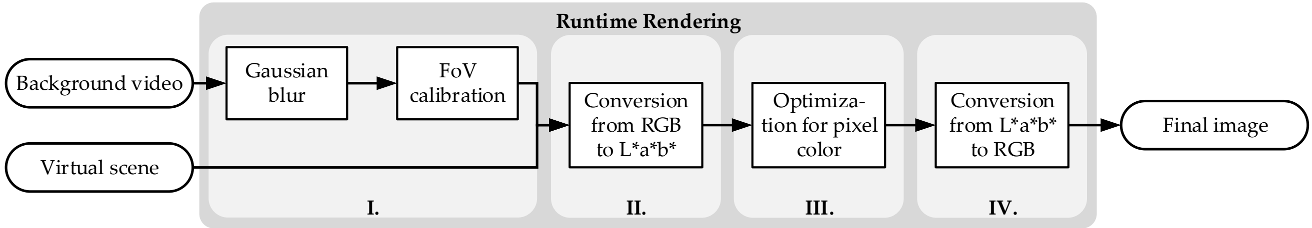


Fig. 5: Algorithm Overview. Our algorithm takes the rendered virtual scene and the streaming background video as input. First, a Gaussian blur and the FoV calibration are applied to the background video. Second, we transform the blurred video and the virtual scene from RGB to the L*a*b* color space and then optimize the displayed color for all pixels. Finally, the pixel color of the virtual scene is converted back to RGB color space and output to displays.

the **Luminance Constraint**:

$$\Delta L^*(l_{opt}, l_d) \leq \lambda_L, \tag{8}$$

where ΔL^* is the luminance difference, and λ_L denotes a non-negative luminance threshold. This constraint alleviates reduced visibility of the displayed content caused by the bright background, as well as the contrast reduction of virtual objects in a dim environment.

Finally, we introduce an evident constraint between l_{opt} and l_b, named the **Just Noticeable Difference Constraint**:

$$\Delta E^*_{ab}(l_{opt}, l_b) \geq \lambda_{JND}, \tag{9}$$

where λ_{JND} represents the just noticeable difference (JND). This constraint indicates that the optimal color l_{opt} and the background color l_b require a minimum color difference that the HVS can distinguish. Note that this constraint is generally satisfied automatically on account of the objective of our optimization. Considering that JND is a statistical rather than an exact quantity [47], [48], if users exhibit an above-average JND, applying this constraint is mandatory.

4 Algorithm

We design a real-time algorithm for color contrast enhancement in various environments. Figure 5 shows an overview of this algorithm, which mainly includes four procedures:

- **I. Preprocessing:** We perform a Gaussian blur and the field of view (FoV, see Section 4.1) calibration for the streaming video of the background.
- **II. Conversion:** We convert the display and background colors from RGB to the L*a*b* space.
- **III. Optimization:** Utilizing the aforementioned four constraints, we optimize the displayed colors on the basis of the background colors.
- **IV. Displaying:** We convert the optimized display colors back to RGB space for displaying them.

4.1 Preprocessing

It is generally believed that multiple individual pattern analyzers contribute to the contrast sensitivity of humans [49]. These pattern analyzers are often called spatial-frequency channels, which filter the perceived image into spatially localized receptive fields with a limited range of spatial frequencies. For example, the low-spatial-frequency channels receive the color and outlines of objects, whereas the high-spatial-frequency channels perceive details. Considering the low-pass nature of color vision [50] and the blurring characteristic of the non-focal field in the HVS, our method does not need a pixel-precise camera-to-display calibration.

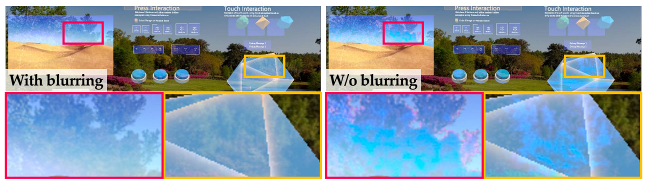


Fig. 6: Demonstration of the necessity of image blur. We use a landscape photo as the test background. The left side is the optimized result of our algorithm with background blurring enabled. Optimal displayed colors without background blurring are represented on the right side. The zoomed regions emphasize details.

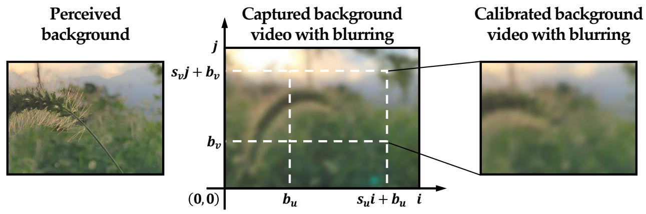
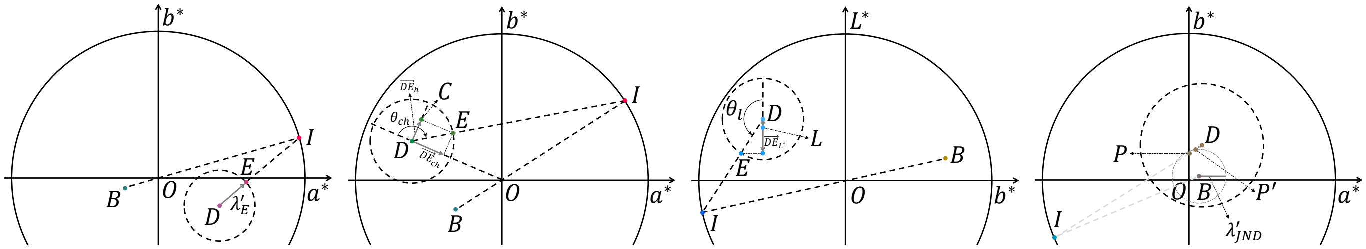


Fig. 7: Illustration of FoV calibration. This process is performed when the FoV between the captured background video and the frame buffer are different. We crop and scale the captured video to match the position and size of the background seen through OST-HMDs.

Instead, we use a Gaussian blur to extract the spatial color information and filter out details of the background video as $G(x, y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$, where x and y are the horizontal and vertical distances from a pixel to its center pixel, respectively, and σ is the standard deviation of the normal distribution. Such blurring simulates the non-focal effect of the HVS. In this manner, the displayed color obtains the weighted average of multiple background colors in the corresponding region. This filter also reduces the flicker in optimized display colors caused by the high-spatial-frequency details in the background (see Figure 6 for examples).

Given the low-frequency characteristic of the blurred background video and the difference of FoV between captured videos and OST-HMDs, we subsequently use a screen-space coordinate mapping called FoV calibration between the background video and the frame buffer of the rendering system, in order to approximate pixel-precise calibration. On this basis, we apply the following coordinate mapping to map the background video to the frame buffer:

$$\begin{cases} u = s_u i + b_u \\ v = s_v j + b_v, \end{cases} \tag{10}$$



(a) (b) (c) (d)

Fig. 8: We illustrate our algorithm on the 2D plane formed by two of the three axes of the L*a*b* color space for simplicity. When coloring the points in (a), (b), and (d), their L^* value is set to 50. Similarly, a^* value of all colored points in (c) is set to 0. **(a)** Illustration of calculating coordinates I and E. All points in this subfigure are plotted on the plane a^*Ob^*. In the unit circle, I is the point farthest from B. All points with the same distance λ'_E from D form a circle, which intersects the line DI at point E. **(b)** Illustration of determining the coordinates C. All points in this subfigure are plotted on the plane a^*Ob^*. If θ_{ch} is an obtuse angle, the change in chroma $\overrightarrow{DE}_{ch}$ of vector $\overrightarrow{DE}$ will be discarded. Thus, the chroma-constrained vector $\overrightarrow{DC}$ only contains the component $\overrightarrow{DE}_h$ of $\overrightarrow{DE}$. **(c)** Illustration of determining the coordinates L. All points in this subfigure are plotted on the plane b^*OL^*. The change in luminance $\overrightarrow{DE}_{L^*}$ of vector $\overrightarrow{DE}$ will be attenuated with $\cos\theta_l$. Thus, $\overrightarrow{DE}_{L^*}$ is shortened to $\overrightarrow{DL}$. **(d)** Illustration of finding the new optimal displayed color P' under the Just Noticeable Difference Constraint. All points in this subfigure are plotted on the plane a^*Ob^*. Sometimes the coordinates P of optimal displayed color l_{opt} falls into the circle B on account of a large radius λ'_{JND}, meaning that the HVS cannot distinguish between l_{opt} and l_b. Therefore, we find the intersection of line DP and circle B as P'.

where (u, v) and (i, j) represent the 2D texture coordinates of the frame buffer and the background video, respectively. (s_u, s_v) denotes the scale factor in the horizontal and vertical directions of the background video coordinates, and (b_u, b_v) is corresponding offsets. Figure 7 shows the illustration. We assume that the FoV of captured videos is greater than that of OST-HMDs. Using this calibration, the low-frequency information of background videos and that of real scenes seen through displays of OST-HMDs is as similar as possible.

4.2 Conversion and Displaying

After preprocessing, all colors are stored and represented in RGB space. However, the widely used RGB space is not perceptually uniform, in that the same amount of numerical change does not correspond to the same amount of color difference in perception. We convert the background color and the displayed color from RGB to the L*a*b* color space, to utilize its (near) uniformity in perception and independence of luminance and chromaticity. After performing the color contrast enhancement, we transform L*a*b* colors back to RGB to display them appropriately on OST-HMDs.

Additionally, the gamut of the L*a*b* color space is more extensive than displays and even the HVS, indicating that many of the coordinates in the L*a*b* color space, especially those located at the edge, cannot be reproduced on typical displays. Therefore, we scale the range of the original L*a*b* color space to $[-1, 1]$ and take its inscribed sphere as the solution space of our method.

4.3 Optimization

Given a blurred background color l_b and an original displayed color l_d in the scaled L*a*b* space, the objective of our optimization is to find the optimal displayed color l_{opt}. On the basis of the complementarity law [44], we shift l_d in the complementary direction of l_b.

In this work, we denote points by capital italic letters. First, we calculate the coordinates I of the ideally optimal displayed color corresponding to the blurred background color without considering any constraint:

$$I = -\frac{B}{\text{dist}(B, O)} = -\text{norm}(\overrightarrow{OB}), \tag{11}$$

where B represents the coordinates of the blurred background color l_b. $\text{dist}(B, O)$ denotes the distance between B and the center O of the unit sphere. This formula can also be rewritten to the latter form, where $\text{norm}(\overrightarrow{OB})$ means to normalize the vector $\overrightarrow{OB}$. That is, the farthest point from B in the unit sphere is the intersection of the extension line of BO and the sphere (see Figure 8a).

Given the ideally optimal displayed color I, we then consider the aforementioned four constraints: Color Difference Constraint, Chroma Constraint, Luminance Constraint, and Just Noticeable Difference Constraint, in color optimization in Equation (5). First, we calculate the coordinates E of the initial optimal display color by applying the Color Difference Constraint to I:

$$\overrightarrow{DE} = \min\left(\text{dist}(D, I), \lambda'_E\right) \cdot \text{norm}(\overrightarrow{DI}), \tag{12}$$

where $\overrightarrow{DI}$ is the ideal change vector starting from the coordinates D of original displayed color l_d, and λ'_E is the scaled color difference threshold specified by users. Now, we have the change vector $\overrightarrow{DE}$ of the displayed color constrained by the color difference. Figure 8a presents a two-dimensional illustration of this step.

Let $\mathbf{v}'$ be the projection of vector $\mathbf{v}$ onto the plane a^*Ob^* of the L*a*b* color space. Thus, the change in chroma $\overrightarrow{DE}'_{ch}$ of the vector $\overrightarrow{DE}'$ can be obtained by calculating the projection of $\overrightarrow{DE}'$ onto the vector $\overrightarrow{OD}'$ (see Figure 8b for a two-dimensional example). For the Chroma Constraint, we discard the chroma reduction of $\overrightarrow{DE}'$ as:

$$\overrightarrow{DC} = t_{ch} \cdot \overrightarrow{DE}'_{ch} + \overrightarrow{DE}'_h, \quad t_{ch} = \begin{cases} 1, & \text{if } \theta_{ch} \le 90^\circ \\ 0, & \text{if } \theta_{ch} > 90^\circ. \end{cases} \tag{13}$$

Here $\overrightarrow{DE}'_h = \overrightarrow{DE}' - \overrightarrow{DE}'_{ch}$, which refers to the component of $\overrightarrow{DE}'$ perpendicular to $\overrightarrow{DE}'_{ch}$. θ_{ch} represents the angle (0°–180°) from $\overrightarrow{OD}'$ to vector $\overrightarrow{DI}'$. This angle describes the deviation in hue and chroma between the optimal displayed color and the original displayed color. In this manner, the adaptive parameter t_{ch} provides a smooth visual effect when hue and chroma change.

As for the Luminance Constraint, we also use an adaptive parameter to scale the alterations in luminance:

$$\overrightarrow{DL} = (1 - |\cos\theta_l|) \cdot \overrightarrow{DE}_{L^*}, \tag{14}$$

where $\overrightarrow{DE}_{L^*}$ refers to $\overrightarrow{DE}$'s component on the L^* axis of the L*a*b* color space. θ_l represents the angle (0°–180°) from the positive L^* axis to $\overrightarrow{DE}$ (see Figure 8c for a two-dimensional example). Correspondingly, this angle indicates the difference in luminance between the optimal displayed color and the original displayed color. The adaptive coefficient $(1 - |\cos\theta_l|)$ smooths the results and makes it redundant to specify the luminance threshold λ_L by users.

According to the above steps, the coordinates P of optimal displayed color l_{opt} can be obtained by:

$$P = D + \overrightarrow{DC} + \overrightarrow{DL}. \tag{15}$$

In contrast to the other three constraints, the Just Noticeable Difference Constraint is only applied to some extreme cases (Figure 8d), wherein we reduce the length of $\overrightarrow{DP}$ so that the distance from the coordinates P' of the new optimal displayed color to the coordinates B of the blurred background color is equal to a scaled JND λ'_{JND} in the unit sphere. Such cases are found in users who have an above-average JND, leading to a larger radius of circle B in Figure 8d. Algorithm 1 gives the pseudocode of the entire optimization process above, where Vector3(0, $E.yz$) generates a three-dimensional vector of which the first component is 0, and the last two components are $E.y$ and $E.z$.

Algorithm 1 Finding the coordinates P of the optimal displayed color l_{opt}

Input: $D, B, \lambda'_E, \lambda'_{JND}$
Output: P
Components x, y, and z denote L^*, a^*, and b^*, respectively.
dir = Vector3(0, 0, 0);
Color Difference Constraint :
$I = -$ norm(B);
$E = \min(\text{dist}(D, I), \lambda'_E) \cdot \text{norm}(I - D)$;
Chroma Constraint :
E' = Vector3(0, $E.yz$);
E'_{ch} = change in chroma of E';
$E'_h = E' - E'_{ch}$;
θ_{ch} = angle from Vector3(0, $D.yz$) to E';
$dir.yz = ((\cos\theta_{ch} \geq 0\ ?\ 1 : 0) \cdot E'_{ch} + E'_h).yz$;
Luminance Constraint :
θ_l = angle from Vector3(1, 0, 0) to E;
$dir.x = (\cos\theta_l \geq 0\ ?\ (1 - \cos\theta_l) : (1 + \cos\theta_l)) \cdot E.x$;
Optimal display color :
$P = D + dir$;
Just Noticeable Difference Constraint :
if (dist(P, B) $< \lambda'_{JND}$) **then**
  P = intersection of line DP and circle B of radius λ'_{JND};
end if
return P

5 Implementation

We implemented our algorithm as a full-screen post-processing effect using the *Unity 2019* rendering engine and validated it using a simulated OST-HMD environment and a commercially available OST-HMD. The simulated environment is based on the *Unity Editor* with a resolution of 960×540, using a series of images as the backgrounds. For the real environment, we used the Microsoft HoloLens MR headset [51] to evaluate our algorithm. Both environments take the modified *Hand Interaction Examples* scene as the virtual scene, which includes text, photographs, user interfaces, and 3D models with plain or intricate materials. We attenuated the luminance of background images in the simulated environment and streaming videos in the HoloLens by 60% to simulate the real background perceived through the translucent lens in the HoloLens. It is a tweakable parameter for other types of OST-HMDs with different levels of transparency. In both environments, the kernel size and σ of our Gaussian filter are 3 and 1.5, respectively. These values apply to all participants in our user studies (Section 6.2, 6.3, and 6.4). We captured streaming videos at 24 frames per second (FPS) by a built-in RGB camera located in the center of the HoloLens in real-time. The screen brightness of the HoloLens was fixed at 70%. The data of our FoV calibration was determined by careful manual calibration. Specifically, $(s_u, s_v) = (0.65, 0.65)$, and $(b_u, b_v) = (0.13, 0.17)$.

6 Results

In our context, we take the scaled color difference threshold (λ'_E) to regulate our color contrast enhancement. All experiments on the HoloLens were conducted indoors, with illuminants of daylight or LED lamps. Besides, we found that features of the physical environment (such as scene illuminance) have no significant impact on our experimental results, on account of the auto exposure and the auto white balance of the built-in camera of the HoloLens.

6.1 Evaluation

We used different images as the background of the simulated environment to evaluate our method in various practical scenarios. Figure 9 shows some of the results. Virtual objects and dimmed background are blended directly as Equation (2) instead of mixed by alpha blending, to simulate the optical properties of OST-HMDs. Pixels that meet the following condition are marked as cyan in Figure 9, 11a, and 11b, meaning that these pixels have an increased ΔE^*_{ab} from the corresponding background in perception:

$$(\Delta E^*_{ab}(l_b, l_{opt}) > \Delta E^*_{ab}(l_b, l_d)) \cap (\Delta E^*_{ab}(l_{opt}, l_d) \geq \lambda_{JND}). \tag{16}$$

Here, λ_{JND} is about 2.3 in the L*a*b* color space [47]. Note that not all virtual content can gain perceptual color contrast enhancement. Given the constraints mentioned above, the color difference between certain optimal colors and their original displayed colors is less than a JND. Moreover, some original displayed colors are almost the complementary colors of the background colors and do not require further

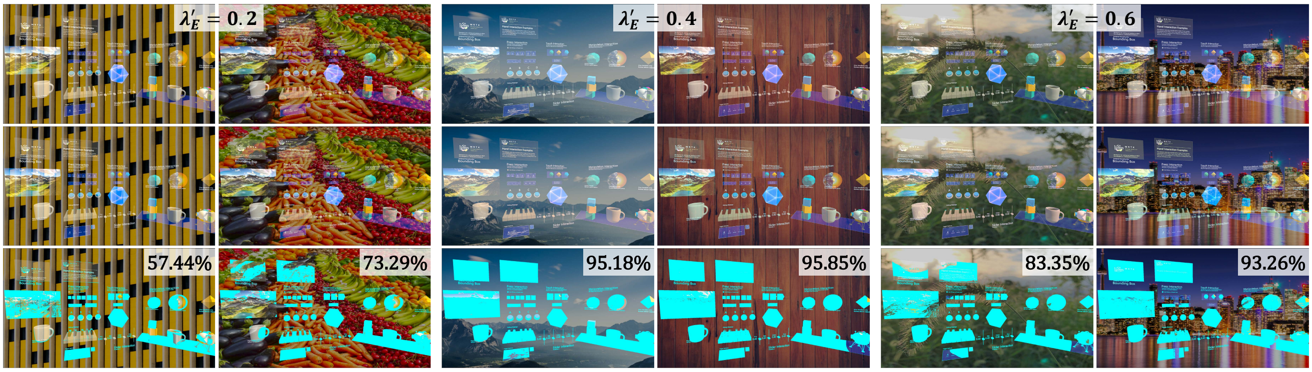


Fig. 9: Selected result images in the simulated OST-HMD environment. Each group with a different λ'_E shows two background images. The first row shows the original blending images, and the second row shows the results of our method. In the last row, pixels colored in cyan indicate a perceptually increased ΔE^*_{ab} between the background color and the displayed color, wherein the numbers in the figures represent the corresponding percentage of these pixels in all foreground pixels. Please refer to the supplementary video for real examples captured by the HoloLens.

Fig. 10: Results of different λ'_E for the same background in the simulated environment. Numbers indicate the percentage of enhanced pixels in all foreground pixels. One can see that the white text in front of yellow backgrounds looks bluish, whereas that in front of blue backgrounds looks yellowish. Please refer to the supplementary video for real examples of different λ'_E captured by the HoloLens.

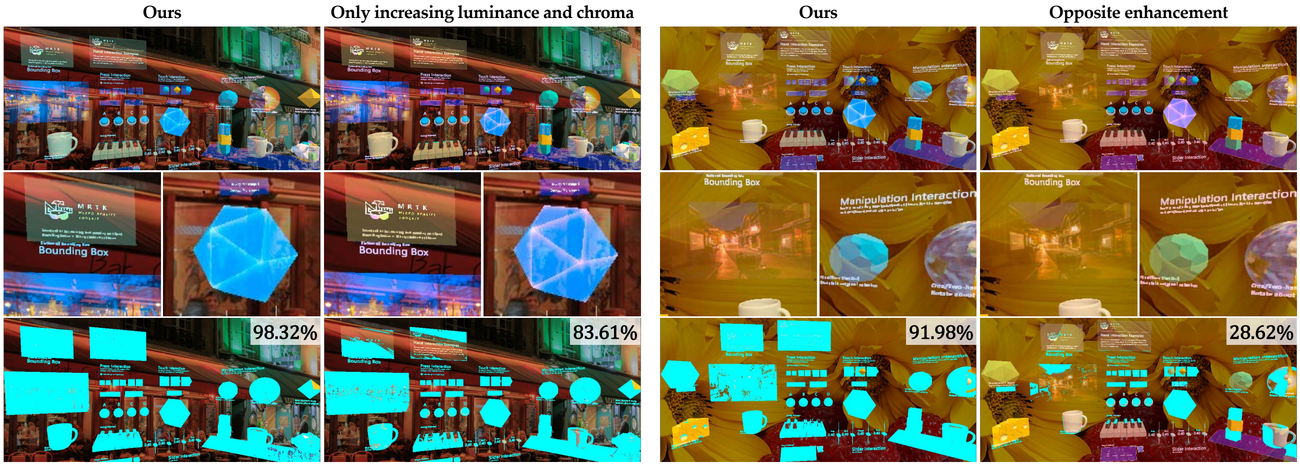
enhancement. We also validated our method with varying λ'_E with the same background, as shown in Figure 10. Generally, a larger color difference leads to an increase of the enhanced pixels in quantities and makes the displayed color more complementary to the background color.

To evaluate the enhancement contributed by hues, we compared results produced by our method and by only increasing luminance and chroma. Figure 12a shows a 2D illustration of this enhancement in chromaticity as $\overrightarrow{DC^{\diamond}} = ||\overrightarrow{DC}|| \cdot \text{norm}(\overrightarrow{OD})$. To ensure increased luminance, when $\theta_l > 90^\circ$, we set $\overrightarrow{DL}$ to $-\overrightarrow{DL}$. In this manner, the changes in ΔE^*_{ab} in the two enhancements are identical, while the shift directions are different. Comparison results are shown in Figure 11a with the same λ'_E of 0.6. The zoomed insets highlight several virtual objects in front of reddish-yellow backgrounds. Our method produces complementary greenish-blue shifts, which makes objects more distinguishable from the surrounding background than that produced by only increasing the color difference regardless of hues.

The difference in the direction of hue shifts between the two color contrast enhancements is shown in Figure 12b, where one direction is $\overrightarrow{DC}$, and the other direction is $\overrightarrow{DC^{\star}} = t_{ch} \cdot \overrightarrow{DE'_{ch}} - \overrightarrow{DE'_h}$. $\overrightarrow{DL}$ is set to a zero vector. Thus, the optimized display colors of the two enhancements are the same as the original display colors in luminance. The changes in ΔE^*_{ab} and chroma in the two enhancements are identical, but the shift directions are opposite. In fact, there is only one direction ($\overrightarrow{DC^{\star}}$) that is iso-luminance, iso-color-difference, and iso-chroma compared to $\overrightarrow{DC}$. Figure 11b shows comparison results, where λ'_E is both set to 0.4. Along our direction, the photograph, polyhedron, and the earth gain blue hue shifts, thereby appear more clear and visible in front of yellow backgrounds. By contrast, the opposite direction produces yellow hue shifts similar to the background, making objects more transparent.

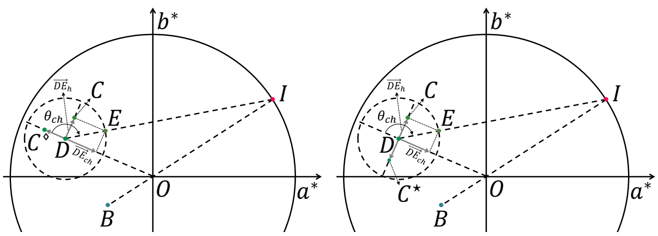
We also evaluated our algorithm on the HoloLens. Figure 1 shows one of the experimental results, where the scaled color difference threshold λ'_E we used was 0.4. In front of a yellow background, our method shifts the displayed color to the complementary direction (i.e., blue) of the background color, to enhance the color difference for better visual distinctions. For example, the sky and the ground of the landscape photograph look bluish, as well as the text in the scene. However, subject to the aforementioned constraints (Section 3.1), the chroma of yellow cheese and mantle does not decrease. Overall, our method is able to keep the consistency from the original displayed color.

Although our goal is to improve the distinctions between virtual contents and surrounding backgrounds rather than correcting colors, we compared our color contrast enhancement with a typical compensation method [5]. Figure 13a presents the experimental results. Note that this result of our enhancement shows the maximum color contrast under a given λ'_E. If one needs less contrast but more consistency with the original displayed color, the threshold λ'_E is tweakable. As for the subtraction compensation, it reduces the luminance of rendered pixels, leading to low visual distinctions between virtual objects and the physical environment. We note that more sophisticated methods based on subtrac-



(a) (b)

Fig. 11: Comparison results in the simulated environment. Numbers in figures denote the corresponding percentage of the enhanced pixels in all foreground pixels. Pixels that have an increased perceptual color difference are marked as cyan. **(a)** Different enhancements with the same change in color difference. Left: Our method. Right: Enhancing color contrast by increasing luminance and chroma only. **(b)** Different hue shifts with the same control conditions (i.e., iso-color-difference, iso-luminance, and iso-chroma). Left: Color contrast enhancement produced by hue shifts of our method. Right: Opposite enhancement produced by hue shifts but in the opposite direction.



(a) (b)

Fig. 12: 2D illustrations of two enhancements. **(a)** Only increasing luminance and chroma. $C^\diamond$ is the coordinates of the optimized display color following this enhancement. $\overrightarrow{DC^\diamond}$ is equal in length to $\overrightarrow{DC}$, but its direction is along $\overrightarrow{OD}$. **(b)** Enhancing color contrast in the other direction of the hue shift. $C^\star$ is the coordinates of the optimized display color following this enhancement. $\overrightarrow{DC^\star}$ is equal in length to $\overrightarrow{DC}$, while it contains a component opposite to $\overrightarrow{DE}_h$.

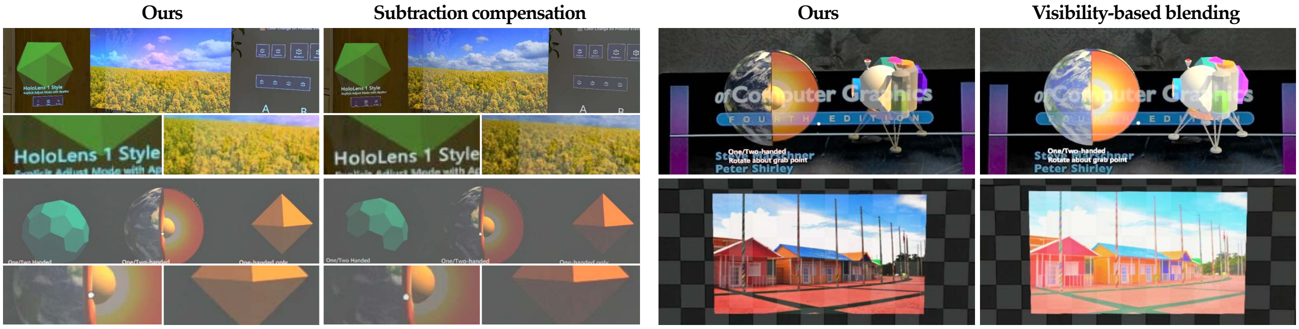
tion compensation (e.g., [7], [36]) also fail to maintain the original visibility of virtual content on common OST-HMDs. Additionally, we compared our method with the visibility-based blending approach [9], as shown in Figure 13b. This blending method increases the brightness at the expense of contrast within surfaces, resulting in a washed-out visual effect of the virtual content. We manually conducted the camera-to-display calibration for these comparison methods so that the background captured by the built-in camera is as similar as possible to that perceived by the human eye.

6.2 User Study I: Threshold Preference

We conducted three user studies to evaluate our method subjectively. Before the start of these studies, there was a training stage to help subjects familiarize themselves with the interface and usage of the HoloLens. During these experiments, participants wore the HoloLens and explored the aforementioned virtual scene freely in various environments (see Figure 14 for an example). In this scene, dozens of virtual objects were placed surrounding the user. Owing to the small FoV of the HoloLens, participants needed to rotate their heads (i.e., change the camera viewport) to see different virtual objects with different backgrounds. To ensure that their view was correct, participants were asked to look at the virtual object and center it by rotating their heads, then describe this object.

User Study I is a two-alternative forced choice (2AFC) subjective experiment, wherein the participants were asked to compare the results of our approach and original rendering. To calibrate the color perception, participants first perceived the reference display color of the virtual object, i.e., the original displayed color in front of a white background (see the center bottom image of Figure 1 for an example). We put a whiteboard in front of the HoloLens to occlude the real background. Then, participants performed full comparisons by freely toggling between the two rendering results of this object, and were asked two questions. First, which of the results was more distinguishable from surrounding backgrounds. Second, which result looked more consistent. This consistency is compared between the perceived display color and the reference display color of a virtual object. For example, white keys that appear to be red have a relatively low consistency. We asked each participant to look at three randomly picked objects, where each object corresponds to one camera viewport. We provided five levels of λ'_E, ranging from 0.2–1.0, with a step of 0.2. Thus, each participant compared results in three random viewports under five λ'_E and gave a total of 15 choices for each question.

A total of 15 participants, 12 males and 3 females, with a mean age of 23.6 (range 21–28), took part in the experiments. All participants had normal or corrected-to-normal vision without any form of color blindness. Six of them had experience with mixed reality prior to this study. Participants gave informed consent to take part in this study. Right before the study, we had every participant perform an eye-to-display calibration through a built-in calibration application of the HoloLens. Finally, we received 45 choices



(a) (b)

Fig. 13: Rendering results of different methods on the HoloLens, in front of various backgrounds. **(a)** Left: Our method ($\lambda'_E = 0.4$). Right: Subtraction compensation [5] ($k_v = 1.0$ and $k_b = 0.4$, where $k_v = 1.0$ means the same display color in both methods, and $k_b = 0.4$ denotes 40% lens transparency). When the background is light gray (such as a white wall), our method has limited optimization for rendered pixels. By contrast, the subtraction compensation causes degradation, making the virtual contents more transparent. **(b)** Left: Our method ($\lambda'_E = 0.4$). Right: Visibility-based blending [9] ($V_t = 1.5$ as specified by the authors in their paper). When the background is an achromatic pattern, our method has little optimization for rendered pixels. By contrast, visibility-based blending improves the distinctions between virtual contents and physical background but decreases contrast within surfaces of objects and leads to a possibly unintended change in color. Please refer to the supplementary video for comparison results between these methods.

of each question under each λ'_E. We used a binomial test to evaluate the choice of users, wherein the hypothesized probability of success is 0.5 and the confidence level is 95%. These values are applied to all tests in this work.

Figure 15a shows the complete results of the preferences of the participants. When $\lambda'_E \geq 0.4$, our method is preferred more often than the original rendering in distinction. On the other hand, the color consistency of our results significantly decreases as λ'_E increases when $\lambda'_E \geq 0.8$. These results show that the value of λ'_E seems to be positively correlated with the distinction but negatively correlated with the consistency. However, when $\lambda'_E = 0.4$, the optimized virtual content is more distinguishable from the background (31 of 45, $p = 0.016$, Binom. test), whereas the difference of color consistency between the optimized and the original virtual content is not statistically significant (21 of 45, $p = 0.766$, Binom. test). We tend to believe that our approach has found a trade-off between contrast and consistency, but further studies are needed for validation.

6.3 User Study II: Hue Evaluation

To subjectively evaluate the effect of hues, we performed another 2AFC experiment. Participants were asked to perform two comparison tasks between results of 1) our enhancement and enhancing color contrast by increasing luminance and chroma only; 2) two different directions of hue shift with the same control conditions (see Section 6.1). For each comparison, participants freely toggled between the two results and then were asked which of the results was more distinguishable from the surrounding background. Each participant compared the results of three random viewports and gave a choice in each setting. We fixed λ'_E at 0.4.

Sixteen participants, including 4 females and 12 males with an average age of 24.6 (range 23–28), volunteered. All participants had normal or corrected-to-normal vision without color blindness. Nine participants had prior experience with mixed reality, and seven of them took part in the previous study 9 months ago. Participants gave informed consent to take part in this study. We required each participant to

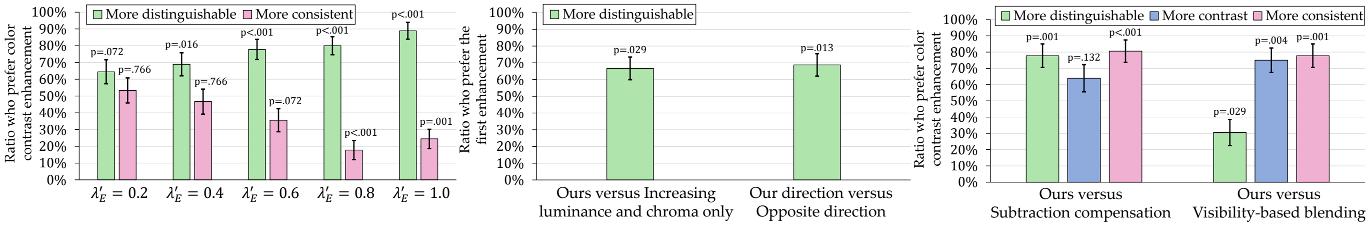
Fig. 14: Scene perceived from the HoloLens by participants in a daily environment, with our color contrast enhancement enabled ($\lambda'_E = 0.4$). One can see that the photograph in front of the green background looks reddish.

perform the eye-to-display calibration before the study.

For each comparison, a total of 48 choices were collected. Figure 15b shows the complete results of the preferences. For both comparisons, the first enhancement was often preferred to the second one. Specifically, our enhancement is more distinguishable than increasing luminance and chroma only (32 of 48, $p = 0.029$, Binom. test). Also, the hue shift direction we proposed is more distinguishable than that along the opposite direction under the same conditions (33 of 48, $p = 0.013$, Binom. test). These results indicate that in most cases, hue shifts in an appropriate direction can further improve the visual distinctions between rendered images and the surrounding environment.

6.4 User Study III: Method Comparison

We conducted another 2AFC experiment in which the participants were asked to compare the results of ours with those of 1) subtraction compensation [5]; 2) visibility-based blending [9]. Participants first perceived the reference display color of an object to calibrate the color perception, and then fully compared the rendered images of this object between these methods and ours by freely toggling. For each method, participants were asked three questions. First, which result was more distinguishable from surrounding backgrounds. Second, which result had higher contrast. Third, which result looked more consistent. Each participant



(a) (b) (c)

Fig. 15: Results of our first (a), second (b), and third (c) subjective experiments. **(a)** The participants compared our color contrast enhanced rendering with the original rendering. λ'_E represents the scaled color difference threshold. **(b)** The participants compared the results of different enhancements. **(c)** The participants compared our color contrast enhanced rendering with subtraction compensation [5] ($k_v = 1$ and $k_b = 0.4$) and visibility-based blending [9] ($V_t = 1.5$). In all figures, participants' preferences are shown as percentages. p-value (Binom. test) is shown above the column, respectively. The error bars represent the standard error.

compared rendered images of three random viewports and gave six choices for each question. We fixed λ'_E at 0.4.

Twelve participants consisting of 5 females and 7 males volunteered, ages 18–33 (mean 22.6). No one exhibited any signs of color blindness. All participants had normal or corrected-to-normal vision. No participant took part in the previous studies, and only one participant had prior experience with mixed reality. Participants gave informed consent to participate in this study. We had every participant perform the eye-to-display calibration before the study.

Finally, we collected 36 choices for every question of each compared method. Figure 15c shows the complete results of the preferences. For the first question, participants preferred our method more often than subtraction compensation (28 of 36, $p = 0.001$, Binom. test). For the second question, our results are considered to have higher contrast compared with visibility-based blending (27 of 36, $p = 0.004$, Binom. test). For the third question, our method is statistically more consistent than the other two methods (29 of 36, $p < 0.001$, and 28 of 36, $p = 0.001$, Binom. test). These results mean that, in most cases, our approach is regarded as more distinguishable and consistent than subtraction compensation, whereas the contrast and consistency of rendered images are significantly higher than visibility-based blending.

Results show that participants regarded our approach as less distinguishable than the visibility-based blending [9] (11 of 36, $p = 0.029$, Binom. test), as the latter significantly increases the luminance of virtual contents. Nevertheless, for applications that pay more attention to texture details, it may not be an optimal solution.

6.5 Performance

Our method does not rely on any precomputation. Therefore, it supports real-time rendering under various scenarios with dynamic camera viewports. Parameters such as color difference threshold can be tuned in real-time. We measured the runtime performance on the HoloLens, which has a display area of 1268×720 pixels for each eye. We rendered displayed contents at different viewports in the *Hand Interaction Examples* scene, covering percentages of the display area that ranges from 0%–100%, with a step of 10%. Different surrounding backgrounds are also used in our measurement. For each level of percentages, we sampled ten times and calculated the average FPS. Generally, reported FPS from the HoloLens varies from 31–55, depending on the number of rendered pixels, as shown in Figure 16.

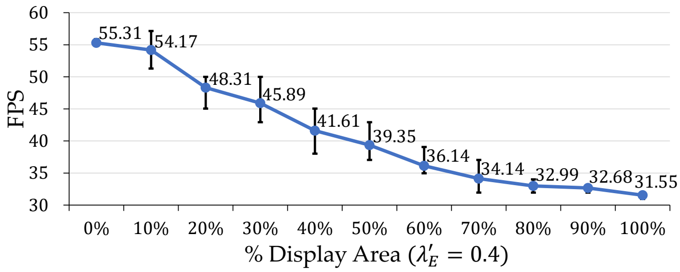


Fig. 16: Performance of our color contrast enhancement as a full-screen post-processing effect on the HoloLens, measured in FPS. The error bars represent the maximum and minimum FPS.

7 Limitation and Discussion

Our approach, though not free from limitations, constitutes a promising avenue to future work. Although the experimental results indicate that our color contrast enhanced rendering for OST-HMDs works well in a proper threshold of color difference, finding an adaptive threshold corresponding to the current physical environment remains an unexplored valuable topic.

There are several aspects in our implementation requiring elaboration. First, our method currently does not consider achromatic backgrounds. Although uncommon, it is reasonable to assume that the physical background is completely achromatic. One possibility to incorporate this achromaticity situation into our optimization would be to increase the chroma of displayed colors. Second, colors in some applications (e.g., for military and industry) have typical meanings, so hue shifts may lead to unexpected results. Reducing the color difference threshold or increasing luminance and chroma are helpful for these applications. Third, later standards have supplanted the L*a*b* color space we used. For example, CAM16 has different adapting surrounds, while L*a*b* do not consider these. However, CAM16 ignores chromatic simultaneous color induction [24], which is common in MR scenarios. We regard more appropriate color models as future work. Finally, the mechanism that lies behind simultaneous color induction has not been well-understood. Diving into how MR could leverage the phenomenon to further improve color perception is also worth exploring.

Note that for those applications giving more weight to geometry than to color and texture, the scope of current work does not fit well. We consider this as an orthogonal problem and save it for future work. In addition, it is still challenging to display black or darker colors due to the hardware limitations of OST-HMDs. Solving such problems is beyond the scope of software-based solutions.

8 Conclusion

On the basis of another perspective for color blending, we present an end-to-end, real-time color contrast enhancement algorithm for OST-HMDs that takes both chromaticity and luminance of displayed colors into account. Existing methods focus solely on changing the luminance of displays or environments to improve the distinction between virtual objects and real background, or compensating displayed colors to approximate the colors originally intended. In this work, we further consider the impact of chromaticity to improve the color perception of rendered images from the perspective of color contrast. Specifically, we use the complementary color of the background as the search direction in the $L^*a^*b^*$ color space to determine the optimal color of the display pixel under various constraints.

To summarize, we strive to inspire future studies on color contrast enhanced rendering. Our current implementation achieves several key technical characteristics. First, it adaptively enhances the visibility of virtual objects. Second, it is real-time. Third, there is no need for any hardware change. Our algorithm is implemented on the GPU using pixel shaders, allowing users to tweak the parameters, such as the threshold of color difference, in real-time. We present results in simulated and real OST-HMDs. In addition, both the objective evaluations and user studies confirm that in most cases, our approach can improve the perceptual contrast between displayed content and physical background in OST-HMDs, making virtual objects more distinguishable from surrounding backgrounds while achieving the maximum level of consistency with the original displayed color.

Acknowledgment

We would like to thank all reviewers and editors for their insightful comments. We also want to thank Hongyu Lu and other volunteers for taking part in our studies under the spread of COVID-19. This research was partially funded by NSFC (No. 61872319), Zhejiang Provincial NSFC (No. LR18F020002), National Key R&D Program of China (No. 2017YFB1002605), and Zhejiang University Education Foundation Global Partnership Fund.

References

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

[2] O. Cakmakci, Y. Ha, and J. P. Rolland, "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*, ser. ISMAR '04. Washington, DC, USA: IEEE Computer Society, 2004, pp. 16–25.

[3] C. Gao, Y. Lin, and H. Hua, "Occlusion capable optical see-through head-mounted display using freeform optics," in *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2012, pp. 281–282.

[4] A. Wilson and H. Hua, "Design and prototype of an augmented reality display with per-pixel mutual occlusion capability," *Opt. Express*, vol. 25, no. 24, pp. 30 539–30 549, 2017.

[5] C. Weiland, A.-K. Braun, and W. Heiden, "Colorimetric and photometric compensation for optical see-through displays," in *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, C. Stephanidis, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 603–612.

[6] S. K. Sridharan, J. D. Hincapié-Ramos, D. R. Flatla, and P. Irani, "Color correction for optical see-through displays using display color profiles," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '13. New York, NY, USA: ACM, 2013, pp. 231–240.

[7] T. Langlotz, M. Cook, and H. Regenbrecht, "Real-time radiometric compensation for optical see-through head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 11, pp. 2385–2394, 2016.

[8] J. Kim, J. Ryu, S. Ryu, K. Lee, and J. Kim, "[POSTER] Optimizing background subtraction for OST-HMD," in *2017 IEEE International Symposium on Mixed and Augmented Reality*, 2017, pp. 95–96.

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

[10] J. M. Wolfe, K. R. Kluender, D. M. Levi, L. M. Bartoshuk, R. S. Herz, R. L. Klatzky, S. J. Lederman, and D. M. Merfeld, *Sensation and perception, 4th ed.* Sinauer Associates, 2015.

[11] L. Gruber, D. Kalkofen, and D. Schmalstieg, "Color harmonization for augmented reality," in *2010 IEEE International Symposium on Mixed and Augmented Reality*, 2010, pp. 227–228.

[12] S. Liu, D. Cheng, and H. Hua, "An optical see-through head mounted display with addressable focal planes," in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2008, pp. 33–42.

[13] K. Oshima, K. R. Moser, D. C. Rompapas, J. E. Swan, S. Ikeda, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato, "SharpView: Improved clarity of defocussed content on optical see-through head-mounted displays," in *2016 IEEE Virtual Reality (VR)*, 2016, pp. 253–254.

[14] C. Menk and R. Koch, "Truthful color reproduction in spatial augmented reality applications," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 2, pp. 236–248, 2013.

[15] Y. Itoh, M. Dzitsiuk, T. Amano, and G. Klinker, "Semi-parametric color reproduction method for optical see-through head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 11, pp. 1269–1278, 2015.

[16] T. Oskam, A. Hornung, R. W. Sumner, and M. Gross, "Fast and stable color balancing for images and augmented reality," in *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization Transmission*, 2012, pp. 49–56.

[17] G. Wetzstein, W. Heidrich, and D. Luebke, "Optical image processing using light modulation displays," *Computer Graphics Forum*, vol. 29, no. 6, pp. 1934–1944, 2010.

[18] S. Mori, S. Ikeda, A. Plopski, and C. Sandor, "BrightView: Increasing perceived brightness of optical see-through head-mounted displays through unnoticeable incident light reduction," *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings*, pp. 251–258, 2018.

[19] Y. Itoh, T. Langlotz, D. Iwai, K. Kiyokawa, and T. Amano, "Light attenuation display: Subtractive see-through near-eye display via spatial color filtering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 5, pp. 1951–1960, 2019.

[20] J. D. Hincapié-Ramos, L. Ivanchuk, S. K. Sridharan, and P. P. Irani, "SmartColor: Real-time color and contrast correction for optical see-through head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 12, pp. 1336–1348, 2015.

[21] K.-H. Lee and J.-O. Kim, "Visibility enhancement via optimal two-piece gamma tone mapping for optical see-through displays under ambient light," *Optical Engineering*, vol. 57, no. 12, pp. 1 – 13, 2018.

[22] M. J. Murdoch, "Brightness matching in optical see-through augmented reality," *J. Opt. Soc. Am. A*, vol. 37, no. 12, pp. 1927–1936, 2020.

[23] M. A. Livingston, J. H. Barrow, and C. M. Sibley, "Quantification of contrast sensitivity and color perception using head-worn augmented reality displays," in *2009 IEEE Virtual Reality Conference*. IEEE, 2009, pp. 115–122.

[24] N. Hassani and M. J. Murdoch, "Investigating color appearance in optical see-through augmented reality," *Color Research & Application*, vol. 44, no. 4, pp. 492–507, 2019.

[25] H.-P. Huang, M. Wei, and S. Chen, "White appearance of virtual stimuli produced by augmented reality," *Color Research & Application*, vol. 46, no. 2, pp. 294–302, 2021.

[26] T. Uchida, K. Sato, and S. Inokuchi, "An optical see-through mr display with digital micro-mirror device," *Transactions of the Virtual Reality Society of Japan*, vol. 7, no. 2, 2002.

[27] Y. Itoh, T. Hamasaki, and M. Sugimoto, "Occlusion Leak Compensation for Optical See-Through Displays Using a Single-Layer Transmissive Spatial Light Modulator," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 11, pp. 2463–2473, 2017.

[28] T. Hamasaki and Y. Itoh, "Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 5, pp. 1961–1969, 2019.

[29] B. Krajancich, N. Padmanaban, and G. Wetzstein, "Factored occlusion: Single spatial light modulator occlusion-capable optical see-through augmented reality display," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, pp. 1871–1879, 2020.

[30] K. Kim, D. Heo, and J. Hahn, "Occlusion-capable head-mounted display," in *PHOTOPTICS*, 2019, pp. 299–302.

[31] K. Rathinavel, G. Wetzstein, and H. Fuchs, "Varifocal occlusion-capable optical see-through augmented reality display based on focus-tunable optics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 11, pp. 3125–3134, 2019.

[32] Q. Y. J. Smithwick, D. Reetz, and L. Smoot, "LCD masks for spatial augmented reality," in *Stereoscopic Displays and Applications XXV*, A. J. Woods, N. S. Holliman, and G. E. Favalora, Eds., vol. 9011, no. March 2014, 2014, p. 90110O.

[33] T. Rhodes, G. Miller, Q. Sun, D. Ito, and L.-Y. Wei, "A transparent display with per-pixel color and opacity control," in *ACM SIGGRAPH 2019 Emerging Technologies*, ser. SIGGRAPH '19. New York, NY, USA: ACM, 2019, pp. 5:1–5:2.

[34] A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2013, pp. 29–38.

[35] M. A. Livingston, J. L. Gabbard, J. E. Swan, C. M. Sibley, and J. H. Barrow, "Basic perception in head-worn augmented reality displays," in *Human factors in augmented reality environments*. Springer, 2013, pp. 35–65.

[36] J. Ryu, J. Kim, K. Lee, and J. Kim, "Colorimetric background estimation for color blending reduction of OST-HMD," in *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2016, pp. 1–4.

[37] R. O. Brown and D. I. MacLeod, "Color appearance depends on the variance of surround colors," *Current Biology*, vol. 7, no. 11, pp. 844–849, 1997.

[38] M. A. Webster, G. Malkoc, A. C. Bilson, and S. M. Webster, "Color contrast and contextual influences on color appearance," *Journal of Vision*, vol. 2, no. 6, pp. 7–7, 2002.

[39] V. Ekroll and F. Faul, "Transparency perception: the key to understanding simultaneous color contrast," *J. Opt. Soc. Am. A*, vol. 30, no. 3, pp. 342–352, 2013.

[40] S. Klauke and T. Wachtler, ""Tilt" in color space: Hue changes induced by chromatic surrounds," *Journal of Vision*, vol. 15, no. 13, pp. 17–17, 2015.

[41] D. Jameson and L. M. Hurvich, "Perceived color and its dependence on focal, surrounding, and preceding stimulus variables," *J. Opt. Soc. Am.*, vol. 49, no. 9, pp. 890–898, 1959.

[42] J. Krauskopf, Q. Zaidi, and M. B. Mandlert, "Mechanisms of simultaneous color induction," *J. Opt. Soc. Am. A*, vol. 3, no. 10, pp. 1752–1757, 1986.

[43] D. Jameson and L. M. Hurvich, "Opponent chromatic induction: Experimental evaluation and theoretical account," *J. Opt. Soc. Am.*, vol. 51, no. 1, pp. 46–53, 1961.

[44] V. Ekroll and F. Faul, "New laws of simultaneous contrast?" *Seeing and Perceiving*, vol. 25, no. 2, pp. 107 – 141, 01 Jan. 2012.

[45] S. Ratnasingam and B. L. Anderson, "What predicts the strength of simultaneous color contrast?" *Journal of Vision*, vol. 17, no. 2, pp. 13–13, 02 2017.

[46] J. L. Gabbard, J. E. Swan, J. Zedlitz, and W. W. Winchester, "More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces," in *2010 IEEE Virtual Reality Conference (VR)*, 2010, pp. 79–86.

[47] M. Mahy, L. Van Eycken, and A. Oosterlinck, "Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV," *Color Research & Application*, vol. 19, no. 2, pp. 105–121, 1994.

[48] J. Y. Lin, L. Jin, S. Hu, I. Katsavounidis, Z. Li, A. Aaron, and C.-C. J. Kuo, "Experimental design and analysis of JND test on coded image/video," in *Applications of Digital Image Processing XXXVIII*, A. G. Tescher, Ed., vol. 9599, International Society for Optics and Photonics. SPIE, 2015, pp. 324 – 334.

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

[50] S. J. Anderson, K. T. Mullen, and R. F. Hess, "Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors." *The Journal of Physiology*, vol. 442, no. 1, pp. 47–64, 1991.

[51] B. C. Kress and W. J. Cummings, "11-1: Invited Paper: Towards the ultimate mixed reality experience: HoloLens display architecture choices," *SID Symposium Digest of Technical Papers*, vol. 48, no. 1, pp. 127–131, 2017.

Yunjin Zhang Yunjin Zhang received the bachelor's degree in Software Engineering from University of Electronic Science and Technology of China, in 2016. He is working toward a Ph.D. degree in the State Key Laboratory of CAD&CG, Zhejiang University. His interests are mainly in the optimization framework for real-time rendering, especially power-budget rendering, VR/AR optimization, and cloud-client co-computing rendering.

Rui Wang Rui Wang is a professor at the State Key Laboratory of CAD&CG, Zhejiang University. He received a bachelor's degree in Computer Science and a Ph.D. degree in Mathematics from Zhejiang University. His research interests are mainly in real-time rendering, realistic rendering, GPU-based computation, and 3D display techniques. Now, he is leading a group working on the next-generation real-time rendering engine and techniques.

Yifan (Evan) Peng Yifan (Evan) Peng is a Postdoctoral Research Fellow at Stanford Electrical Engineering. He received a Ph.D. in Computer Science from University of British Columbia, and an M.Sc. in Optical Science & Engineering from Zhejiang University. His research focuses on incorporating optical and computational techniques for enabling new imaging modalities. He has been working on computational imaging & display with wave optics.

Wei Hua Wei Hua is an associate professor with the State Key Laboratory of CAD&CG, Zhejiang University. He got a Ph.D. in Applied Mathematics from Zhejiang University in 2002. He joined the CAD&CG State Key Lab in 2002. His main interests include real-time simulation and rendering, virtual reality, and software engineering.

Hujun Bao Hujun Bao is a professor with the State Key Laboratory of CAD&CG and the College of Computer Science and Technology, Zhejiang University. He leads the 3D graphics computing group in the lab, which mainly makes researches on geometry computing, 3D visual computing, real-time rendering, and their applications. His research goal is to investigate the fundamental theories and algorithms to achieve good visual perception for interactive digital environments, and develop related systems.