https://doi.org/10.2352/CIC.2022.30.1.05

Adaptive Display White Point for Enhancing Viewing Experience of Mixed Reality Headsets

Minchen Wei[1,], Wenyu Bao[1], Zheng Huang[1], Jan Oberländer[2], Stefan Rüffer[2], Jerry Jia[2]*
[1]Color, Imaging, and Illumination Laboratory, The Hong Kong Polytechnic University; [2]Meta Reality Labs
minchen.wei@polyu.edu.hk

Abstract

The development of various new imaging systems introduces new viewing conditions that did not exist in the past. Mixed reality systems, which capture the real environment and render the captured image on display with virtual objects superimposed, require more immersive and realistic feeling than virtual and augmented reality systems, especially when users just put on the headsets. Therefore, the rendering shown on the display needs to be carefully adjusted to match the appearance in the real environment. In this study, we specifically focus on reproducing the overall color tone of the real environment under different ambient illumination colors by shifting the white point of the display. The human observers viewed a real environment under different ambient illumination conditions, in terms of CCT and chromaticities, and evaluated the rendering of the captured scene with 44 white points. The results clearly suggested that the display white point should be adaptive to the ambient illumination color, especially when the ambient illumination had a CCT below 4000 K, to provide a good user experience.

Introduction

Mixed reality (MR) is a feature for virtual reality (VR) headsets. In VR, the direct view of outside real world is completely blocked and replaced by a display and a lens placed in front of the eyes. MR is achieved using front-facing cameras mounted on VR headset to capture the outside real world and show that on the VR display. Frequently, virtual content (independently generated by computer graphics) is overlayed on top of the captured environment. Color and 3D depth performances are two critical considerations in MR pipeline, with Figure 1 showing a representative color pipeline.

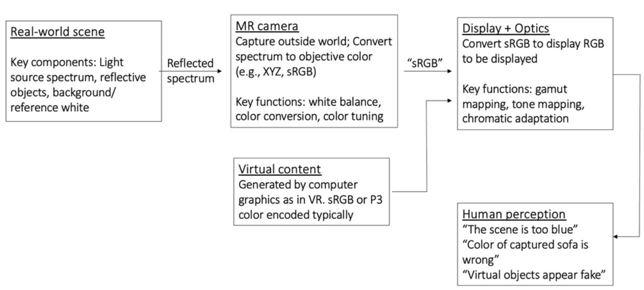

Figure 1 Representative color processing pipeline for mixed reality headsets.

We believe MR's color performance needs to consider the following three aspects: 1) scene consistency (e.g., reference white) between real and captured environments, 2) object color consistency (e.g., the captured object color needs to match that of the real object), and 3) virtual objects matching captured scene (e.g., graphically generated virtual objects need to fit naturally in the captured scene). Our work in this paper focuses on 1) and briefly touches on 2) and 3). It's important to define the terms: "real" refers to real objects/scenes in the real world, "virtual" refers to graphically generated objects or scenes such as those used in VR, and "captured" refers to real-world objects/scenes captured by the MR camera and shown on the VR display.

The chromatic adaptation mechanism in the human visual system in real world can automatically account for the variation of ambient illumination colors, with the perceived color appearance of surfaces remaining relatively constant. Electronic displays, however, are self-luminous, with the spectral composition of the light emitted from displays not varying with the ambient illumination. Such a difference motivates consumer products to have displays that are adaptive to ambient conditions [1]. This is a complex topic both perceptually and as a product feature, since users do not expect self-luminating displays to behave exactly like real-world reflective objects [2-6].

MR is different from traditional displays viewed under an ambient illumination condition. In MR, users see the rendering of the captured environment merged with the graphically generated virtual contents on the headset display, as if the real environment and the residing virtual contents are viewed directly without the headset. However, users do not experience the real ambient lighting and display simultaneously as in traditional consumer products (e.g., phones, tablets). For MR, we expect similar adaptive display with its color (brightness and chromatic components) dynamically changing with the ambient illumination, particularly at the moment of putting on headsets. A more tolerant adaptive performance, however, is expected.

Moreover, the near eye displays in MR headsets generally have a lower luminance and cover a much larger field of view (FOV) with a darker surround. The effects of these differences have not been investigated in the past and are expected to affect the feature algorithm.

In this study, while we aim to understand the human vision preference in the new MR environment, we have a strong emphasis to build a practical and complete solution to bring the MR feature to mass-manufactured products.

Specifically, we are discussing: 1) a new imaging and display pipeline that optimizes color accuracy instead of color "pleasantness" as in traditional digital photography, 2) a method of using camera to find and understand the light source in the real environment for camera's white balance and display's white point, and 3) a new simpler chromatic adaptation hinged on the blackbody/daylight curve that is perceptually adequate and engineering-wise practical, with algorithm and specification of how headset display white point should be adjusted under different ambient illuminations.

Methods

Apparatus

A viewing booth with interiors painted using Munsell N7 spectrally neutral paint under the illumination of two 11-channel spectrally tunable LED lighting devices was used to simulate a real environment under different ambient illumination conditions. A chin rest was mounted just outside the viewing booth, centered on the opening, and the top part of the front opening was partially covered using black felt to prevent the observer from seeing the LED devices directly. The observer was asked to fix his or her chin on the rest when viewing the real environment, so that the real environment in the viewing booth covered his or her entire field of view. Man-made objects and replica of familiar objects were selected to cover different hues and placed inside the booth, as shown in Figure 2.

An MR simulator based on a virtual light booth was built to simulate a similar viewing condition as an MR headset. The virtual light booth is a light booth and a 43-in 4K display placed behind the booth. The interiors of the viewing booth were covered using black diffuse coating with an average reflectance of 4%. A 15 cm × 15 cm opening was cut at the center of the front panel, with a chin rest mounted outside the booth aligned with the opening. During the experiment, the display was viewed by the observer through the opening, with a field of view approximately 86° (horizontal full) × 53° (vertical full). Such a setup was considered acceptable and almost desired, as the color perception can be somewhat separated from 3D depth perception, so that we do not need a complex 3- or 6-degree of freedom (DOF) headset. Also, it allows us to accomplish very accurate color calibration, capture and reproduction, which is critically important to carry out user studies as described in this work. No commercially available headsets can achieve such high color accuracy and across FOV. The viewing booth and the MR simulator were placed at 90° from each other with respect to observer, as illustrated in Figure 3, so that the observer can switch between the real environment and the MR simulator easily.

Figure 2 A photograph of the real viewing booth under the illumination of the spectrally tunable LED lighting devices. This photograph is then reproduced with a validated color accuracy on the MR simulator booth (see Figure 5 and description).

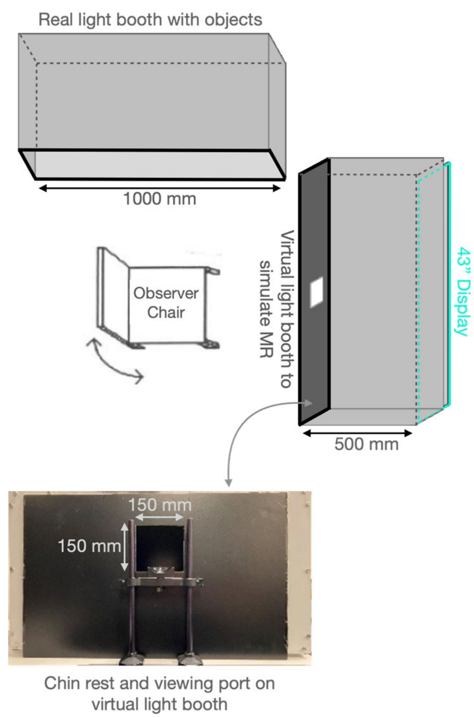

Figure 3 Illustration of a light booth to simulate real environment and the MR simulator by a virtual light booth, comprising of a viewing booth and a 43-inch display. The display was placed behind the booth, and the observer viewed the display through a 15 cm × 15 cm opening, with the interiors covered using diffuse black coating.

Display calibration, lighting condition design, and reproduction of captured scene on display

The display had a nominal peak luminance of 350 cd/m^2 and an sRGB color gamut. To characterize the uniformity across field of view, the display was divided into 24 regions (4 columns × 6 rows) and 25 RGB combinations were shown at each region, with the spectral power distribution (SPD) measured using a PhotoResearch PR-655 spectroradiometer. The uniformity was characterized using the mean color difference from the mean (MCDM) by calculating the difference between the chromaticities measured at each region and the average chromaticities of the 24 regions in the CIE 1976 *u*'*v*' chromaticity diagram. The MCDM values of the 25 RGB combinations ranged between 0.0005 (dE00 of 0.40) and 0.0018 (dE00 of 1.42), with an average of 0.0011 units (dE00 of 0.88). A gain-offset-gamma (GOG) model was then used to calibrate the display at the center region, where the observer would view when looking straight forward.

Based on the past work [4,5], the display white point between 2700 and 7000 K was believed to be wide enough for producing a similar overall color appearance of a real environment under the illumination between 2700 and 6500 K, which is the range of general illumination. Forty-four lighting conditions, with a CCT from 2700 to 7000 K with an interval of 100 K, were created to illuminate the viewing booth. The scene inside the viewing booth was then captured using a PhaseOne XF IQ3 camera in the RAW format to produce images at these 44 white points. All the lighting conditions were calibrated by adjusting the intensities of the 11 channels in the LED devices, so that their chromaticities were close to the blackbody locus and daylight locus when below and beyond the 5000 K respectively, and their CIE general color rendering index (CRI) values were as high as possible. The colorchecker placed in the viewing booth was used to calibrate each photograph, with a transformation matrix $M_{RGB2XYZ}$ derived using the measured tristimulus values XYZ and the extracted RAW RGB values of the 24 color patches. This matrix, together with the GOG model of the display, was then used to produce an accurate rendering of the real environment under each lighting condition, so that the XYZ values of the image shown on the MR simulator display were similar to those in the real environment under the corresponding illumination, which was verified through measurements. Though the display had a nominal peak luminance of 350 cd/m^2, the luminance levels of the images were adjusted with the white patch on the colorchecker being 57 cd/m^2, so that all the images, especially those under the low CCT levels, were within the display color gamut, and such a level is similar to the display luminance level in an MR headset (50-120 cd/m^2).

During the first study, the real light booth (simulating real environment) was viewed under eight lighting conditions having the chromaticities on the blackbody/daylight locus, including three (i.e., 2700, 3000, and 3200 K) having the same luminance as the images shown on the display (i.e., L_w = 57 cd/m^2, E ≈ 179 lx at booth floor surface) and five (i.e., 3200, 3500, 4000, 5000, and 6500 K) having a slightly higher luminance than the images (i.e., L_w = 89 cd/m^2, E ≈ 279 lx at booth floor surface). Though such a slight difference in luminance was not expected to cause any difference, the experiments were repeated under the CCT of 3200 K with the two light levels for verification. These two light levels were within the range of typical general illumination. In addition, it is worthwhile to point out that the smaller CCT interval between 3000 and 3500 K was purposely designed, as past work suggested that the degree of chromatic adaptation changed fast within this range [4,5]. Figure 4 shows the chromaticities of the 44 light settings that were used to calibrate the camera, and those of the eight light settings in the real environment. Figure 5 shows the examples under the 2700 and 6500 K illumination conditions, illustrating that the overall color tone of the real environment and that of the calibrated rendering shown on the display were similar.

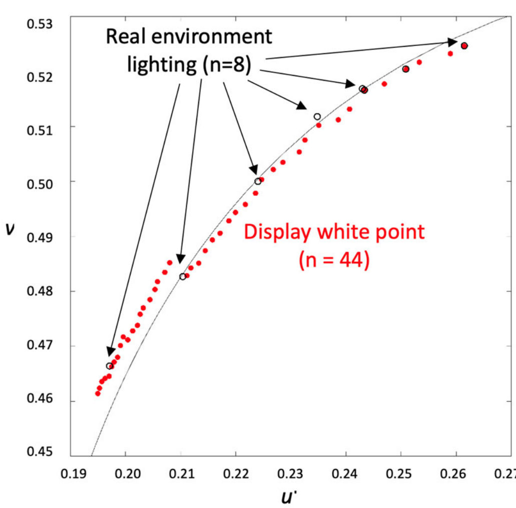

Figure 4 Chromaticities of the 44 light settings that were used for the camera calibration and those of the eight light settings in the real environment in the CIE 1976 u'v' chromaticity diagram.

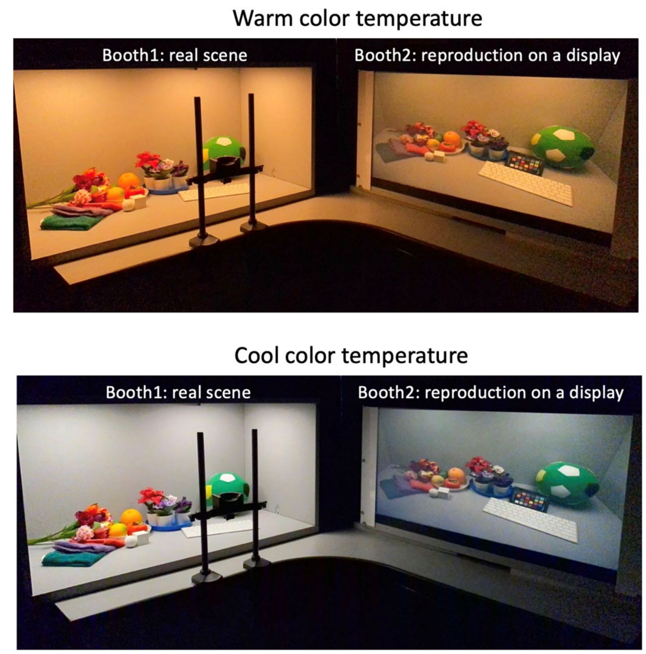

Figure 5 Illustration of the similar color tone between the real environment and calibrated rendering shown on the display when the light setting was 2700 or 6500 K. (note: the setup of the booth on the right was modified here for illustration. The setup used in the experiment is shown in Figure 1. For taking these photographs, we temporarily took off the cover on the virtual light booth and moved the monitor to the front opening).

Experiment Procedure

The experiment was designed by asking the human observers to evaluate the similarity between the overall color tone of the ambient illumination in the real scene and that shown on the display based on their short-term memory [2]. Under each lighting condition, each observer was asked to be seated in front of the viewing booth simulating the real environment first, with his or her chin fixed on the rest, to look at the viewing booth for two minutes for chromatic adaptation, and to remember the overall color appearance of the environment rather than the color appearance of the individual objects. After two minutes, the observer immediately switched to the MR simulator and looked at the display, with his or her chin fixed on the rest. The image with a certain white point was shown on the display for four seconds, then the image disappeared and the observer was asked to rate how similar the overall color appearance of the content appeared in comparison to the real environment he or she just viewed in the other viewing booth, using a five-point rating scale (i.e., 1: Identical; 2. A little difference; 3. Noticeable difference but acceptable; 4. Large difference; 5. Completely different). The four-second duration was purposely designed to simulate the condition that the MR headset is just put on and the visual system has not been adapted to the display white point and it was long enough for the observer to make the evaluation. After making the evaluation, the observer switched back to the real viewing booth to view the environment under the same lighting condition for 15 seconds, which was long enough for chromatic adaptation as he or she just left the same condition for less than 10 seconds. The same procedure was repeated for all the 45 images (i.e., 44 white points and one evaluated twice for characterizing the intra-observer variations) in a random sequence with the same lighting condition in the viewing booth. The order of the eight lighting conditions was also randomized, with each observers made 360 evaluations in total (i.e., 45 evaluations x 8 ambient lighting).

Results

User preference: ambient illumination with chromaticities on the blackbody curve

With each of the eight real environment lighting conditions shown in Figure 4, we varied the display white points and evaluated observers' preferences. Figure 6 shows the average rating of the images with the 44 white points evaluated by nine observers for each ambient lighting condition. Among the nine observers, the average standard deviation of the ratings was 0.29 across all the 360 evaluations. As highlighted in Figure 7, the optimal white point (i.e., the white point having the lower average rating) generally shifts with the CCT of the ambient lighting condition. It can be observed that the two light levels for the 3200 K did not cause much difference to the observers' evaluations.

The results obtained under the seven CCT levels can be classified into three groups—2700 K, 3000 to 3500 K, and 4000 to 6500 K. If the white point of the rendering was set to the traditional display white point (i.e., D65), when the real environment was viewed under the lighting conditions below 4000 K, the difference of the overall color tone between the rendering and the real environment was rated very large and not acceptable. In particular, a 300 K increase from 2700 to 3000 K introduced a significant difference, which corroborated the findings in the past studies [4-6]. These results clearly suggest that it is necessary to shift the white point of the MR headset display according to the lighting condition in the real environment, and special attention is needed when ambient lighting CCT is below 4000K (warm lighting).

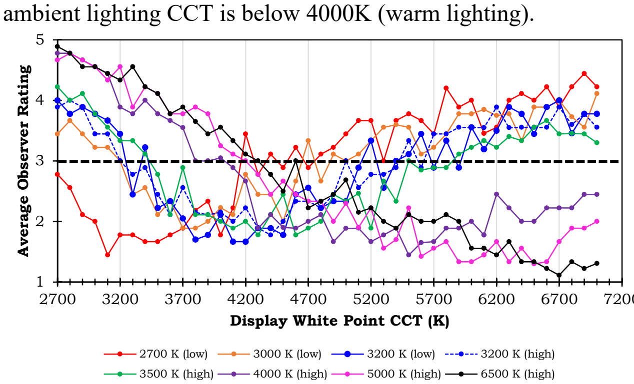

Figure 6 Average observer ratings on the similarity between each of the images with the 44 white points shown on the display and real environment under the eight ambient illumination conditions, in terms of the overall color tone of scene.

With the user preference result, it is straightforward to adjust the display white point according to the ambient illumination CCT. More precisely, the full image's colors (not just white) are simultaneously adjusted by a chromatic adaptation transform (e.g., CAT02). Figure 8 illustrates such an effect with the same physical scene under different ambient illumination colors. Note that we are using the higher CCT bound highlighted in the green region in Figure 7. For example, when the real environment has a CCT of 2700 K, the display white point is set to 4000 K, instead of 3100 K (the values following the red line in Figure 7). There are a couple of reasons. One is object color consistency mentioned in introduction. The captured object color needs to match that of the real object. A similar user experiment focusing on object color consistency indicated a preference of higher CCT for white point than the results shown in Figure 7. Another reason is a product choice with balance in power efficiency, perceived brightness, and user demographics.

Optimal display White Point range

Display White Point CCT (virtual light booth)	2700	3000	3200	3200	3500	4000	5000	6500
6900	4.4	3.6	3.8	3.8	3.4	2.4	1.9	1.2
6700	3.9	4.0	4.0	3.9	3.4	2.2	1.7	1.1
6500	4.0	3.9	3.4	3.6	3.6	2.0	1.3	1.3
6300	4.0	3.8	3.9	3.6	3.4	2.2	1.3	1.7
6100	3.5	3.9	3.2	3.6	3.3	1.8	1.4	1.6
5900	3.9	3.8	2.9	3.6	3.1	1.9	1.3	2.0
5700	3.4	3.2	2.9	3.4	2.9	1.7	1.6	2.0
5500	3.8	3.6	3.1	3.3	3.0	1.4	2.2	2.1
5300	3.0	3.6	2.6	2.8	2.7	1.8	1.6	2.0
5100	3.7	3.1	2.9	2.6	2.5	1.9	1.9	2.2
4900	3.2	3.1	2.3	2.4	2.4	1.7	2.0	2.4
4700	2.9	3.3	2.6	2.3	1.9	2.0	2.3	2.2
4500	2.9	2.0	1.8	2.0	2.4	1.9	3	2.5
4300	2.8	2.4	1.9	1.9	1.8	1.9	2.8	3.0
4100	2.2	2.1	1.7	2.0	1.9	2.9	3.1	3.3
3900	2.3	2.0	1.8	2.1	2.1	3.0	3.8	3.4
3700	1.9	1.9	2.1	2.6	2.9	3.6	3.8	3.9
3500	1.7	2.1	2.2	2.4	2.8	3.8	4.1	4.1
3300	1.8	2.4	2.4	2.8	3.3	3.8	3.9	4.6
3100	1.4	3.2	3.7	3.4	3.6	4.4	4.3	4.4
2900	2.1	3.4	3.9	3.9	4.1	4.7	4.7	4.6
2700	2.8	3.4	4.0	3.9	4.2	4.8	4.7	4.9

Ambient Illumination CCT (real light booth)

Figure 7 Average observer ratings on the similarity between each of the images with the 44 white points shown on the display and real environment under the eight ambient illumination conditions, in terms of the overall color tone of scene. The green region highlights the ranges of display white points that were rated below 2.5, and the red line shows the trend of the lowest rating (most consistent white point under each illumination).

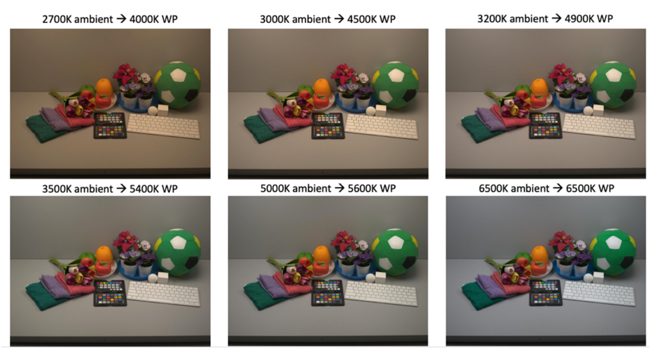

Figure 8 Illustration of adjusting display white point and the image color for a same physical scene under different ambient illumination colors.

User preference: ambient illumination with chromaticities off the blackbody curve

The development of LED lighting in recent years has allowed the adjustment of SPDs and chromaticities much easier. Scientific studies have found that the illumination with chromaticities off the blackbody locus can improve the color quality of the illumination, which has led to the development of commercial products [7,8,9]. Thus, we carried out a follow-up experiment by shifting the chromaticities of the illumination in the real environment away from the blackbody locus, while fixing the display white point on the blackbody/daylight locus, as shown in Figure 9. The LED devices were adjusted to produce 10 light settings, comprising five CCTs (i.e., 2700, 3000, 3500, 4000, and 6500 K) and two D_{uv} levels (i.e., +0.01 and -0.02), as shown in Figure 9. These two D_{uv} levels generally covered all the possible ranges of conditions in daily life [9]. The conditions having a CCT of 2700 and 3000 K were under the lower light level, while the others were under the higher light level. The same set of images with the 44 white points were used. The experiment procedure was the same as the main experiment.

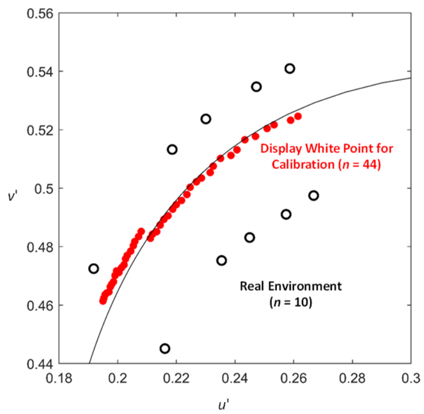

Figure 9 Chromaticities of the 44 light settings that were used for the camera calibration and those of the 10 light settings in the real environment with the chromaticities far from the blackbody locus.

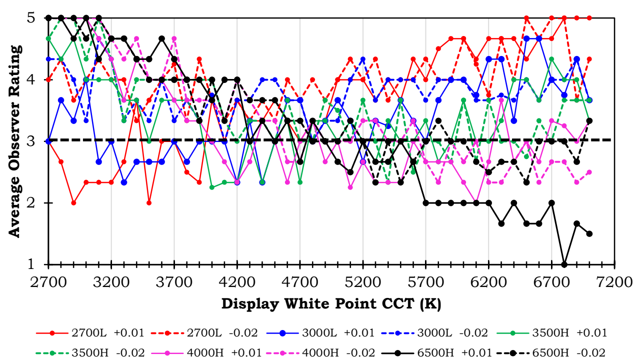

Figure 10 Average observer ratings on the similarity between each of the images with the 44 white points shown on the display and real environment under the 10 ambient illumination conditions, in terms of the overall color tone of scene.

Figure 10 shows the average observer rating on each of the 44 images under the 10 light settings in the real environment. It can be observed that the general trend was consistent to that in the main experiment. However, the observers generally evaluated a larger difference (i.e., higher rating) between the color tone of the rendering and that of the real environment, regardless of the CCT, which was caused by shifting the chromaticities of the ambient illumination off the blackbody locus. Also, the difference between the results obtained under the two conditions having the same CCT but different D_{uv} was the largest when the ambient illumination was 2700 K.

Compared to viewing a standalone display under an ambient illumination, MR feature in a VR headset generally has a less stringent requirement on the accuracy for chromatic adaptation. In MR, observers never see ambient light (or reference white) directly and the adaptation of eye in the headset can be fully controlled. The lighting conditions selected in Figure 9 represent the boundary of typical conditions in daily life. The result shown in Figure 10, the consideration of MR uniqueness, and its simplicity led us to the decision to move the white point along the blackbody/daylight locus.

Imaging and display pipeline that maximizes color accuracy

We know how to adjust display white point based on the CCT of the ambient illumination, but how do we know the CCT reliably? The imaging/display pipeline also heavily relies on correct white balancing during the capture of the real scene. Traditional auto white balancing sometimes varies when the camera view is changed. This can be undesirable in MR, especially when the ambient lighting is expected to be constant (e.g., people wearing MR headset in their living room). To solve both challenges, we developed a method to train the camera to detect the location and CCT of the light source.

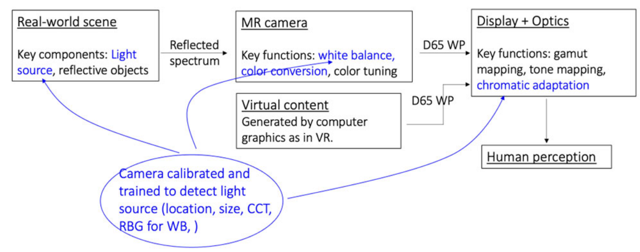

Figure 11 Modified color pipeline for adaptive display in MR. Camera is calibrated and trained to estimate the CCT of the ambient light source and reliable white balance, both of which are key to accurate and pleasant color experience in MR.

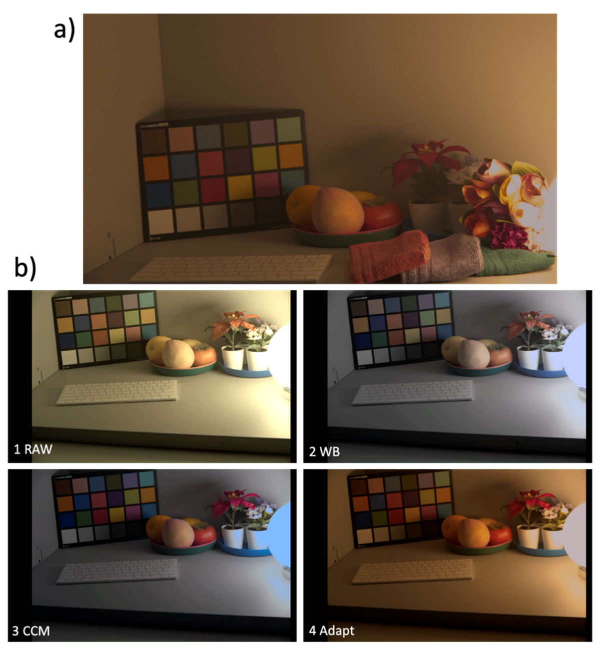

Figure 12 a) Ambient environment inside the real viewing booth, b) the imaging and display pipeline realized on the MR simulator/virtual light booth. From 1-4: 1) the raw image taken by the MR camera, 2) in previous frames (not shown here) the camera sensitivity was adjusted to detect the light source (in the right part of the images above) without being overexposed. The RGB readings inside a region of interest were used to adjust the white balance, and to estimate the CCT of the light source, 3) image after applying the color transformation matrix which also utilized the estimated CCT of the light source, and 4) after applying the chromatic adaptation transform before showing on the display. It is a good reproduction of the real world in a).

Fortunately, in MR, unlike instant photography, we have more time to make things right. At the initial setup period, we lower the sensitivity of the camera to detect the light sources in the environment. The RGB ratios on the source can be used to perform white balance, which is more reliable than making the estimations based on the captured image. This method can be combined with traditional auto white balance and auto brightness in typical camera pipeline (e.g., when the algorithm fails to detect a light source) or combined with signal from ambient light sensor. When calibrated, the RGB readings from the camera can further measure the CCT of the light source. The CCT can then be used to conduct chromatic adaptation transformation at the display, after combining captured scene and virtual content under the same D65 white point. We find this method most straightforward and reliable among other potential methods of combining virtual and captured contents. Figure 11 shows the modified imaging/display color pipeline to achieve such a process, and Figure 12 shows an example of this pipeline realized in the MR simulator/virtual light booth.

Conclusion

Two experiments were designed to investigate how MR headset display should be adapted based on the ambient illumination color in the real environment for a better viewing experience. An MR simulator was built to provide a similar viewing condition as in an MR headset. A viewing booth simulating a real environment was illuminated under a wide range of ambient illumination colors, in terms of CCT and chromaticities (i.e., D_{uv} level). This setup provides necessary color accuracy for the experiments that is not readily achievable in a real headset.

The observers viewed the real environment under an illumination and then evaluated the rendering of the environment with a certain white point shown on the display in the MR simulator, which was well calibrated. The results suggest that the display white point and color transformation should be adaptive to the ambient light color instead of being fixed to D65, especially when the ambient light has a lower CCT level (i.e., below 4000 K).

We also demonstrated a new imaging and display pipeline for MR that optimizes color accuracy instead of color "pleasantness" as in traditional digital photography. A method of using camera to find and estimate the CCT of the light source in the environment for performing camera white balance and display white point adaptation was introduced. This method and the simpler CCT-based chromatic adaptive display represent a good product solution for a delightful MR color experience.

Reference

[1] Wu J, Zhang L, Isikman S, Chen C. 2019. Enhanced viewing experience considering chromatic adaptation. SID 2019:857-860.

[2] Fairchild MD. 2013. Color Appearance Models, 3rd edition.

[3] Huang HP, Wei M, Ou LC. 2018. White appearance under different ambient lighting conditions. Optics Express. 26(4):5018-5030.

[4] Wei M and Chen S. 2019. Effects of adapting luminance and CCT on appearance of white and degree of chromatic adaptation. Optics Express. 27(6):9276-9286.

[5] Huang Z and Wei M. 2021. Effects of adapting luminance and CCT on appearance of white and degree of chromatic adaptation, part II: extremely high adapting luminance. Optics Express. 29(25):42319-42330.

[6] Zhai Q and Luo MR. 2018. Study of chromatic adaptation via neutral white matches on different viewing media. Optics Express. 26(6):7724-7739.

[7] Wang Y and Wei M. 2018. Preference among light sources with different Duv but similar color rendition: a pilot study. Lighting Research & Technology. 50:1013-1023.

[8] Wei M and Houser KW. 2016. What is the cause of apparent preference for sources with chromaticity below the blackbody locus? LEUKOS. 12:95-99.

[9] Wei M and Houser KW. 2012. Status of solid-state lighting based on entries to the 2010 US DOE Next Generation Luminaires Competition. LUEKOS. 8:237-259.

Author Biography

Minchen Wei is an associate professor at The Hong Kong Polytechnic University. He obtained his PhD degree at The Pennsylvania State University. His research interest focusses on color science and imaging science.

Wenyu Bao just obtained her PhD degree at The Hong Kong Polytechnic University, and she obtained her bachelor's degree at Fudan University.

Zheng Huang is a PhD student at The Hong Kong Polytechnic University. He obtained his bachelor and master's degree at Wuhan University.

Jan Oberländer is a software engineer in mixed reality at Meta Reality Lab in Zürich Switzerland. He got his master's degree in computer science at Karlsruhe Institute of Technology.

Stefan Rüffer is a technical program manager in mixed reality at Meta Reality Lab in Zürich Switzerland. He got his master's degree at Technische Universität Ilmenau.

Jerry Jia is a system engineer and human perception specialist at Meta Reality Lab in California US. He advocates for experience-centric product design and an integration of hardware system, algorithms, and human vision in product development of VR/AR. He received his Ph.D. in Material Science and Engineering from Columbia University in the City of New York.