Towards Predictive Virtual Prototyping: Color Calibration of Consumer VR HMDs

Olaf Clausen*†, Gregor Fischer *, Arnulph Fuhrmann * and Ricardo Marroquim ⋆

* TH Köln, Germany
olaf.clausen@th-koeln.de
gregor.fischer@th-koeln.de
arnulph.fuhrmann@th-koeln.de

† UFRJ, Brasil

⋆ TU Delft, Netherlands
r.marroquim@tudelft.nl

Abstract: Nowadays, virtual prototyping is an established and increasingly important part of the development cycle of new products. Often CAVEs and Powerwalls are used as Virtual Reality (VR) systems to provide an immersive reproduction of virtual content. These VR systems are space and cost-intensive. With the advent of the recent consumer Virtual Reality Head Mounted Displays (VR HMDs), HMDs got more attention from the industry. To increase the acceptance for HMDs as VR system for virtual prototypes, color consistency has to be improved. In this paper, we present an approach to characterize and calibrate displays of consumer VR HMDs. The approach is based on a simple display model, which is commonly used for calibration of conventional displays, but has not yet been applied for VR HMDs. We implemented this approach with the HTC Vive Pro and the Pimax 5k+. In combination with our calibration approach, the Vive Pro provides a color reproduction without perceivable color differences.

Keywords: Color Calibration, Color Consistency, VR HMD, Virtual Prototyping

1 Introduction

The development of a new product is a time and cost-intensive process. Decisions made during the design stage are particularly critical and have a determining effect on the success of the final product. The fact that consumers nowadays desire individual and customized products renders the design process even more challenging. Consequently, an important aim is to do prototyping as early as possible in the development cycle. Prototypes enable to optimize product design and, thus, to eliminate weaknesses in the early development stage, leading to reduced development time and overall costs.

Prototypes can be roughly grouped into two categories: physical and virtual. A physical prototype is the representation of an idea often made out of wood, clay, foam, or metal, which does not necessarily have the same properties and functionality as the final product. A virtual prototype (VP), as defined by Jimeno and Puerta [JP07], is the construction of product models using computers, frequently in a virtual environment. A more detailed definition from the computer graphics perspective is given by de Sa and Zachmann [GdSZ98], where a VP is the application of virtual reality for prototyping physical mock-ups with all relevant characteristics being simulated and rendered as precise and realistic as possible.

Although physical prototypes shorten the production cycle leading to cost reductions, they are still less time and cost-effective than virtual prototypes. VPs have various benefits [GSA01], where the most significant are the further reduction in development time and manufacturing costs, and the possibility for a collaborative design that overcomes the geographical distance. However, VPs can not completely replace physical prototypes yet due to some of their still unresolved weaknesses, such as the limited fidelity of rendered images reproduced on a display device, and the limited immersion of VR systems. Both can lead to wrong design decisions, which might not have been made with the physical prototypes. Nevertheless, the increasing use and the already important role of VPs in many industrial areas (e.g. automotive, aerospace and architecture), show that these weaknesses are being overcome and the benefits of VPs progressively outweigh its disadvantages.

In the past Powerwalls or CAVEs were often used as VR systems for VPs, since they provide stereo reproduction and head tracking. Unfortunately, these VR systems require a large physical space and are expensive. Furthermore, the immersion is often disturbed by the mixture of virtual and real environments. Virtual Reality Head Mounted Displays (VR HMDs) have become more attractive for the industry with the advent of the recent consumer devices, such as the Oculus Rift and the HTC Vive. Current consumer HMDs provide a reasonable display resolution, a large field of view, and an accurate head tracking for a low price while delivering similar or even higher immersion as CAVEs [SPL+16] or Powerwalls.

There has been extensive research on tracking, display resolution, and immersion of VR HMDs. However, to the best of our knowledge, so far there has been no work dealing with the display characterization and calibration from a color perspective. Referring to Greenberg et al. [Gre99] the visual display algorithm is one of three research areas handling the fidelity of physical simulations besides the local light reflection model and the energy transport simulation. In this work, we address this last topic by presenting a simple display model for VR HMDs and two frameworks for verifying the simulated and the real application of the display model. We implemented the display model and the proposed frameworks using the HTC Vive Pro and the Pimax 5k+.

The work is divided into five sections. In Section 2 literature dealing with the color calibration of VR systems is presented. In Section 3 first, the display model is presented and second, the two verification frameworks are described. In Section 4 the Vive Pro and the Pimax are calibrated and the two described frameworks are conducted. Finally, conclusions and future work are given in the last section.

2 Related Work

There are only a few works that deal with the color calibration of VR systems, and none considers the color calibration of VR HMDs. Some approaches describe the characterization and/or calibration of a stereoscopic projection system based on the Infitec color separation technology [KRK03, GBR+08, JJB10]. The Infitec technology requires two projectors with slightly different primary valences, which is often achieved by using interference filters in front of the projectors. The user wears special glasses with the corresponding filters, which separates the full-color image for the left and right eye [JSF08].

Kresse et al. [KRK03] built two multi-projector displays using the Infitec technology; the digital CAVE and the HEyeWall (a high-quality, high-resolution and stereo display). To provide color consistency they determined the common color gamut of all projectors and the gamma curves for each color channel. Adjusting the input-RGB values accordingly improves the color consistency but still leads to clearly discernible color differences between different projectors. Furthermore, using a common color gamut decreases the color gamut and contrast ratio compared to the native properties of the projectors.

Gadia et al. [GBR+08] present a virtual reality theater, which uses a stereo multi-projector display similar to the HEyeWall. In their work, they determine an accurate spectral and colorimetric characterization of the VR theater. They concluded that the filters reduce the maximum luminance from $44\frac{cd}{m^2}$ to $15\frac{cd}{m^2}$, and that in particular the common gamut negatively affects the red and blue channels, whereas for the green it does not impose great restrictions.

Gerhardt et al. [JJB10] extend the approach of Kresse et al. [KRK03] by adding a color difference threshold. In an optimization process, the common gamut is maximized while the color differences are within the defined threshold. In their work they recommend a threshold of $\Delta E_{ab} = 20$ for a white uniform patch. This threshold is clearly above the just noticeable difference (JND) of $\Delta E_{1976} = 2.7$.

Projector based VR systems are costly and require a lot of space. It is difficult to calibrate them properly and even calibrated systems are insufficient for any appearance critical virtual prototyping application. Recently, Infitec published new filters with more spectral bands. Referring to their website [Inf] these filters lead to more accurate color reproduction. Unfortunately, no evaluation results are available yet.

3 Definition and verification of the display model

3.1 Display Model

To calibrate a VR HMD we use the display model as shown in Figure 1. Our model assumes that each pixel consists of three independent light sources with the constant spectral compositions $r(\lambda)$, $g(\lambda)$ and $b(\lambda)$. Their intensities are controlled by linear gain factors RGB_{Lin} within the range $[0, 1]$. The relationship between the RGB_{Lin} and the input RGB signals is described by the gamma curves, and is, consequently, non-linear. The three additive primary colors are overlaid by an independent black component caused by the residual light from the display panel and by a reflection component caused by ambient light. For HMDs the reflection component is negligible, due to the encapsulation of the displays from ambient light.

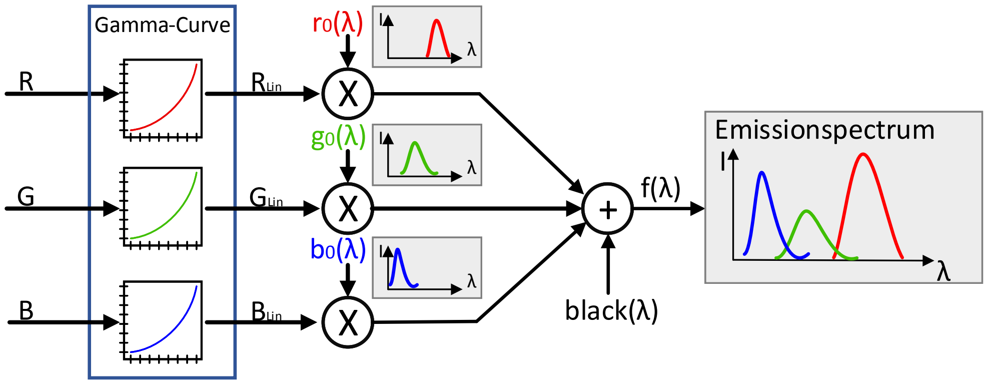

Figure 1: Display model.

According to the display model in Figure 1 the color stimulus $f(\lambda)$ is linearly composed of the intensities of the modulated three primary colors and the black spectrum, as shown in Equation 1:

$$f(\lambda) = black(\lambda) + R_{Lin} \cdot r_0(\lambda) + G_{Lin} \cdot g_0(\lambda) + B_{Lin} \cdot b_0(\lambda) \tag{1}$$

Based on this equation a relationship between the RGB_{Lin} signals and the resulting display XYZ (XYZ_{Disp}) values can be derived. The XYZ_{Disp} values are computed by multiplying and integrating the color stimulus $f(\lambda)$ with each of the three CIE color matching functions:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{Disp} = \int\limits_{\lambda} \begin{pmatrix} \overline{x}(\lambda) \\ \overline{y}(\lambda) \\ \overline{z}(\lambda) \end{pmatrix} \cdot f(\lambda) d\lambda \tag{2}$$

Substituting $f(\lambda)$ from Equation 1 into Equation 2 leads to Equation 3, which can be then resolved in terms of the RGB_{Lin} signals as detailed in Equation 4.

$$\begin{aligned} &\int\limits_{\lambda} \begin{pmatrix} \overline{x}(\lambda) \\ \overline{y}(\lambda) \\ \overline{z}(\lambda) \end{pmatrix} \cdot black(\lambda) d\lambda + \begin{pmatrix} \overline{x}(\lambda) \\ \overline{y}(\lambda) \\ \overline{z}(\lambda) \end{pmatrix} \cdot R_{Lin} \cdot r_0(\lambda) d\lambda + \begin{pmatrix} \overline{x}(\lambda) \\ \overline{y}(\lambda) \\ \overline{z}(\lambda) \end{pmatrix} \cdot G_{Lin} \cdot g_0(\lambda) d\lambda + \begin{pmatrix} \overline{x}(\lambda) \\ \overline{y}(\lambda) \\ \overline{z}(\lambda) \end{pmatrix} \cdot B_{Lin} \cdot b_0(\lambda) d\lambda \\ &= \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{black} + R_{Lin} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{Red} + G_{Lin} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{Green} + B_{Lin} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{Blue} \end{aligned} \tag{3}$$

$$\begin{pmatrix} R_{Lin} \\ G_{Lin} \\ B_{Lin} \end{pmatrix} = \begin{pmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{pmatrix}^{-1} \cdot \left(\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{Disp} - \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{black} \right) \tag{4}$$

Finally, the nonlinear RGB control signals can be computed by applying the inverse gamma curves. In conclusion, to fully describe the display model the primary valences, the gamma curves, and the black signal have to be acquired.

3.2 Framework 1: Verification of the display model

To verify how well a display respects the assumptions made in the display model, the framework depicted in Figure 2 is used. It is divided into two paths: the measurement and the simulation path. In the measurement path, RGB test colors are displayed and the resulting emission spectra are measured with an X-Rite i1 Pro 2. The measured spectra are converted into the CIE XYZ colorspace, and, in turn, the XYZ values are converted into the CIE Lab space.

In the simulation path, the display model simulates the emission spectra resulting from the RGB test colors. The simulated emission spectra are converted into the CIE Lab colorspace in the same way as described before. The display model is validated by conducting a physical and colorimetric comparison. In the physical comparison, the measured and simulated spectra are compared, while in the colorimetric comparison the $\Delta E2000$ (dE00) and Δab values are computed. The DeltaE metric describes the perceptional difference between two colors. Its original formulation is defined by the euclidean distance of two color stimuli in the CIE Lab colorspace, which is approximately perceptually linear. To reduce the error caused by the non-linearity of the Lab colorspace, the dE00 was introduced. The dE00s are calculated as described in [CIE01], where the weighting factors KL, KC, and KH are set to the default value 1. The dE00 values are classified by the rating scale given in the book Color Imaging [RKAJ08, p.461]. The rating scale is actually defined for the CIE DeltaE 1994 metric, but it can also be used for the dE00 metric since they are very similar. In the Color Imaging book, a DeltaE of 1 is considered as the just noticeable color difference (JND); a value of 2 leads to discernible color differences for patches that are next to each other; and color differences larger than 5 are easily perceived in a side-by-side image comparison.

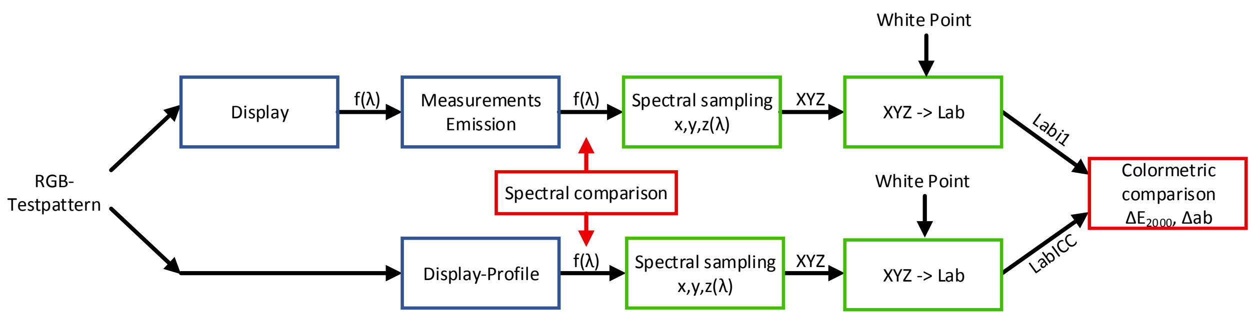

Figure 2: The framework paths: measurement (upper) and simulation (lower).

3.3 Framework 2: Verification of the display model application

The previously described framework used the display model to simulate the color reproduction of the display. To verify the real application of the display model, the framework depicted in Figure 3 is used. It determines the perceived color differences between measured colors and its reproduction on a display. Therefore, the reflectance spectra of a Color Checker are measured, multiplied with an illumination spectrum and converted to the absolute CIE XYZ values (XYZ_{Ref}). Then the XYZ_{Ref} values are normalized with the display black and white point as described in [Ado05]. Applying the display model backward on the normalized XYZ_{Ref} values leads to the display control values (RGB_{Disp}), which are displayed and the resulting emission spectra are captured ($f_{Display}(\lambda)$). These emission spectra are converted to CIE XYZ (XYZ_{Disp}) and normalized in the same way as the initial spectra. Both, the normalized XYZ_{Ref} and $XYZ_{Display}$ values are converted with a given white point to CIE Lab and the dE00s between them are calculated. The resulting dE00s provide information on how well the display model in real application works.

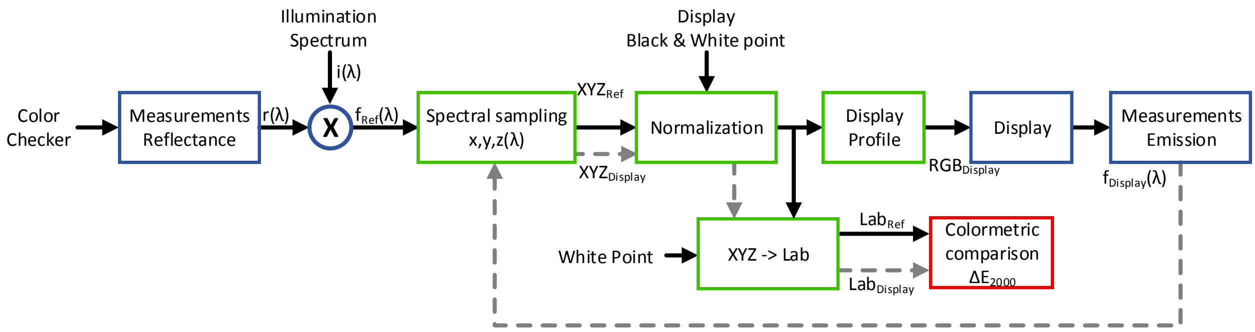

Figure 3: Verification of the reproduction of measured colors on a display using a display profile.

4 Calibration of HTC Vive Pro and Pimax 5k+

In this section, we describe how we implemented the previously described display model and frameworks with state-of-the-art consumer VR HMDs: HTC Vive Pro and Pimax 5k+. First, the display model for each device and each display is determined. Second, the framework as described in Figure 2 is conducted to verify how good the displays respect the display model assumptions. And finally, the real application of the display model is verified by conducting the framework in Figure 3.

4.1 Determination of the display model

Each display has different characteristics, thus, the display models of the left and the right display of the Vive Pro and Pimax have to be acquired separately. To determine the display model the three gamma curves, the primary valences, and the black and white level have to be acquired. Therefore, each display is controlled by 10 RGB signals per color channel, giving the non-linear gamma curves and the primary valences, and by the maximum (255,255,255) and minimum (0,0,0) RGB signals, giving the black and white level. The resulting emission spectra are measured with an X-Rite i1 Pro 2.

The display model defines the colorspace and consequently the quality of the display from a color perspective. A comparison of the display quality of the Vive Pro and Pimax is

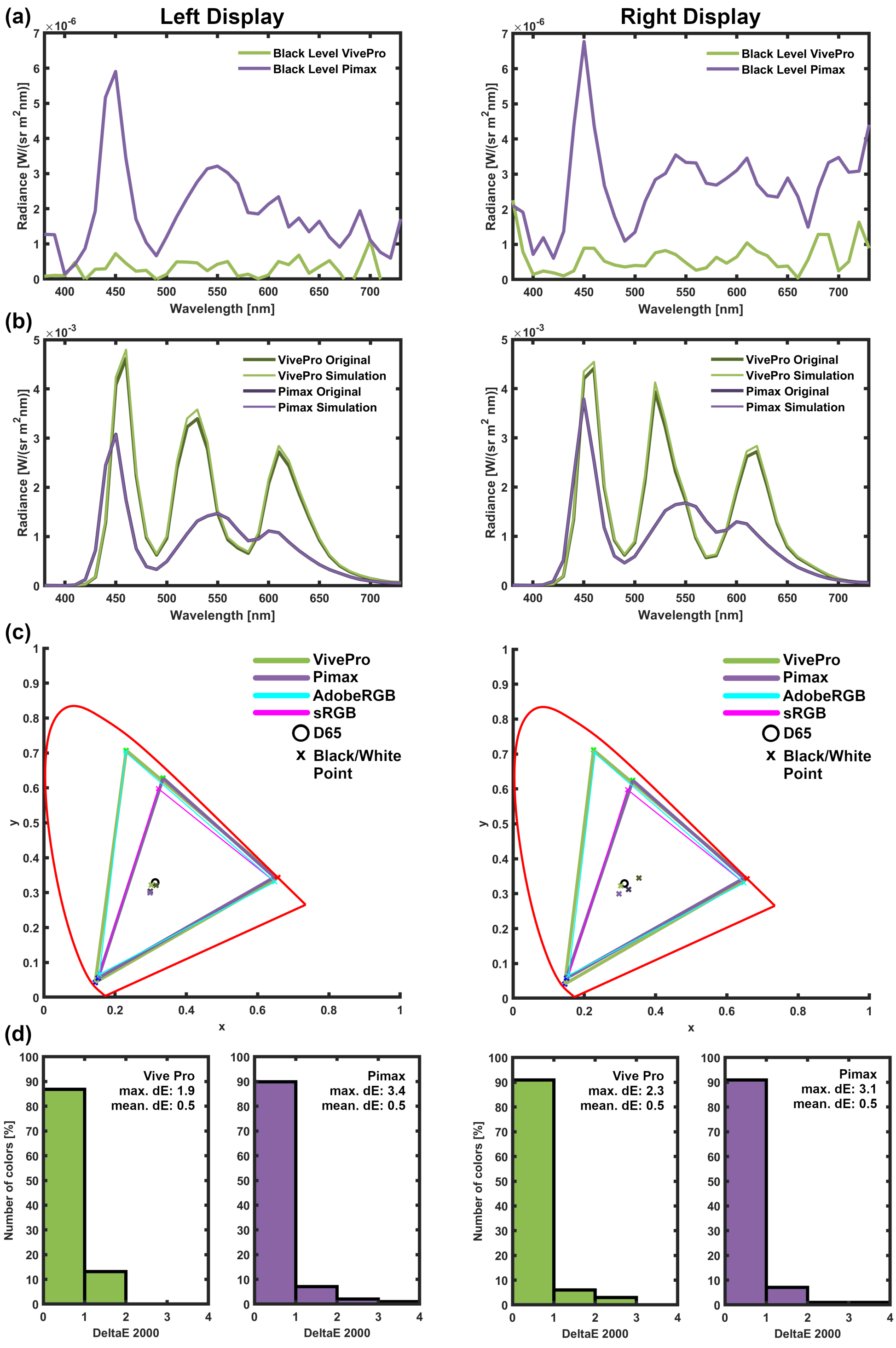

Figure 4: Display characterization of the HTC Vive Pro and Pimax 5k+. The left and right column show the results for the left and right displays, respectively. (a) Black level. (b) Original white spectrum and the simulation with the display-model. (c) Chromaticity diagram including the sRGB, AdobeRGB, and display colorspaces. (d) DeltaE histogram (see Figure 2).

shown in Figure 4, where the Diagrams (a)-(c) characterize the display colorspaces and (d) verifies the used display models.

Fig. 4(a) shows that the Vive Pro has a much lower black level than the Pimax. This can be traced back to the different display technologies. The AMOLED display used in the Vive Pro can entirely turn off each pixel separately, whereas the backlight of the Pimax can not be fully blocked by the LCD panel. This results in a lower and consequently better black level of $0.04\frac{cd}{m^2}$ for the Vive Pro compared to $0.21\frac{cd}{m^2}$ for the Pimax.

In Fig. 4(b) the measured (original) and simulated white spectra are shown. For the simulation, the display model as described in Figure 1 is used. The Vive Pro has a brighter white spectrum than the Pimax, which is confirmed by the maximum luminance of $130\frac{cd}{m^2}2$ compared to $80\frac{cd}{m^2}$. However, the Pimax shows a better match between the original and simulated white spectrum, which indicates that the Vive Pro does not entirely meet the assumptions made in the display model.

The chromaticity diagrams (Fig. 4c) depict the primary valences of the display, sRGB, and AdobeRGB colorspaces. The primary valences of the Vive Pro and Pimax are in great accordance with the AdobeRGB and sRGB colorspace, respectively.

Summarizing, the Vive Pro almost matches the AdobeRGB colorspace. Only the maximum luminance of $130\frac{cd}{m^2}$ is below the $160\frac{cd}{m^2}$ value of the AdobeRGB colorspace. The colorspace of the Pimax covers only the sRGB colorspace and, consequently, is much smaller than the Vive Pro's.

4.2 Verification of the display model

To verify the display model the framework of Figure 2 is conducted, where as RGB-test pattern equally sampled RGB values with a step size of 85 are used. The resulting dE00s between the measured and simulated test colors are shown in the histograms of Fig. 4(d). Both, the Vive Pro and the Pimax, have an average dE00 of 0.5, which is, referring to Section 3, below the JND. However, approximately 10% of the test colors lead to dE00s larger than 1, but only a few of them are larger than 2. Hence, most of the color differences are not discernible when compared side-by-side. The verification of the display model shows that both VR HMDs are in great accordance with the assumptions made in the display model, therefore, the presented display model can be used to describe the displays.

Table 1 gives an overview of the display characterization of the VivePro and Pimax. We conclude that, from a color perspective, the VivePro is more suitable for virtual prototyping than the Pimax due to its larger colorspace and lower maximum deltaEs.

4.3 Verification of the display model application

To verify the real application of the display model the framework from Figure 3 is conducted. We only present the results of the left display of the Vive Pro, since, as previously shown, the Vive Pro is more suitable for virtual prototyping and the left and right displays have very similar characteristics.

As input, we used the X-rite ColorChecker Digital SG with 96 patches and chose the illumination spectrum in a way that the spectrum of the illuminated white patch is equal to

	HTC Vive Pro	Pimax 5k+
Display technology	AMOLED	LCD
FOV	110°	200°
Resolution	1440 x 1600	2560 x 1440
Colorspace	$\sim$ AdobeRGB	$\sim$ sRGB
Max. Luminance	130 cd/m^2	80 cd/m^2
Min. Luminance	0.04 cd/m^2	0.21 cd/m^2
Contrast ratio	3250:1	377:1
Max. dE	2.3	3.4
Mean dE	0.5	0.5

Table 1: Results of our display characterization

the display white spectrum. This guarantees that all ColorChecker spectra are within the display's colorspace.

Usually, VR HMDs are not calibrated and the HMDs are controlled directly by sRGB or AdobeRGB values. To demonstrate the differences between a workflow with and without color management, we conducted the framework with the sRGB, AdobeRGB and measured display profiles.

The results of the three approaches are shown in Figure 5, where the first row shows the color comparison between the reference colors (outer square) and the displayed colors (inner square), and the second row shows the corresponding deltaE histograms. It demonstrates well that the standard workflow with the sRGB profile leads to a clearly visible color difference between the reference and displayed color for all patches. Although the Vive Pro and the AdobeRGB colorspaces are similar, the AdobeRGB profile insufficiently reproduces the reference colors achieving an average dE00 of 5.7. On the other hand, our approach leads to a very convincing color reproduction with an average dE00 of 1. This again confirms that our display model works and that the Vive Pro in combination with our color management can be used for color-critical applications, such as virtual prototyping.

5 Conclusion & Future Work

We presented a simple display model, which can be used to characterize the displays of current consumer VR HMDs. We concluded that both, the HTC Vive Pro and Pimax 5k+, match well the assumptions made in the display model. However, we recommend for color-critical applications the Vive Pro, due to its considerably larger colorspace. The Vive Pro almost covers the AdobeRGB colorspace, while the Pimax only covers the sRGB colorspace. Besides, we demonstrated that with our calibration approach measured colors can be on average reproduced with a dE00 of 1, which is below the perception threshold. The commonly used sRGB and AdobeRGB profiles without color management lead to average

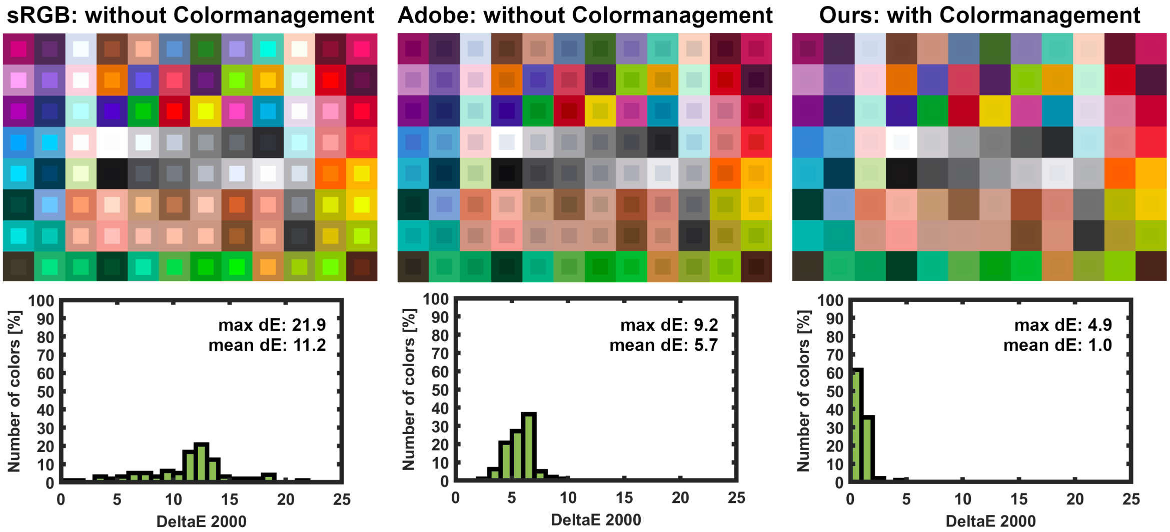

Figure 5: Comparison of measured colors with its reproduction with the sRGB, the AdobeRGB, and the measured display profile on the Vive Pro.

dE00s of 11.2 and 5.7, which is insufficient for color-critical applications of VR HMDs. Furthermore, our display model can be easily incorporated into a real-time rendering system with only marginal computational overhead.

In our display model we do not consider the Fresnel lenses and consequently, omit the lens aberrations. For future work, it would be interesting to extend the display model by a detailed characterization of the Fresnel lenses. Some research in this direction has been proposed by Maxwell et al. [MO18]. Furthermore, our display model assumes that the brightness and color reproduction over the whole display is homogeneous. However, especially OLED-displays do not fulfill this requirement. To use our display model it would be important to quantify the color consistency between different pixels and, if necessary, correct it. We also intend to further evaluate the use of VR HMDs for virtual prototyping. Therefore, we plan an interactive comparison setup based on the data published by Clausen et al. [CMF18], where a defined scene can be observed in reality and virtual reality using a VR HMD.

6 Acknowledgments

We would like to thank the anonymous reviewers for their very detailed and constructive reviews. This work was partially funded by the German Federal Ministry of Education and Research (BMBF) under the grant number 02K16C232 as part of the project “Retail 4.0”.

References

[Ado05] Adobe Systems Incorporated. Adobe®RGB(1998) Color Image Encoding, May 2005.

[CIE01] CIE. Improvement to industrial colour difference Evaluation. Technical report, CIE 14x-2001, Commission internationale de l'éclairage, 2001.

[CMF18] Olaf Clausen, Ricardo Marroquim, and Arnulph Fuhrmann. Acquisition and validation of spectral ground truth data for predictive rendering of rough surfaces. *Computer Graphics Forum*, 37(4):1–12, 2018.

[GBR+08] Davide Gadia, Cristian Bonanomi, Maurizio Rossi, Alessandro Rizzi, and Daniele Marini. Color management and color perception issues in a virtual reality theater. In *Eletronic Imaging*, volume 6803, 2008.

[GdSZ98] Antonino Gomes de Sa and Gabriel Zachmann. Integrating virtual reality for virtual prototyping. In *Proceedings of DETC'98, 1998 ASME Design Engineering Technical Conferences*, Atlanta, Georgia, USA, 1998.

[Gre99] Donald P. Greenberg. A framework for realistic image synthesis. *Commun. ACM*, 42(8):44–53, August 1999.

[GSA01] Enrico Gobetti, Riccardo Scateni, and Marco Agus. Exploring virtual prototypes using time-critical multiresolution rendering. *ERCIM News No*, 2001.

[Inf] Infitec. Color management in 3d visualization. `https://infitec.net/color-management-in-3d-visualization/`. Accessed: 2019-05-03.

[JJB10] Gerhardt Jérémie and Thomas Jean-Baptiste. Toward an automatic color calibration for 3d displays. In *Color and Imaging Conference*, volume 18, pages 5–10. Society for Imaging Science and Technology, 2010.

[JP07] Antonio Jimeno and Alberto Puerta. State of the art of the virtual reality applied to design and manufacturing processes. *The International Journal of Advanced Manufacturing Technology*, 33(9-10):866–874, 2007.

[JSF08] Helmut Jorke, Arnold Simon, and Markus Fritz. Advanced stereo projection using interference filters. In *2008 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pages 177–180, May 2008.

[KRK03] Wolfram Kresse, Dirk Reiners, and Christian Knöpfle. Color consistency for digital multi-projector stereo display systems: The heyewall and the digital cave. In *Proceedings of the Workshop on Virtual Environments 2003*, EGVE '03, pages 271–279, New York, NY, USA, 2003. ACM.

[MO18] Douglas Maxwell and Evan Oster. Evaluating the applicability of repurposed entertainment virtual reality devices for military training. In *MODSIM World*, 2018.

[RKAJ08] Erik Reinhard, Erum Arif Khan, Ahmet Oguz Akyz, and Garrett M. Johnson. *Color Imaging: Fundamentals and Applications*. A. K. Peters, Ltd., Natick, MA, USA, 2008.

[SPL+16] Volker Settgast, Johanna Pirker, Stefan Lontschar, Stefan Maggale, and Christian Gütl. Evaluating experiences in different virtual reality setups. In *International Conference on Entertainment Computing*, pages 115–125. Springer, 2016.