Dark Stereo: Improving Depth Perception Under Low Luminance

KRZYSZTOF WOLSKI, MPI Informatik, Germany
FANGCHENG ZHONG, University of Cambridge, UK
KAROL MYSZKOWSKI, MPI Informatik, Germany
RAFAŁ K. MANTIUK, University of Cambridge, UK

Standard rendering

Proposed stereo constancy method

Fig. 1. Crossed fusion stereoscopic pairs demonstrating the proposed stereo constancy algorithm. Left: original scene. Right: after stereo constancy processing. The images should be seen on a dimmed display (about 5 cd/m^2 peak brightness). To dim the image, either see it through a neutral density filter (1.0 D), strong sunglasses or use the minimum screen brightness on a smartphone phone and observe it in a dark room. Enlarge the images to the point that both pairs stretch the width of the screen. The assets are a part of the POLYGON series prepared by Synty Store.

It is often desirable or unavoidable to display Virtual Reality (VR) or stereoscopic content at low brightness. For example, a dimmer display reduces the flicker artefacts that are introduced by low-persistence VR headsets. It also saves power, prolongs battery life, and reduces the cost of a display or projection system. Additionally, stereo movies are usually displayed at relatively low luminance due to polarization filters or other optical elements necessary to separate two views. However, the binocular depth cues become less reliable at low luminance. In this paper, we propose a model of stereo constancy that predicts the precision of binocular depth cues for a given contrast and luminance. We use the model to design a novel contrast enhancement algorithm that compensates for the deteriorated depth perception to deliver good-quality stereoscopic images even for displays of very low brightness.

CCS Concepts: • **Computing methodologies** → **Perception**; *Virtual reality*.

Additional Key Words and Phrases: depth perception, binocular disparity, stereo constancy, contrast constancy, low luminance, mesopic vision, screen-space postprocessing

Authors' addresses: Krzysztof Wolski, kwolski@mpi-inf.mpg.de, MPI Informatik, Campus E1 4, Stuhlsatzenhausweg 4, Saarbruecken, Germany, 66123; Fangcheng Zhong, fangcheng.zhong@cst.cam.ac.uk, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, Cambridge, UK, CB3 0FD; Karol Myszkowski, karol@mpi-inf.mpg.de, MPI Informatik, Campus E1 4, Stuhlsatzenhausweg 4, Saarbruecken, Germany, 66123; Rafał K. Mantiuk, rafal.mantiuk@cl.cam.ac.uk, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, Cambridge, UK, CB3 0FD.

0730-0301/2022/7-ART146
https://doi.org/10.1145/3528223.3530136

ACM Reference Format:
Krzysztof Wolski, Fangcheng Zhong, Karol Myszkowski, and Rafał K. Mantiuk. 2022. Dark Stereo: Improving Depth Perception Under Low Luminance. *ACM Trans. Graph.* 41, 4, Article 146 (July 2022), 12 pages. https://doi.org/10.1145/3528223.3530136

1 INTRODUCTION

There are a number of benefits of using darker displays for VR headsets. An obvious benefit is reduced power consumption, as the display itself can be responsible for half of the power usage in a stand-alone, battery-powered headset. It has been also postulated that a lower display brightness in VR/AR can reduce VR sickness, however, no evidence for that has been found so far [Vasylevska et al. 2019]. A less obvious benefit is an improved quality of the reproduced motion. Most VR displays are equipped with low persistence, in which an image is displayed at a higher intensity for a fraction of frame duration and the display remains blank for the rest of the frame. Without the stroboscopic effect introduced by such low persistence displays, we would see a significant amount of blurring caused by eye gaze moving over the image, which is stationary on the display over the duration of a frame. However, low persistence can also introduce visible flicker if the refresh rate is not high enough [Hoffman and Lee 2019]. Lowering display luminance can reduce the visibility of such flickering without the need to increase the refresh rate [Chapiro et al. 2019].

However, dimming a display also introduces a number of undesirable effects: it reduces the perceived contrast of an image [Barbur and Stockman 2010; Wanat and Mantiuk 2014], makes the image appear less colorful [Cao et al. 2008; Kwak et al. 2003; Shin et al. 2004], and affects our ability to make depth judgements based on

stereoscopic depth cues. While the former two effects have been well studied and addressed in the literature, the effect of absolute luminance on our ability to see depth has received relatively less attention.

The main goals of this work are: (a) to measure and quantify the effect of display luminance and contrast on our ability to make depth judgements from binocular disparity depth cues, and (b) to propose an image contrast enhancement technique that can enhance depth perception on dimmed stereoscopic displays. Based on a series of psychophysical measurements on our prototype stereoscopic high dynamic range (HDR) display, we propose a model of *stereoscopic constancy* (Sections 3 and 4), which predicts the amount of physical contrast needed to maintain the same precision of binocular disparity depth cues across the luminance range of 0.1 cd/m^2 to 1000 cd/m^2. The model is then used to develop a multi-scale contrast compensation method (Section 5) that attempts to preserve the precision of binocular depth cues at different display luminance levels. The method has been implemented in GPU shaders and it can be used in real-time applications. Finally, we test our algorithm in a low-brightness VR rendering application, in which our method is both preferred and gives a better impression of depth than non-processed rendering and existing methods (Section 6). The main contributions of our work are:

- Psychophysical measurements of the effect of luminance on our ability to infer 3D shapes from binocular disparity cues.
- A model of stereoscopic constancy.
- A contrast enhancement algorithm that improves depth perception on dimmed displays, with applications to VR and 3D movies.
- An implementation of the algorithm as an efficient post-processing GPU shader, which can be used directly in the Unity Engine[1].

2 RELATED WORK

In this section, we first consider works that discuss the advantages and disadvantages of dimming a display. Next, we discuss studies on the functioning of the human visual system (HVS) and color appearance in dark and bright conditions. Finally, we outline the works that manipulate image content to improve stereo vision.

Effect of display dimming on user experience. Several works have studied the impact of display brightness on user experience and power consumption. Schuchhardt et al. [2015] proposed an optimal dimming scheme to reduce mobile display brightness while ensuring good legibility on the screen. Erickson et al. [2020] investigated the effect of color mode on visual acuity and fatigue with VR head-mounted displays. They found that a dark background used in dark mode can reduce visual fatigue and increase visual acuity in a dim VR environment. Mantiuk et al. [2009] argued for using amber and red colors on dark displays as they induce the least amount of disability glare or photophobia. They also found that the preferred display brightness was between 20 and 40 cd/m^2 in a dark environment. Chapiro et al. [2019] reported that, in low luminance conditions, judder is less visible, leading to better-perceived motion quality. These works provide solid ground for the merit of dimming VR displays. However, it is well recognized that the visual performance, including contrast and depth perception, is substantially degraded at low luminance. Although the visual system can preserve the appearance of contrast through a range of conditions [Georgeson and Sullivan 1975], the contrast appears weaker and eventually disappears as the luminance is reduced, particularly the contrast that is close to the threshold [Kulikowski 1976; Peli et al. 1991]. Lower luminance levels cause the pupil to dilate. This could result in a larger defocus blur in fixed-focus displays. Singh et al. [2018] found that matching the brightness of the displayed and real object on an AR display results in more accurate depth estimation when focus cues are consistent (no vergence-accommodation conflict) or when focused on the mid-point of the tested depth range. No absolute luminance levels were reported so we cannot compare their finding to ours. According to Frisby et al. [1978], the global stereopsis mechanism requires a certain level of suprathreshold contrast to detect a binocular disparity signal. Since the contrast detection thresholds are much higher at low luminance, our ability to see depth in low contrast content is greatly reduced [Livingstone and Hubel 1994]. Our work focuses on solving this issue.

Color appearance on dimmed displays. Despite the ability of our visual system to maintain color perception across a very wide range of illumination (color constancy), some changes in appearance are inevitable. Kim et al. [2009] observed, that at higher luminance levels perceived colorfulness and lightness increases. The opposite effect can be observed, when light levels are low, in particular when the visual system transitions from cone-mediated vision (photopic) to cone- and rod-mediated vision (mesopic) [Barbur and Stockman 2010]. Indeed, color appearance in mesopic vision (0.01 - 3 cd/m^2) can be influenced by the change in rod activity [Stabell and Stabell 1998]. As a consequence, luminance levels alter the perception attributes such as hue, chroma, and lightness [Kwak et al. 2003; Shin et al. 2004]. Brightness and colourfulness also reduce with decreasing luminance [Fu et al. 2012]. Several models explaining the changes in colour appearance at mesopic light levels have been proposed [Cao et al. 2008; Shin et al. 2004], including an extension of CIECAM02 color appearance model [Luo and Li 2013].

Impact of color on depth perception. Multiple studies tried to investigate impact of colored stimuli on perceived depth. Trościanko et al. [1991] found out that certain isoluminant color gradients (the ones occurring in nature) might be used as a monocular depth cue, but its input to stereopsis is weak. The follow-up work [Bailey et al. 2007] established that influence of colors on depth judgements is even weaker when natural images are used as a stimuli. Overall, chromatic (isoluminant) contrast has little influence on the disparity depth cues [Simmons and Kingdom 1994]. For that reason, the vast majority of binocular vision studies rely on achromatic stimuli (stereoacuity thresholds are ten times larger for chromatic stimuli than for achromatic ones [Krauskopf and Forte 2002]). Because chromatic contrast has little influence on the depth perception, our considerations are limited to luminance.

[1]The code of model, postprocessing algorithm, and the experimental data can be found at the project web page https://dark-stereo.mpi-inf.mpg.de and at https://github.com/gfxdisp/dark-stereo.

Simulation and compensation of night vision. Some works tried to simulate and compensate for changes between day and night visions. Wanat et al. [2014] proposed a luminance re-targeting method to match the appearance of different luminance levels by altering perceived contrast and modelling of hue and saturation shifts of an image. Kellnhofer et al. [2014] argued that for stereoscopic displays, such changes are not sufficient to fully simulate dark conditions. Their proposed solution involves the manipulation of binocular disparity so that a scotopic stereo content displayed on a photopic monitor is perceived as the scene was scotopic. Contrary to the mentioned studies, instead of improving or simulating the appearance of a dark screen, our work focuses on improving stereo vision in low luminance conditions.

Disparity manipulation. Several works have proposed techniques for altering image disparity, mostly intending to reduce vergence-accommodation conflict which can make images uncomfortable to view [Wann et al. 1995]. Oskam et al. [2011] described a method that controls the camera convergence and interaxial separation over time to optimally map a dynamically changing scene to the desired depth range, which improves comfort. Lang et al. [2010] proposed a method that controls and re-targets the depth of a stereoscopic scene in a nonlinear and locally adaptive fashion. The solution employs computed disparity and saliency estimates to compute a deformation of the input views so that they meet the desired disparities. To avoid undesirable distortions from disparity manipulation, Didyk et al. [2011] introduced a perceptual model of disparity which provides a metric to evaluate perceived disparity change for stereo images. The follow-up work [Didyk et al. 2012b] studies the interplay of contrast and disparity on the depth discrimination. Based on their disparity-perception model, they jointly manipulate luminance contrast and disparity to reduce depth in stereoscopic images. However, these models ignore the impact of display luminance, which is central to our work. Didyk et al. [2012a] also proposed a depth-enhancement technique that relies on Cornsweet illusion in the disparity domain. While all these methods show the potential of manipulating disparity to improve the perception of depth in a displayed image, we argue that manipulating disparity in VR content might affect visual feedback to egomotion and contribute to an intensified VR sickness [Jacobs et al. 2019].

Depth enhancement by image content manipulation. It has been shown that certain image manipulations can enhance the apparent depth. Luft et al. [2006] proposed a technique that enhances contrast and color near depth discontinuities to improve the perceptual quality of monoscopic images. Our intention is to preserve the perception of depth in stereoscopic content across different luminance levels rather than to enhance it.

3 EXPERIMENT 1: 3D SHAPE PERCEPTION

Binocular disparity is one of the most important depth cues [Cutting and Vishton 1995] that is commonly employed in VR and cinematographic applications to trigger stereo 3D scene appearance. In this work, we develop a computational model of the precision of binocular disparity cues as a function of image contrast and luminance. To this end, we designed an experiment in which observers were asked to judge the angle of a 3D hinge-like shape (Figure 2, top), reproduced on the display using only disparity depth cues (Figure 2, bottom). Our 3D shape perception experiment was inspired by the study of Watt et al. [2005], where a similar hinge-like shape was used to examine whether focus cues have an indirect effect on depth interpretation.

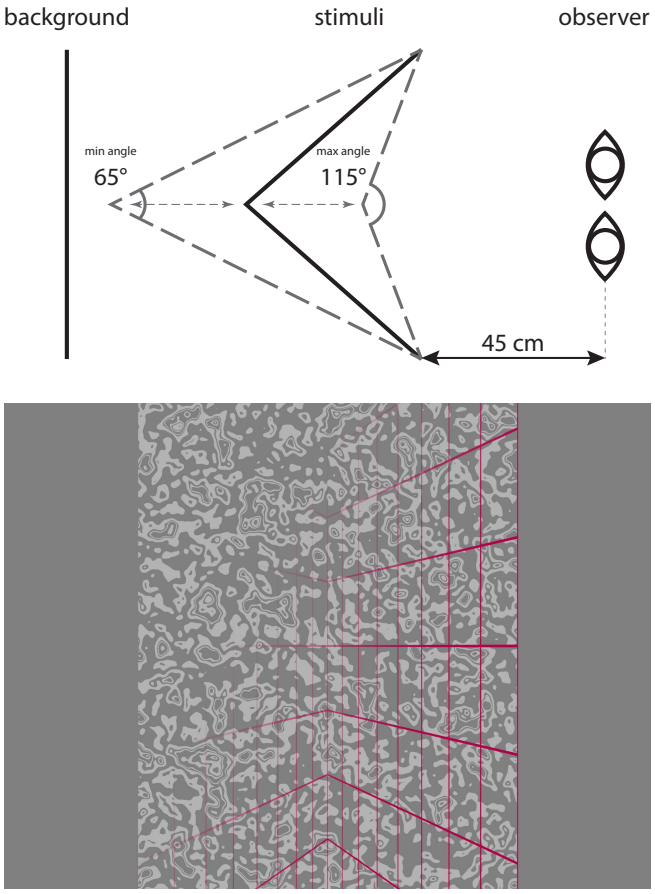

Fig. 2. Top: The stimulus used in Experiment 1. The observer is presented with a hinge-like concave shape. The angle is changed by moving the hinge part towards or away from the observer (depicted by the arrows). Bottom: Procedural organic pattern on a uniform background. The superimposed grid depicts the three-dimensional shape of the stimuli. Note that the superimposed grid was only added to this figure to facilitate its 3D interpretation, while originally the hinge shape was reproduced only by the disparity cue.

Apparatus. The experiment was conducted on a custom-built stereo high dynamic range (HDR) display, which allowed a single observer to view a pair of stereo images through an optical arrangement similar to the Wheatstone mirror stereoscope, as illustrated in Figure 3. We used a stereoscopic HDR display to explore a much larger range of luminance than that available on regular displays. The stereoscopic stimuli were shown by two projector-based 9.7" HDR displays, 2048×1536 resolution each. The peak luminance of each display was above 3,000 cd/m^2 and the black level

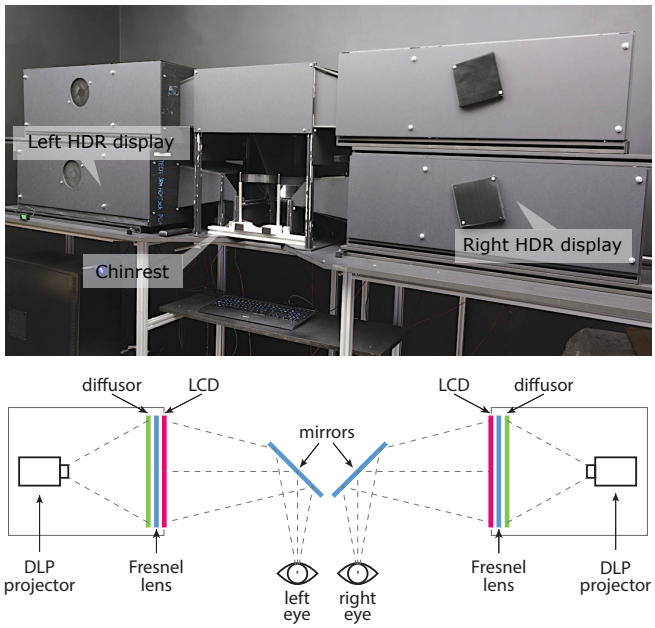

Fig. 3. Top: The photograph of a prototype stereo HDR display. Each observer sat at the display with their head stabilized by both the chin-rest and the head-rest. Note that the upper part of the display, intended for multi-focal-plane presentation, was not used in this project. Bottom: The schematic diagram of the display, showing the portion of the stereo display.

was below 0.01 cd/m^2. Further details on the display design, control software and its colorimetric and geometric calibration can be found in [Zhong et al. 2021]. The display algorithm used spatio-temporal dithering to avoid banding artifacts caused by insufficient bit-depth. Additional information on HDR display calibration for psychophysical measurements, which were conducted with similar but monoscopic display, can be found in [Wuerger et al. 2020]. We used HDR rather than a standard display as it allowed us to reproduce both very low and very high luminance while maintaining sufficient contrast and bit-depth accuracy. The virtual images of the HDR content were placed 45 cm in front of the observers, with a resolution of 82 pixels per visual degree.

Stimuli. To study the influence of low luminance on binocular depth perception, we designed a stimulus that contained a controlled binocular disparity cue while minimizing the effect of other depth cues. The observers were presented with a concave, hinge-like shape on a uniform gray background. The stimuli were textured with a procedurally generated pattern (see Figure 2, bottom) with only 2 shades of gray. To isolate only the disparity cue, the texture was projected on a surface from a position of a cyclopean eye, thus eliminating the perspective projection cue (the texture density did not change with the distance). It was also rendered without a reflection model to remove shading cues. During the experiment, observers were asked to use a chinrest to prevent head movements. The stimuli and setup are presented in Figure 2.

It is well known that contrast sensitivity strongly changes in the scotopic and mesopic range, hence we decided to investigate the effect for luminances of 0.1, 1 and 10 cd/m^2. However, recent work by Wuerger et al. [2020] has shown that there are also significant changes in contrast sensitivity of above 100 cd/m^2 which motivated is to test two additional luminance levels of 100 and 1000 cd/m^2. For each luminance, 4 contrast levels of the texture were measured: 0.05, 0.1, 0.2, or 0.4. The Weber contrast of the texture is expressed as:

$$C_{\mathrm{w}} = \frac{Y_{\mathrm{max}} - Y_{\mathrm{min}}}{Y_{\mathrm{min}}}, \tag{1}$$

where Y_{max} and Y_{min} are maximum and minimum luminance of the stimuli. In the case of our stimuli, Y_{min} was the luminance of the background.

Because in the pilot experiment observers were not able to see the stimuli at 0.1 cd/m^2 and Weber contrast of 0.05, we decided to remove this condition from the main experiment. The order of conditions was randomized for each observer. If the luminance decreased between two conditions, we displayed a uniform field with the target luminance for a minute to ensure the observer was adapted to the new luminance level.

Experimental procedure. The task was to assess whether the angle is greater or smaller than 90 degrees and confirm the decision by pressing a corresponding key on the keyboard. For each of the 19 conditions, we used the method of constant stimuli to estimate the probability of judging the angle as acute or obtuse. The tested angles were: 65°, 75°, 85°, 95°, 105°, and 115°. Six trials were collected for each angle and each observer. We changed the angle by moving the hinge part towards and away from the observer while keeping the side edges stationary (see Figure 2). This was done to avoid additional depth cues. The tested angles were randomized between the trials.

Each observer was asked to complete a training session at 10 cd/m^2, in which they were given feedback on whether their answer was correct. Such feedback was not given in the main experiment. The training session helped the observers to familiarize themselves with the task, get accustomed to disparity-only stimuli and was also used to screen observers. We excluded three observers who were unable to do the task in the training session from further experiments. The entire experiment took each participant around two hours and was split into 3–5 short sessions.

Observers. Eleven volunteers, who passed the training session, (four females and seven males, aged 24 to 36, including three authors) participated in the experiment. All were recruited from among students and researcher working in the field of computer graphics. Nine of them completed trials for all luminance levels, while two completed only the trials for luminance levels from 0.1 cd/m^2 to 10 cd/m^2. All observers had a normal or corrected-to-normal visual acuity (self-reported). All passed the Titmus stereoacuity test. All observers except the authors were naïve to the purpose of the experiment. Before the experiment, each observer read and signed the consent form. The observers were rewarded for their participation. The experiment was approved by the departmental ethics board.

Results. The experiment explained how well the observers could see a geometric angle at several luminance and contrast levels. The

data averaged over all observers, plotted as the probability that an observer reports an obtuse angle, is shown in Figure 4 as red stars.

The first important observation is that the psychometric curves formed by the data points cross 50% probability point at about 90-degree angle regardless of luminance and contrast. This means that the perceived angles were not distorted by lower luminance and contrast. However, the slopes of the psychometric functions differ substantially between the conditions. The shallower slopes indicate that the observers more often mistook the angle at low luminance and low contrast. This means that lower luminance does not reduce the accuracy of the shape assessment task, but it reduces the precision of that task.

4 STEREO CONSTANCY MODEL

In this section, we propose a model that can predict how contrast needs to be altered to preserve the same precision of the stereo task across different luminance levels. As the first step, we assume the collected data can be explained by a psychometric function that follows the Weibull cumulative distribution function [Wichmann and Hill 2001]. We used this psychometric function to describe the probability p of an observer perceiving the hinge-like shape (Figure 2) as an obtuse angle given the actual angle α, represented in degrees:

$$p(\alpha, \beta) = 1 - \exp\left(\log(0.5) 10^{\beta(\alpha - \alpha_{thr})}\right) \tag{2}$$

where α_{thr} is the angle at which the probability of detection p is 0.5, which we assumed to be 90 degrees. The value of β controls the steepness of the function, which reflects the precision of the user performance in this task. A higher value of β means that the observer is more sensitive to the variations in the perceived angle and also that the task is easier.

Our goal is to find a model of β as a function of contrast and luminance, such that the likelihood of the data observed in the 3D shape perception experiment is maximized. We found that β can be explained by a quadratic function of contrast and log-luminance:

$$\beta(c, L; \mathbf{w}) = w_1 L + w_2 c + w_3 L^2 + w_4 c^2 + w_5 \tag{3}$$

where L is the logarithm of luminance ($L = \log_{10}(Y)$), c is logarithmic contrast, and $\mathbf{w} = [w_1, ..., w_5]$ denote unknown free parameters. Note that contrast was recorded as Weber contrast C_w (1) in our experiment. However, our contrast enhancement method (Section 5) can be implemented more efficiently if it operates on logarithmic contrast. The Weber contrast C_w can be converted into logarithmic contrast c with the formula:

$$c = \log_{10}(C_w + 1). \tag{4}$$

To have a better control over free parameters, and to ensure that the function is monotonic, we used the maximum a posteriori (MAP) estimation to find the values of $\mathbf{w}$. We assume $\mathbf{w} \sim N(\boldsymbol{\mu}, \mathrm{diag}(\boldsymbol{\sigma})^2)$ for some $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, where μ_i and σ_i^2 are the mean and the variance of w_i respectively.

The likelihood of observing k out of n trials (of selecting obtuse angle) can be explained by a binomial distribution, with a latent probability of p of perceiving the angle as obtuse. As indicated by (2), the value of p is dependent on the presented angle α and detection sensitivity β, which is then parameterized by contrast c, luminance L, and $\mathbf{w}$ in (3). Under the MAP framework, free parameters $\mathbf{w}$ can be found by minimizing the negated log-likelihood of the binomial distribution:

$$\begin{aligned}\arg\min_{\mathbf{w}} - \sum_s \sum_{\mathbf{d}} \log\left(\binom{n_{s,\mathbf{d}}}{k_{s,\mathbf{d}}} p_{\mathbf{d}}^{k_{s,\mathbf{d}}} \left(1 - p_{\mathbf{d}}^{(n_{s,\mathbf{d}} - k_{s,\mathbf{d}})}\right)\right) \\ + \sum_{i=3,4} \frac{1}{2\sigma_i^2}(w_i - \mu_i)^2\end{aligned} \tag{5}$$

where s is the index of the observer, $\mathbf{d} = [\alpha, c, L]$ are the parameters of each condition, and $p_{\mathbf{d}} = p(\alpha, \beta(c, L; \mathbf{w}))$ is given by Equations 2 and 3. $n_{s,\mathbf{d}}$ in the binomial coefficient is the total number of measurements collected for observer s and condition $\mathbf{d}$ and $k_{s,\mathbf{d}}$ is the number of measurements in which obtuse angle was selected. Table 1 shows the final estimated parameters we found for $\mathbf{w}$, and our choices for $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. Under the MAP framework, $\frac{1}{2\sigma_i^2}$ becomes the weights on the regularization terms. Note that we only regularize 2^{nd}-order terms w_3 and w_4 to ensure monotonicity. With these parameters, we plot the corresponding fitted psychometric functions parameterized by β under various luminance and contrast conditions on top of the original data points in Figure 4 as blue curves. The plots demonstrate that the model explains well most conditions. The worse fit for some conditions (e.g. 0.1 cd/m^2, $C_W = 0.4$) is due to the regularization, which was necessary to make the model monotonic and thus invertible.

Table 1. Estimated values of free parameters of (3) and the priors for the Maximum a Posteriori (MAP) estimation. Symbol "/" means that no prior was used.

	1	2	3	4	5
w	0.0050	0.1849	-0.0010	0.3994	0.0263
μ	/	/	-0.4	0.4	/
σ^2	/	/	0.001	0.001	/

Next, we use the fitted model to find the lines of equal precision of the task (constant β). Such lines are plotted as continuous lines in Figure 5. The figure shows that to maintain the same precision of the task (the same β), we need to increase the contrast at low luminance and that such an increase should be smaller for higher contrast. We will refer to this model as a stereo constancy model and use it in the next section to derive our contrast enhancement technique for dark stereo displays. To demonstrate that the effect cannot be predicted by a contrast constancy model, we plot in the same figure the contrast constancy model of Kulikowski [Kulikowski 1976] (used in [Wanat and Mantiuk 2014]). The comparison shows that stereo constancy requires stronger contrast enhancement between 0.1 and 10 cd/m^2 (mesopic and photopic range) than contrast constancy. The difference between both models will be further corroborated in our validation experiments in Section 6.

5 STEREO-PRESERVING CONTRAST ENHANCEMENT METHOD

We use our model to design a local contrast enhancement method that preserves the precision and difficulty of stereo perception at low

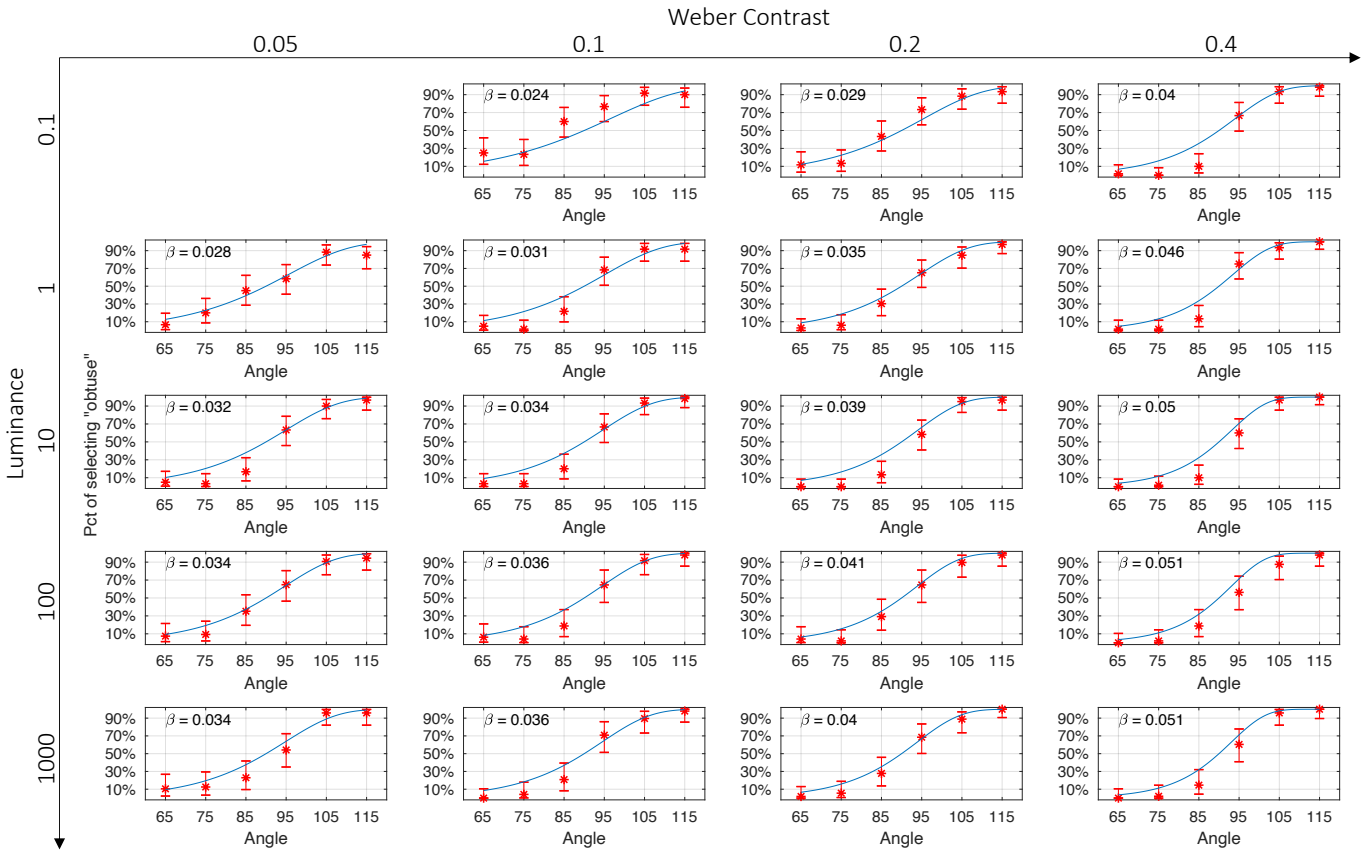

Fig. 4. The red stars are the original data points collected from the 3D shape perception experiment (Section 3). They represent the frequency at which the participants assessed the angle as obtuse under various luminance and contrast conditions. The error bars denote the 99% confidence intervals. The blue curves represent our fitted psychometric model (Section 4). The top-left plot has no data points as it was impossible to see the stimuli at this condition.

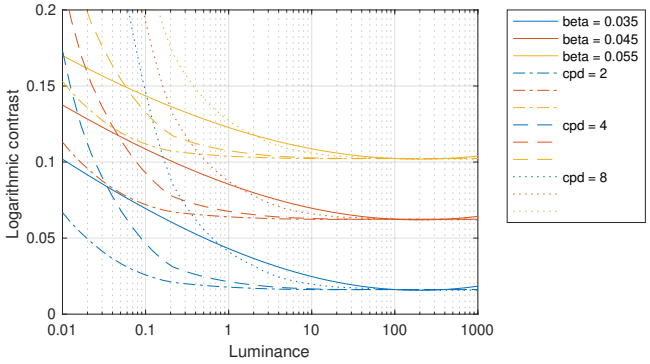

Fig. 5. The solid lines, or equivalent-β lines, connect the contrast values that result in the same precision of perceiving depth (the same β of the psychophysical function) as a function of different display luminance levels. The lines are derived from our model of stereo task difficulty. The dashed lines represent the equivalent perceived contrast for three different spatial frequencies (2, 4, and 8 cpd) according to Kulikowski's model.

luminance. Our method is inspired by the compensation for dark displays proposed by Wanat et al. [2014], yet it differs in its goal as our method is meant to improve stereo perception rather than contrast and color appearance. We employ the stereo constancy model from the previous section rather than Kulikowski's contrast matching model. We also improve on a few processing steps for better real-time performance and temporal stability. The processing diagram of the method is presented in Figure 6. The following subsections present the details of each step of our algorithm.

5.1 Color space transformation

Because our stereo constancy model is defined in terms of physical (linear) luminance units, we need to convert the rendered frame from a gamma-encoded to a linear color space. Assuming ITU-R BT.709-6 RGB primaries and the standard gamma ($\gamma = 2.2$), the relative luminance, y_{input}, is computed as:

$$y_{\text{input}}(\mathbf{x}) = \sum_{k=1}^{3} v_k \, I'^{\gamma}_{input}(\mathbf{x}, k), \tag{6}$$

where $I'(\mathbf{x}, k)$ is the gamma-encoded input value at pixel $\mathbf{x}$ and in color channel k (in the range 0–1), while $v_k = [0.212656, 0.715158,$

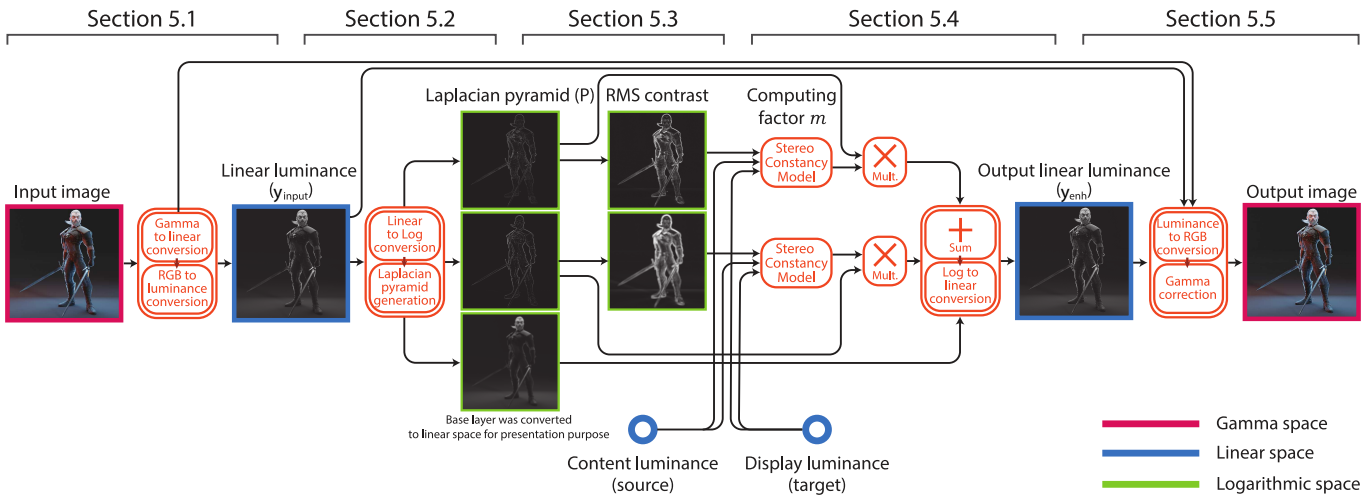

Fig. 6. Schematic diagram of the proposed method. As the first step, the input image is linearized and color channels are used to compute luminance. Then, a Laplacian pyramid is built using the luminance converted to logarithmic space. Next, we compute a magnitude of contrast which is fed to the stereo constancy model. Per-pixel enhancement factor m produced by the model is multiplied with the contrast acquired in the process of Laplacian pyramid creation. Later all the levels of the pyramid are summed together and converted back to linear space. As the last step, the enhanced luminance channel is combined with the color of the input image, and the image is gamma encoded.

0.072186]. Note that we use lower-case y for relative luminance to make it distinct from absolute luminance, Y.

5.2 Multi-scale decomposition

The proposed method compensates for the deteriorated depth perception by enhancing local image contrast. In order to operate on local image contrast, we decompose an image into frequency bands using the Laplacian pyramid. We use the classical Burt and Adelson method [Burt and Adelson 1983] with the coefficient $a = 0.4$ used to construct the filters. In our implementation, we construct a Laplacian pyramid consisting of 3 levels: two band-pass levels and one low-pass level (baseband). The two band-pass levels are sufficient because of the limited effective resolution of VR headsets (in terms of pixels per degree). Such a shallow decomposition also improves the performance in real-time applications.

For computational convenience, the decomposition is performed on logarithmic values of luminance $l = \log_{10}(y_{\text{input}})$. This ensures that the coefficient of the pyramid represents logarithmic contrast (they approximate the logarithm of ratios between two levels). The Laplacian pyramid coefficient at level i is then computed as:

$$P_i(\mathbf{x}) = (g_i * l)(\mathbf{x}) - (g_{i+1} * l)(\mathbf{x}), \tag{7}$$

where g_i is the kernel of a Gaussian pyramid at the level i and $*$ is the convolution operator. The pyramid was not subsampled in our implementation. Instead, we used dilated convolutions.

5.3 Measure of local contrast

Our stereo constancy model requires an estimate of the local contrast to find a corresponding equivalent contrast in the target image. Although we could use the coefficients of the Laplacian pyramid for this purpose, this results in over-enhancement and artifacts at sharp contrast edges, as explained in [Wanat and Mantiuk 2014] and shown in Figure 7. We follow the same approach as in [Wanat and Mantiuk 2014] and compute a root-mean-squared (RMS) measure of local contrast, but we do it more efficiently by reusing the Gaussian pyramid from the multi-scale decomposition step.

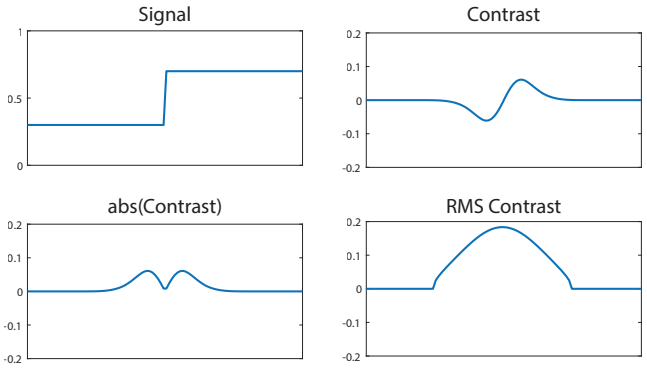

Fig. 7. Comparison of two different local contrast measures. When a sharp-contrast edge (top-left) is decomposed on a Laplacian pyramid, the band-pass levels will contain a signal that resembles a derivative of a Gaussian (top-right). Using the absolute values on that signal (bottom-left) as a measure of contrast has two disadvantages: (a) the measure has zero-crossing at the edge location; and (b) it underestimates the contrast of the edge as it considers the signal in a single band. Using RMS contrast (bottom-right) solves both problems.

The localized root-mean-square (RMS) contrast can be computed as:

$$c_i(\mathbf{x}) = \sqrt{(g_\sigma * l^2)(\mathbf{x}) - ((g_\sigma * l)(\mathbf{x}))^2}, \tag{8}$$

where l is the logarithm of relative luminance and g is a Gaussian kernel with standard deviation σ. We want to use the kernels with

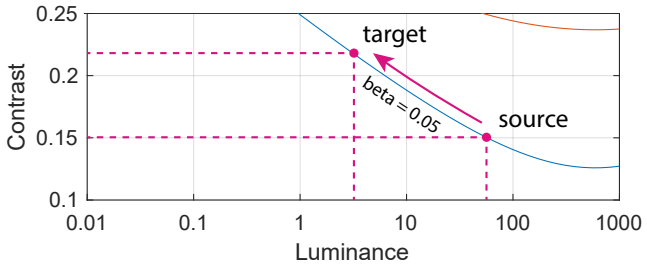

Fig. 8. Method of finding equivalent contrast that preserves the precision of binocular disparity cues. Similar as in Figure 5, for a given input contrast and source luminance, our stereo constancy model gives the curves of equivalent contrast (constant β, blue line). This lets us find the desired contrast for any target display luminance.

larger σ at the lower frequency pyramid levels. To avoid computing additional convolutions, we can instead reuse the Gaussian pyramid and estimate the local RMS contrast as:

$$c_i(\mathbf{x}) = \sqrt{H_i(\mathbf{x}) - G_i^2(\mathbf{x})}\,, \tag{9}$$

where G_i is a Gaussian pyramid built from log-luminance l and H_i is a Gaussian pyramids built from squared log-luminance l^2. Computing a second pyramid H is inexpensive on a GPU, as it can be done by operating on a 2-channel texture, where the first channel contains log-luminance and the second channel contains squared log-luminance.

5.4 Contrast retargeting

Once the Laplacian pyramid and contrast magnitude are computed, we can map the contrast for a given source luminance (Y_{in}) to the contrast that provides the same stereoacuity when seen at target luminance (Y_{out}). This can be done by executing the following steps for every frequency band except the low-pass band (baseband), which does not encode contrast.

Finding contrast enhancement factor. We need to find an equivalent contrast at another (target) luminance level, which results in the same stereo precision (β) as the original contrast. We can rearrange (3) to compute the equivalent contrast c_{eq} for the desired logarithmic contrast c, source (Y_{in}) and target (Y_{out}) luminance:

$$c_{eq}(c, Y_{\text{in}}, Y_{\text{out}}) = \frac{-w_2 + \sqrt{w_2^2 - 4w_4 t}}{2w_4}, \tag{10}$$

where

$$\begin{aligned} t &= w_1\, L_{\text{out}} + w_3\, L_{\text{out}}^2 + w_5 - \beta(c, L_{\text{in}}) \\ L_{\text{in}} &= \log_{10} Y_{\text{in}} \qquad L_{\text{out}} = \log_{10} Y_{\text{out}} \end{aligned} \tag{11}$$

with the parameters $w_1, ..., w_5$ reported in Table 1. Function $\beta(\cdot)$ is given in (3). The process of mapping contrast between luminance levels is further illustrated in Figure 8.

Instead of directly modifying contrast in the Laplacian pyramid, we compute a contrast enhancement factor:

$$m_i(\mathbf{x}) = \frac{c_{\text{eq}}\Big(c_i(\mathbf{x}), Y_{\text{in}}(\mathbf{x}), Y_{\text{out}}(\mathbf{x})\Big)}{c_i(\mathbf{x})}, \tag{12}$$

where c_i is an input RMS contrast computed according to (9) and $c_{\text{eq}}()$ is the equivalent contrast function from (10). Y_{in} and Y_{out} are source and target luminance which are computed as:

$$\begin{aligned} Y_{\text{in}}(\mathbf{x}) &= 10^{G_N(\mathbf{x})} \cdot Y_{\text{peak,src}}\,, \\ Y_{\text{out}}(\mathbf{x}) &= 10^{G_N(\mathbf{x})} \cdot Y_{\text{peak,trg}}\,, \end{aligned} \tag{13}$$

where G_N is the base-band of the Gaussian pyramid (as explained in Section 5.2). $Y_{\text{peak,src}}$ is the peak luminance of the source display (before dimming) and $Y_{\text{peak,trg}}$ is the peak luminance of the target (dimmed) display. We use $Y_{\text{peak,src}} = 80$ cd/m^2 in all our experiments.

Contrast enhancement. Given the local contrast estimate computed in Section 5.2, we retarget it, enhancing locally the Laplacian pyramid:

$$\tilde{P}_i(\mathbf{x}) = P_i(\mathbf{x}) \cdot m_i(\mathbf{x}), \tag{14}$$

where P_i is the i-th level of Laplacian pyramid ($i = 1, ..., N-1$, excluding the base-band) and m_i is a corresponding enhancement factor from (12).

We reconstruct the resulting enhanced luminance channel y_{enh} by summing all N levels of pyramid $\tilde{P}$ including the base band:

$$y_{\text{enh}}(\mathbf{x}) = 10^{\sum_{i=1}^{N} \tilde{P}_i(\mathbf{x})}. \tag{15}$$

5.5 Reconstructing color image

The enhanced color image I_{enh} is produced by multiplying input color (RGB) image in linear space I_{input} by the ratio of enhanced and input luminances:

$$I_{\text{enh}}(\mathbf{x}, k) = I_{\text{input}}(\mathbf{x}, k)\, \frac{y_{\text{enh}}(\mathbf{x})}{y_{\text{input}}(\mathbf{x})}\,, \tag{16}$$

where k is the index of the color channel ($k \in \{1, 2, 3\}$). Such an approach may, however, result in out-of-gamut colors (one of the color channels values greater than 1) and distorted or desaturated colors. To prevent this, we compute how much a particular pixel can be enhanced until the pixel exceeds the gamut:

$$m_{\max}(\mathbf{x}) = \frac{1}{\max_k \Big\{I_{input}(\mathbf{x}, k)\Big\}}\,. \tag{17}$$

Next, we introduce this term into the previous gray to color conversion equation:

$$I_{\text{enh}}(\mathbf{x}, k) = I_{\text{input}}(\mathbf{x}, k) \cdot \min\left\{\frac{y_{\text{enh}}(\mathbf{x})}{y_{\text{input}}(\mathbf{x})}, m_{\max}(\mathbf{x})\right\}. \tag{18}$$

As the last step we convert the linear color channels to display-ready gamma-encoded ones with a gamma function: $I'(\mathbf{x}, k) = I^{1/\gamma}(\mathbf{x}, k)$. Figure 9 shows a comparison between naïve and saturation-aware methods. The left image shows visible, almost white, desaturated petals, while the saturation-aware method preserves colors.

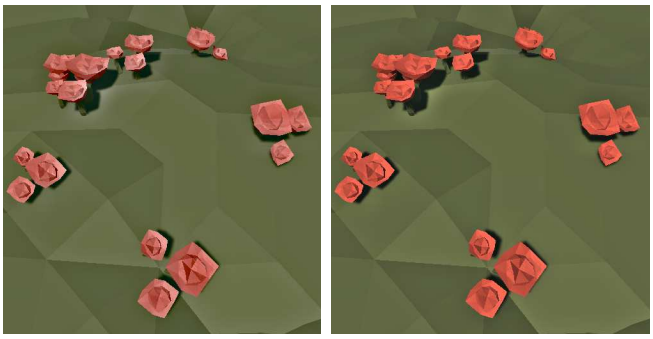

Fig. 9. Comparison between two methods of grayscale to color conversion. Left: naïve. Right: saturation-aware. Contrast enhancement is exaggerated for low-frequency bands for presentation purposes. The assets are a part of the POLYGON series prepared by Synty Store.

5.6 Implementation details

We implemented our enhancement method (Section 5) in the Unity game engine (v2019.2.19) as a post-processing shader, which took an anti-aliased rendered frame as an input. The code was optimized to ensure that we could run at a stable frame rate of 90 frames per second on a laptop with an NVIDIA GeForce 1080 graphics card and Intel® Core™ i9-8950HK 2.90 GHz CPU. The performance was measured with the use of Unity built-in profiler.

6 EXPERIMENT 2: VALIDATION

We evaluated the effectiveness of our method in a validation experiment in which we compared the proposed enhancement algorithm with the most closely related method of Wanat et al. [2014] and standard rendering. The method of Wanat et al. is meant to preserve color and contrast appearance across the luminance range, which can potentially result in an image that also provides better stereoscopic depth cues. The methods were compared in terms of the impression of three-dimensionality and the appearance of the presented scene. We believe that such a preference experiment is required to ensure that proposed method brings practical benefit without any negative effects.

As alternative we considered measuring the response times in a 3D task. However, the design of such an experiment is difficult as the completion time is not only affected by 3D depth perception but also by display and input latency, training, experience, and other effects.

VR headset. The experiments were prepared for the Valve Index VR headset, which offers a relatively high display resolution of a maximum of 16 pixels per visual degree. Additionally, the drivers of this device allow the user to dim its display. Using a luminance meter Minolta LS-100 and manipulating brightness settings, we achieved the desired peak luminance of 5 cd/m^2. While we chose Valve Index due to the convenient settings allowing for display dimming, the effect could be reproduced on other devices (e.g. HTC Vive Pro 2, Oculus Quest 2) after proper dimming (e.g. with the use of neutral density filters).

Fig. 10. Preview of the scene presented to the observer in the preference experiment. The images show a non-enhanced (standard) rendering of the scene. The assets are a part of the POLYGON series prepared by Synty Store.

Stimuli. The test scene was built from stylized assets that provided a good balance between good quality content, similar to those found in most VR experiences, and performance (no complex geometry). An example screenshot from the scene is shown in Figure 10. Figure 11 shows three rendering modes used in the experiment: proposed stereo-constancy model (top), standard rendering (middle), and Wanat's method (bottom). It should be noted that we implemented only the local processing part of the Wanat et al. method to evaluate only that aspect and avoid confounding factors such as changes in color or image brightness. The VR scenes were rendered at 90 Hz.

Procedure. During the experiment, we placed the observers in the virtual environment and teleported them to 5 different locations. The initial observers' orientation was randomized, but then they were allowed to look around freely. Additionally, observers were able to switch between two rendering methods using the trackpad on the right controller. In each trial, they compared our method with either standard rendering (no post-processing) or the method of Wanat et al. [2014].

The experiment was split into two parts. In each part, the observers visited the same locations (in random order), regardless of the question being asked. In the first part, the observers were asked to select the rendering mode that looks *more three dimensional* and in the second part they were asked to select the rendering mode that *looks better*. The observers gave the answer by pressing the

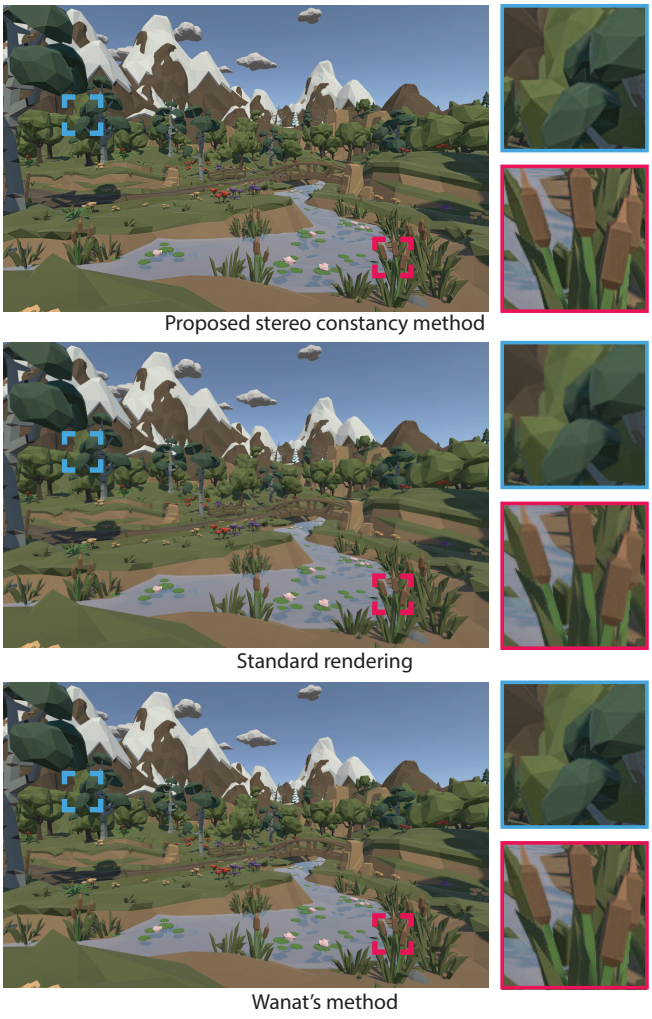

Fig. 11. Three rendering methods used in the preference experiment: image enhanced with the proposed stereo-constancy model (top), standard rendering with no enhancement (middle), and image enhanced with Wanat's method (bottom). The insets show close-ups of the selected image areas. It can be observed that Wanat's and the proposed stereo constancy methods increase local contrast and result in a sharper image. The assets are a part of the POLYGON series prepared by Synty Store.

right trigger while the selected rendering method was active. The participants were allowed to make a selection only after viewing both rendering modes. For both questions and every condition, each location was shown to the participant 5 times, each time using a different direction of the camera (random rotation around the up vector). The order of trials and parts was randomized. We also displayed information about the current progress of the experiment and the assessment criterion (depth or preference) at the bottom of the viewport.

The experiment lasted approximately 20 minutes. According to the post-experiment interviews, it was indicated that the session length was acceptable and did not cause excessive fatigue.

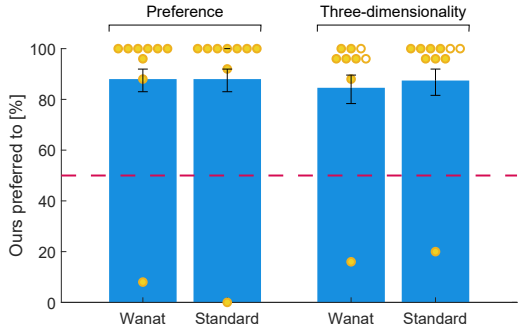

Fig. 12. Results of Experiment 2, assessing the preference and the impression of three-dimensionality. The yellow circles represent the per-observer results and the empty circles denote observers who failed the stereoacuity test. The height of the bars presents the percentage of trials in which our method was chosen over the method given below (excluding disqualified observers). *Standard* stands for a no enhancement and *Wanat* for the method proposed by Wanat et al. The error bars present a 95% confidence interval and the red dashed horizontal line indicates the guess rate.

Observers. Nine observers (age 22 to 30, 1 female and 8 males) were recruited from among students and researchers working in computer graphics and computer vision. All observers had a normal or corrected-to-normal vision and were also naïve to the purpose of the experiment. Before the experiment, each participant read and signed the consent form. The participants were screened for stereoacuity in a test performed in VR, in which they had to choose a closer square from a pair (akin the Titmus fly test). The results of the observers who failed that test were removed.

Results. The bars in Figure 12 show the percentage of trials in which our method was chosen over the alternative method. The yellow circles indicate per-observer results. Eight out of nine observers agreed that the image modified with our method looks more three-dimensional and also better than the image enhanced with Wanat et al.'s method and standard rendering. We further validated these results with a one-sided binomial test and a null hypothesis of random selection for both preference and impression of three-dimensionality. The tests confirmed that the results were significant and our contrast enhancement for stereo-constancy improves the perception of 3D shapes and produces more preferred images.

Even though Wanat et al.'s method was intended to improve image appearance, it was not selected when we asked about the preference. The most likely reason for its worse performance is the weaker strength of its enhancement. This is because the strength of the enhancement of that method depends on the changes in the contrast detection threshold (contrast sensitivity function) with luminance. Detection thresholds get smaller very quickly with luminance for high spatial frequencies, but they drop more slowly for low frequencies [Wuerger et al. 2020] (see also Figure 5). Since VR headset can reproduce relatively low spatial frequencies (up to 8 cycles per degree), the resulting enhancement was moderate.

From the post-experimental verbal survey, we found that observers based their decision on the fact that the preferred rendering mode appeared sharper and much clearer. Additionally, two observers stated that for the rendering with high contrast, colors

seemed to be more vivid and one person mentioned that the sharper image looked like it had a higher resolution. Moreover, no unnatural experiences of our method (including frame rate drop or lack of temporal stability) were reported by the observers.

In the survey, we also asked whether the color seen in the VR headset appeared natural. None of the participants reported any problems with color appearance.

7 DISCUSSION

No distortion of depth. The most important conclusion from our 3D shape perception experiment (Section 3) is that low luminance levels (0.1–10 cd/m^2) do not distort depth. This is in contrast to the observations of Kellnhofer et al. [2014], who reported compression of depth at low luminance. Our experiment showed that, even at 0.1 cd/m^2, observers could correctly assess the angle without bias (high accuracy), however with larger variance in their responses (lower precision). Had the perceived depth been compressed at low luminance, the results would have been biased towards obtuse angles, as the observers compensated for the reduced disparity. However, we did not experiment with light levels below 0.1 cd/m^2, so we cannot confirm whether the perception of 3D shapes is affected at these luminance levels.

Contrast vs. disparity manipulation. Several works [Didyk et al. 2011, 2012b; Kellnhofer et al. 2016] manipulate disparity to improve depth perception. While such an approach is practical for 3D cinema content, it is unsuitable for VR environments, in which depth must be faithfully reproduced to give accurate visual feedback to egomotion. Disparity manipulation in VR is likely to result in conflicting visual and vestibular sensations leading to VR-sickness [Jacobs et al. 2019].

Color appearance in mesopic vision. Degradation in stereoacuity is not the only issue of showing VR content at low brightness. It is well-established that color appearance also degrades as luminance decreases to mesopic vision [Barbur and Stockman 2010; Fu et al. 2012]. Those models could be incorporated into our enhancement technique, as it was done in [Wanat and Mantiuk 2014], however, we did not find the changes in color to be substantial enough to require additional processing. None of the experiment participants reported an unnatural color appearance in our post-experiment questionnaire (see Section 6). Our observation is also supported by the results of Kwak et al. [2003, Fig. 6,7] who reported negligible changes in colorfulness and hue, measured using magnitude estimation, when the reference white was reduced to only 1 cd/m^2. Larger changes in color appearance can be observed when two luminance conditions are presented to the observer simultaneously in an asymmetric (haploscopic) matching experiment [Shin et al. 2004]. Such artificial presentation, however, is not representative for viewing content on a dimmed VR headset.

Limitations. Our model was fitted to the data collected in the luminance range between 0.1 cd/m^2 and 1 000 cd/m^2, which may limit the ability of our model to generalize to very low luminance levels. We cannot generalize our model to the scotopic levels much below 0.1 cd/m^2, but we argue that such low luminance is less relevant for displays. We also do not consider the influence of tone mapping on depth perception. Since tone mapping often involves contrast compression, we expect increased difficulty of inferring depth from tone-mapped stereo images.

8 CONCLUSIONS

Dimming a display can be beneficial for VR experience as it reduces the visibility of flicker, saves power, prolongs battery life, and reduces the cost of the device. The major downside of this approach is the reduced sensitivity to stereoscopic depth cues. Contrary to previous works [Kellnhofer et al. 2014], we do not find the distortion of 3D depth at low luminance (0.1-1 cd/m^2), but instead, we find increased difficulty and lower precision (larger variance) of assessing 3D shapes based on binocular cues. This motivates our method for enhancing contrast at low luminance levels, intended at improving the reliability of stereoscopic depth cues. We demonstrate that such contrast enhancement can be implemented in the real-time rendering of VR environments. We further show the effectiveness of such depth enhancement in a perceptual experiment asking about qualitative aspects of preference and impression of depth. The experiment demonstrates that depth perception can be effectively restored by contrast enhancement and overall image quality can be improved. The proposed method can improve the user experience for VR headsets that need to operate at low power or those that cannot achieve high refresh rates.

ACKNOWLEDGMENTS

The authors wish to thank Minjung Kim, Maryam Azimi, and Joe G. March for their help. This work has received funding from the European Union's Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No. 765911 (RealVision), and from the European Research Council (ERC) grant agreement No. 725253 (EyeCode).

REFERENCES

Reynold Bailey, Cindy Grimm, Christopher Davoli, and Richard Abrams. 2007. The effect of object color on depth ordering. (2007).

J. L. Barbur and A. Stockman. 2010. Photopic, mesopic and scotopic vision and changes in visual performance. *Encyclopedia of the Eye* 3 (2010), 323–331.

P. Burt and E. Adelson. 1983. The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications* 31, 4 (1983), 532–540.

Dingcai Cao, Joel Pokorny, Vivianne C Smith, and Andrew J Zele. 2008. Rod contributions to color perception: linear with rod contrast. *Vision Research* 48, 26 (nov 2008), 2586–92. https://doi.org/10.1016/j.visres.2008.05.001

Alexandre Chapiro, Robin Atkins, and Scott Daly. 2019. A Luminance-Aware Model of Judder Perception. *ACM Transactions on Graphics* 38, 5, Article 142 (2019), 10 pages. https://doi.org/10.1145/3338696

J. Cutting and P. Vishton. 1995. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of Space and Motion (Handbook Of Perception And Cognition)*, W. Epstein and S. Rogers (Eds.). Academic Press, 69–117.

Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, and Hans-Peter Seidel. 2011. A Perceptual Model for Disparity. *ACM Transactions on Graphics*, Article 96 (2011), 10 pages. https://doi.org/10.1145/1964921.1964991

Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, and Hans-Peter Seidel. 2012a. Apparent stereo: the Cornsweet illusion can enhance perceived depth. In *Human Vision and Electronic Imaging XVII*, Bernice E. Rogowitz, Thrasyvoulos N. Pappas, and Huib de Ridder (Eds.), Vol. 8291. International Society for Optics and Photonics, SPIE, 180 – 191. https://doi.org/10.1117/12.907612

Piotr Didyk, Tobias Ritschel, Elmar Eisemann, Karol Myszkowski, Hans-Peter Seidel, and Wojciech Matusik. 2012b. A Luminance-Contrast-Aware Disparity Model and Applications. *ACM Transactions on Graphics* 31, 6 (2012).

A. Erickson, K. Kim, G. Bruder, and G. F. Welch. 2020. Effects of Dark Mode Graphics on Visual Acuity and Fatigue with Virtual Reality Head-Mounted Displays. In *2020*

IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 434–442.

John P Frisby and John E W Mayhew. 1978. Contrast Sensitivity Function for Stereopsis. *Perception* 7, 4 (1978), 423–429. https://doi.org/10.1068/p070423 arXiv:https://doi.org/10.1068/p070423 PMID: 704272.

Chenyang Fu, Changjun Li, Guihua Cui, M. Ronnier Luo, Robert W. G. Hunt, and Michael R. Pointer. 2012. An investigation of colour appearance for unrelated colours under photopic and mesopic vision. *Color Research & Application* 37, 4 (2012), 238–254. https://doi.org/10.1002/col.20691

B Y M A Georgeson and G D Sullivan. 1975. Contrast constancy: deblurring in human vision by spatial frequency channels. *The Journal of Physiology* 252, 3 (1975), 627–656.

David M. Hoffman and Grace Lee. 2019. Temporal Requirements for VR Displays to Create a More Comfortable and Immersive Visual Experience. *Information Display* 35, 2 (mar 2019), 9–39. https://doi.org/10.1002/msid.1018

Jochen Jacobs, Xi Wang, and Marc Alexa. 2019. Keep It Simple: Depth-Based Dynamic Adjustment of Rendering for Head-Mounted Displays Decreases Visual Comfort. *ACM Trans. Appl. Percept.* 16, 3, Article 16 (2019), 16 pages.

Petr Kellnhofer, Piotr Didyk, Karol Myszkowski, Mohamed M. Hefeeda, Hans-Peter Seidel, and Wojciech Matusik. 2016. GazeStereo3D: Seamless Disparity Manipulations. *ACM Transactions on Graphics* 35, 4 (2016). https://doi.org/10.1145/2897824.2925866

Petr Kellnhofer, Tobias Ritschel, Peter Vangorp, Karol Myszkowski, and Hans-Peter Seidel. 2014. Stereo Day-for-Night: Retargeting Disparity for Scotopic Vision. *ACM Trans. Appl. Percept.* 11, 3 (2014).

Min H Kim, Tim Weyrich, and Jan Kautz. 2009. Modeling human color perception under extended luminance levels. In *ACM SIGGRAPH 2009 papers*. 1–9.

John Krauskopf and Jason D Forte. 2002. Influence of chromaticity on vernier and stereo acuity. *Journal of Vision* 2, 9 (2002), 6–6.

J.J. Kulikowski. 1976. Effective contrast constancy and linearity of contrast sensation. *Vision Research* 16, 12 (jan 1976), 1419–1431. https://doi.org/10.1016/0042-6989(76)90161-9

Youngshin Kwak, Lindsay William MacDonald, and M. Ronnier Luo. 2003. Mesopic color appearance. In *Human Vision and Electronic Imaging VIII*, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.), Vol. 5007. International Society for Optics and Photonics, SPIE, 161 – 169. https://doi.org/10.1117/12.477371

Manuel Lang, Alexander Hornung, Oliver Wang, Steven Poulakos, Aljoscha Smolic, and Markus Gross. 2010. Nonlinear Disparity Mapping for Stereoscopic 3D. *ACM Transactions on Graphics* 29, 4, Article 75 (2010), 10 pages. https://doi.org/10.1145/1778765.1778812

Margaret S. Livingstone and David H. Hubel. 1994. Stereopsis and positional acuity under dark adaptation. *Vision Research* 34, 6 (1994), 799–802.

Thomas Luft, Carsten Colditz, and Oliver Deussen. 2006. Image Enhancement by Unsharp Masking the Depth Buffer. *ACM Transactions on Graphics* 25, 3 (2006), 1206–1213. https://doi.org/10.1145/1141911.1142016

Ming Ronnier Luo and Changjun Li. 2013. *CIECAM02 and Its Recent Developments*. Springer New York, New York, NY, 19–58. https://doi.org/10.1007/978-1-4419-6190-7_2

Rafał Mantiuk, Allan G. Rempel, and Wolfgang Heidrich. 2009. Display considerations for night and low-illumination viewing. In *Proc. of Symposium on Applied Perception in Graphics and Visualization - APGV '09*. 53–58. https://doi.org/10.1145/1620993.1621005

Thomas Oskam, Alexander Hornung, Huw Bowles, Kenny Mitchell, and Markus Gross. 2011. OSCAM - Optimized Stereoscopic Camera Control for Interactive 3D. *ACM Transactions on Graphics* 30, 6 (2011), 1–8. https://doi.org/10.1145/2070781.2024223

Eli Peli, Jian Yang, Robert Goldstein, and Adam Reeves. 1991. Effect of luminance on suprathreshold contrast perception. *Journal of the Optical Society of America A* 8, 8 (aug 1991), 1352. https://doi.org/10.1364/JOSAA.8.001352

M. Schuchhardt, S. Jha, R. Ayoub, M. Kishinevsky, and G. Memik. 2015. Optimizing mobile display brightness by leveraging human visual perception. In *2015 International Conference on Compilers, Architecture and Synthesis for Embedded Systems (CASES)*. 11–20.

Jae Chul Shin, Hirohisa Yaguchi, and Satoshi Shioiri. 2004. Change of Color Appearance in Photopic, Mesopic and Scotopic Vision. *Optical Review* 11, 4 (2004), 265–271. https://doi.org/10.1007/s10043-004-0265-2

David R Simmons and Frederick AA Kingdom. 1994. Contrast thresholds for stereoscopic depth identification with isoluminant and isochromatic stimuli. *Vision Research* 34, 22 (1994), 2971–2982.

Gurjot Singh, Stephen R. Ellis, and J. Edward Swan. 2018. The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. *IEEE Transactions on Visualization and Computer Graphics* 26, 2 (2018), 1385–1398. https://doi.org/10.1109/TVCG.2018.2869729 arXiv:1712.00088

Bjørn Stabell and Ulf Stabell. 1998. Chromatic rod–cone interaction during dark adaptation. *J. Opt. Soc. Am. A* 15, 11 (Nov 1998), 2809–2815. https://doi.org/10.1364/JOSAA.15.002809

Tom Troscianko, Rachel Montagnon, Jacques Le Clerc, Emmanuelle Malbert, and Pierre-Louis Chanteau. 1991. The role of colour as a monocular depth cue. *Vision research* 31, 11 (1991), 1923–1929.

Khrystyna Vasylevska, Hyunjin Yoo, Tara Akhavan, and Hannes Kaufmann. 2019. Towards Eye-Friendly VR: How Bright Should It Be?. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 566–574. https://doi.org/10.1109/VR.2019.8797752

Robert Wanat and Rafał K. Mantiuk. 2014. Simulating and compensating changes in appearance between day and night vision. *ACM Trans. Graph.* 33 (2014), 147:1–147:12.

John P Wann, Simon Rushton, and Mark Mon-Williams. 1995. Natural problems for stereoscopic depth perception in virtual environments. *Vision research* 35, 19 (1995), 2731–2736.

Simon J. Watt, Kurt Akeley, Marc O. Ernst, and Martin S. Banks. 2005. Focus cues affect perceived depth. *Journal of Vision* 5, 10 (12 2005), 7–7. https://doi.org/10.1167/5.10.7 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/932834/jov-5-10-7.pdf

Felix A Wichmann and N Jeremy Hill. 2001. The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics* 63, 8 (2001), 1293–1313.

Sophie Wuerger, Maliha Ashraf, Minjung Kim, Jasna Martinovic, María Pérez-Ortiz, and Rafal K. Mantiuk. 2020. Spatio-chromatic contrast sensitivity under mesopic and photopic light levels. *Journal of Vision* 20, 4 (apr 2020), 23. https://doi.org/10.1167/jov.20.4.23

Fangcheng Zhong, Akshay Jindal, Ali Özgür Yöntem, Param Hanji, Simon J. Watt, and Rafał K. Mantiuk. 2021. Reproducing reality with a high-dynamic-range multi-focal stereo display. *ACM Transactions on Graphics* 40, 6 (dec 2021), 1–14. https://doi.org/10.1145/3478513.3480513