Depth from HDR: Depth Induction or Increased Realism?

Peter Vangorp*   Rafał K. Mantiuk[†]   Bartosz Bazyluk[‡]   Karol Myszkowski*

Radosław Mantiuk[‡]   Simon J. Watt[†]   Hans-Peter Seidel*

MPI Informatik*   Bangor University[†]   West Pomeranian University of Technology[‡]

Abstract

Many people who first see a high dynamic range (HDR) display get the impression that it is a 3D display, even though it does not produce any binocular depth cues. Possible explanations of this effect include contrast-based depth induction and the increased realism due to the high brightness and contrast that makes an HDR display "like looking through a window". In this paper we test both of these hypotheses by comparing the HDR depth illusion to real binocular depth cues using a carefully calibrated HDR stereoscope. We confirm that contrast-based depth induction exists, but it is a vanishingly weak depth cue compared to binocular depth cues. We also demonstrate that for some observers, the increased contrast of HDR displays indeed increases the realism. However, it is highly observer-dependent whether reduced, physically correct, or exaggerated contrast is perceived as most realistic, even in the presence of the real-world reference scene. Similarly, observers differ in whether reduced, physically correct, or exaggerated stereo 3D is perceived as more realistic. To accommodate the binocular depth perception and realism concept of most observers, display technologies must offer both HDR contrast and stereo personalization.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Display Algorithms; J.4 [Computer Applications]: Social and Behavioral Sciences—Psychology

Keywords: luminance, contrast, stereo 3D, binocular disparity

1 Introduction

Traditional low dynamic range (LDR) display technologies can only display images with a limited range of luminance. To compress the luminance range of an image for such a display, tone-mapping operators cannot avoid distorting the contrast. High dynamic range (HDR) displays can produce a wider range of luminance and can therefore reproduce nearly undistorted contrast.

Many people who first see an HDR display get the impression that it is a 3D display, even though it does not produce any binocular depth cues like a stereo 3D display. However, contrast is closely related to depth cues such as aerial perspective and translucency, so undistorted (or enhanced) contrast reproduction might induce an additional 3D shape percept. When an image closely resembles a real scene in photometric terms, it is more likely to be interpreted as less flat and more three-dimensional, i.e., less as a rendition on a 2D plane and more like a real scene. This effect can be observed to
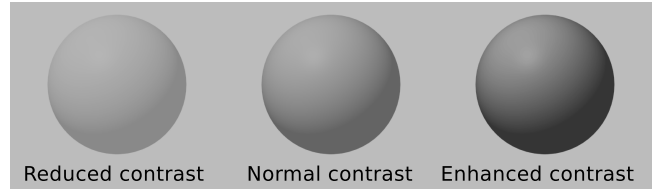
---

*e-mail: {pvangorp,karol,hpseidel}@mpi-inf.mpg.de
[†]e-mail: {mantiuk,s.watt}@bangor.ac.uk
[‡]e-mail: {bbazyluk,rmantiuk}@wi.zut.edu.pl

Figure 1: *The difference in contrast affects how we perceive depth in an image. Three spheres differ only in contrast (γ is equal to 0.5, 1 and 2 for each sphere respectively, refer to Equation 1). The sphere on the left with the lowest contrast appears to be "flat", while the sphere on the right seems to stand out the most from the background and seems to have more "depth". Our goal is to investigate whether contrast actually contributes to the metrical cues of our depth perception.*

a lesser extent in LDR images, for example in Fig. 1. The images of three-dimensional objects that have lower contrast appear more as a flat texture on the plane rather than a 3D shape. This effect extends to HDR where the impression of depth could be further improved.

This effect of HDR on 3D shape perception has been studied before [Ichihara et al. 2007; Rempel et al. 2011; Easa et al. 2013]. However, these studies considered contrast and luminance as an isolated depth cue in a scene, which does not apply to most practical applications. In contrast to these studies, we investigate the depth illusion created by HDR contrast and brightness in the presence of real depth cues such as binocular stereo using a specially constructed HDR stereoscope.

In this paper we investigate how HDR contrast (and luminance) alters the perception of 3D shape and realism in images. In particular, we examine whether enhanced contrast and brightness do in fact contribute to 3D shape perception and induce additional depth, or whether HDR images are just interpreted as a more realistic rendition of real-world scenes but do not affect metric depth judgments.

If contrast can induce additional metric depth in the presence of binocular depth cues, HDR could be used to magnify the perceived depth of stereo 3D displays beyond the range afforded by the vergence/accommodation comfort zone [Shibata et al. 2011]. If on the other hand contrast does *not* affect metric depth but only gives the illusion of 3D shape in 2D displays, it could offer a graceful degradation method for retargeting stereo 3D content to legacy 2D displays for applications where accurate metric depth judgments are not required and partially retaining the impression of 3D shape is sufficient. This includes most stereo 3D content where it is acceptable to compress depth to the vergence/accommodation comfort zone, which also impairs accurate metric depth judgments.

We also investigate how contrast reproduction and binocular depth cues both contribute to the sense of realism of a stereo 3D image. If the sense of realism is strongly correlated to 3D shape perception, then we expect a similar interaction between contrast and binocular depth cues. Otherwise, their relative importance as realism cues could be different from their relative strengths as depth cues. We examine how the optimal contrast reproduction and binocular depth

cues for realism vary across observers and provide guidelines for practical applications on multi-user and personalized HDR stereo 3D displays.

2 Previous Work

Contrast-based depth induction Ichihara et al. [2007] have shown with simple stimuli that image contrast can give the impression of depth on standard dynamic range displays. They suggest that the contrast in their stimuli may be interpreted as aerial perspective. A foggy but homogeneous medium between the observer and the stimulus would reduce contrast based on the viewing distance. Therefore, a reduced contrast can be interpreted as an increased distance.

Ichihara et al. made a distinction between *area* and *texture* contrast. Area contrast is the contrast between objects, or between an object and the background. For example, a nearby building may appear brighter than a building farther away in a fog. Texture contrast is the contrast within an object due to texture and shading. For example, because of the atmospheric effects, a nearby tree appears to have much stronger texture contrast in the leaves than the same tree seen from a distance. Ichihara et al. found that both area and texture contrast create an effective depth impression.

Rempel et al. [2011] confirmed these findings and hypothesized that contrast-based depth induction could explain the depth-from-HDR effect. HDR images differ from LDR images by their increased contrast and luminance. They showed that contrast-based depth induction also occurred with HDR stimuli.

There are also indications that contrast-based depth induction may not completely explain the depth-from-HDR effect. Contrast-based depth induction works equally well in LDR conditions; in fact Ichihara et al. showed that the effect was strongest for low luminance and contrast levels. Yet the depth-from-HDR effect does not appear to be reproducible in LDR conditions.

Easa et al. [2013] evaluated the strength of different monocular depth cues in the scenario in which several layers of an MRI scan are visualized. The depth cues included area contrast ("brightness" cue) and luminance contrast ("contrast" cue) shown on an HDR display. They found that both area contrast and texture contrast provide very strong cues for intuitive depth ordering, comparable to the cue of relative size, which is one of the strongest monocular indicators of depth ordering.

All the discussed studies showed that contrast can induce depth when presented isolated from other depth cues. In contrast to these studies, we want to measure the role of contrast in the presence of another strong depth cue, which is binocular disparity. This allows us to explore their interaction and their relative effect on depth perception and realism.

Contrast enhancement in photography Various studies [Calabria and Fairchild 2003; Daly et al. 2013] have concluded that people prefer images that have higher contrast, which overall eases comprehension of their content and makes them look more like real scenes [Yoshida et al. 2006].

Depth-enhanced sense of naturalness and presence Seuntiëns et al. [2006] investigated stereo 3D image quality as a function of JPEG-compression and various camera interaxial distances (IAD), and found that the resulting depth levels do not affect quality judgements. As confirmed in other studies as well [Kaptein et al. 2008] the added value of perceived depth is not incorporated in the outcome of subjective image quality assessment. Kaptein et al.

[2008] argue that naturalness[1] incorporates both depth and image quality, which makes it suitable for the assessment of both 2D and 3D performance. Overall, the naturalness of unimpaired 3D images is typically rated higher than their 2D counterparts, which is also the case when the same amount of noise [Seuntiëns et al. 2005] or blur [Kaptein et al. 2008] is added to both images. Remarkably, for a given amount of blur the naturalness seems to be unaffected by depth levels. Moreover, the apparent sharpness can even be enhanced in 3D images [Berthold 1997].

IJsselsteijn et al. [1998] found that the sense of presence can be enhanced by an increase in depth, which itself must be perceived as natural. Interestingly, the low correlation between the depth and naturalness of depth might suggest that consistent maintaining of depth naturalness is difficult even in case of professionally prepared stereo 3D footage.

Since the feeling of presence is less appropriate for still images, in this work we focus on naturalness (realism), and we investigate interactions between contrast and depth levels, including the issue of apparent depth induction by 2D/3D HDR images.

Comparisons between reality and displayed images Several works have studied the perceived differences between reality and displayed images produced by rendering algorithms [Meyer et al. 1986; Drago and Myszkowski 2001; McNamara 2006] or HDR photography [Yoshida et al. 2006]. However, these previous works performed only simple uniform luminance calibration of the display and did not produce the correct binocular cues that match the depth of the real scene.

3 Apparatus

All experiments in this paper were performed on an HDR stereoscope setup, consisting of a commercial HDR display and a custom-built Wheatstone stereoscope [Wheatstone 1838]. This setup was calibrated to ensure accurate binocular depth cues, spatially homogeneous luminance, and optimal color reproduction of photographs.

High dynamic range display The HDR display is a commercial 47" SIM2 HDR47E LCD display with 1920×1080 resolution. The white LED backlight produces $4000 \, \text{cd} \, \text{m}^{-2}$ peak brightness and 20 000:1 static contrast.

Wheatstone stereoscope The HDR display was split into left and right halves of 960×1080 resolution which were presented to the corresponding eye using a Wheatstone stereoscope (Fig. 2). The optical path length (from the eyes via the mirrors M1 and M2 to the screen) was 87.3 cm. The stereoscope and the HDR display are enclosed in a black box with an aperture through which observers can watch the screen.

Stereo calibration The mirrors M1 and M2 have a 45° angle to the HDR display. The precise angles can be adjusted to ensure accurate ocular convergence (vergence) and binocular disparity depth cues.

All stimuli were generated with a pair of cameras in an orthostereoscopic setup with a chosen interaxial distance, without toe-in. Such a camera arrangement, when set to the typical interocular distance of 6.5 cm, ensures correct binocular depth cues that match those arising from viewing a real scene.

---

[1] Naturalness refers to the degree in which the image approximates reality, which is conceptually similar to the notion of realism used in this work.
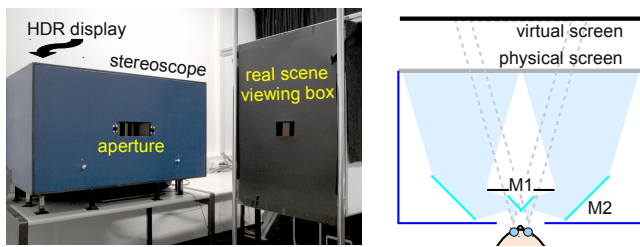
Figure 2: *Hardware setup of the HDR stereoscope and the real scene viewing box. Physical view frusta are shaded in light blue. Virtual view frusta are indicated by dashed gray lines. The cyan lines depict the locations of the mirrors.*

Homogeneous luminance calibration For target luminances of 0.1, 1, 10, 100, and 1000 cd m^{-2}, the true luminance output at 16×9 spatial locations on the display was measured through the mirror optics of the stereoscope with a Konica Minolta LS-100 luminance meter. These measurements were used to compensate for small display non-linearities at high and low luminance, spatial variation of the display, and luminance loss in the mirror optics.

Color calibration For Experiment 2 (Sec. 5) a real scene with controlled lighting was built in a viewing box next to the HDR stereoscope, with a viewing aperture identical to that of the stereoscope. The viewing box is shown in Fig. 2 and a photograph of the real scene is shown in Fig. 3.

Figure 3: *Example photograph of the real scene used in Experiment 2.*

The real scene was photographed with a focal length that matches the field of view of the display to ensure the most accurate size and binocular depth cue reproduction. The scene was photographed with a Canon 550D DSLR camera as a stack of RAW images with exposures from 1/4000 to 30 s in 1-f-stop increments. The exposures were combined into an HDR image using *pfstools* [Mantiuk et al. 2007]. The images were taken at a small aperture ($f/8$) to ensure that all objects were in focus. A low ISO setting (200) was used to avoid noise. To achieve good color match, we performed an iterative color calibration procedure as follows:

1. Capture the target HDR photograph of the real scene and display it on the HDR screen.
2. Capture an iteration HDR photograph of the display through the mirror optics of the stereoscope and compute the homographic alignment to the target photograph.
3. Optimize the RGB-to-RGB color transformation matrix that minimizes the CIE 1976 $(L^\star, u^\star, v^\star)$ $\Delta E^\star_{uv}$ distances between the target and iteration images.

4. Apply that color transformation to the target image and display the result on the HDR screen.
5. Repeat steps 2–4 if necessary until the displayed image looks close enough to the real scene. In practice there was no more visible improvement after the first iteration.

4 Experiment 1: Near-threshold Depth Induction

First we want to verify whether increased luminance difference can affect near-threshold depth judgements. The experiment is inspired by the work of Ichihara et al. [2007] and Rempel et al. [2011]. However, instead of testing whether luminance difference affects depth perception, we want to test whether such a difference interferes with the binocular disparity cues. Burge et al. [2005] have shown that the depth cues of familiarity and convexity influence the metric depth judgements from binocular disparity. If we find that the luminance contrast can enhance or reduce the depth differences from binocular disparity, it could be used as an additional cue for the manipulation of depth in 3D stereo images. Such additional cues are highly useful as the "depth budget" of stereo 3D content is restricted by the vergence vs. accommodation conflict [Shibata et al. 2011].

4.1 Experiment

To confirm the reliability of our setup we started with a pilot experiment in which stimuli, task, and results from previous work [Rempel et al. 2011] were successfully replicated. This confirmed that luminance contrast can indeed induce a sense of depth in the absence of other known depth cues. Then we added near-threshold binocular depth cues to test how they affect this depth induction.

Stimuli The stimulus was an image of two vertically arranged disks (called A and B) with leaf textures (see Fig. 4). It was rendered with multi-sampling to achieve anti-aliasing and sub-pixel disparity precision. Both disks were generated separately for left and right eye with varying disparities corresponding to distances from the virtual screen plane (see Fig. 2). To simulate near-threshold depth visibility we used disparity angles of 4.29°, 4.28°, 4.27° (in front of the virtual screen) and 4.22° (behind the virtual screen) (or depth values of 86.7, 86.9, 87.1, and 88.2 cm, respectively). The possible disparity *differences* between all combinations of these four disparities were 0′, ±0.43′, ±0.85′, ±3.41′, ±3.83′, and ±4.26′. (or 0, ±1.4, ±2.9, ±11.7, ±13.2, and ±14.6 mm, respectively).

We considered three cases: both disks have a low average luminance of 50 cd m^{-2} ("dark–dark" case), both have a high average luminance of 1000 cd m^{-2} ("bright–bright" case), or in the final case one disk is dark and the other is bright ("dark–bright" case).

The leaf texture was a grayscale texture with a roughly Gaussian luminance histogram and a texture contrast of 0.5, defined as the Michelson contrast between the brightest and darkest pixels within a disk. The area contrast was defined as the Michelson contrast between the average luminance of a disk and the uniform "black" background luminance of 1 cd m^{-2}.

Experimental procedure Observers were asked to watch the stimuli on the HDR display and answer the question: "Which of the disks is *closer* to you?" All combinations of disparity values for both disks were tested in two-alternative-forced-choice (2AFC) trials. A training run of 5 pairs of disks covering a representative range of contrasts was first conducted to familiarize observers with the mechanics of the experiment. This was then followed by 220 tri-
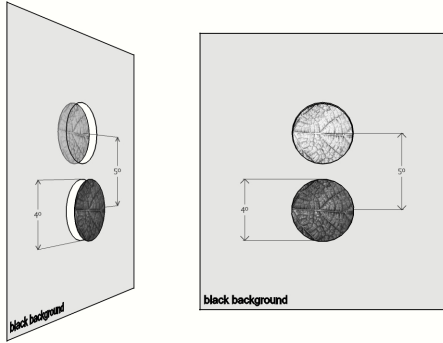
Figure 4: *Perspective and frontal sketch of a stimulus displayed on the screen during the "dark–bright" case of the experiment. Because of the black background there were no black outlines visible in the stimuli. The leaf textures are tonemapped for print reproduction and do not reflect the actual luminance and contrast used.*

als: 8 repetitions × all 28 combinations of 2 different stimuli from the pool of 8 stimuli (4 disparities × 2 brightnesses). The full set of 220 trials was randomized, and disk A or B was randomly presented on top or bottom. Observers went through the experiment at their own pace.

Observers Eleven observers participated in the experiment (2 females and 9 males, aged 20–41 years, average of 23.4 years). They had normal (6 observers) or corrected to normal vision (5 wore glasses) and were not stereo-blind. An experiment session took up to 20 min depending on the observer. All participants were naïve regarding the experiment's purpose.

4.2 Results

Fig. 5 presents the psychometric function that describes the probability of perceiving one disk in front of or behind the other. The results averaged over all $N = 11$ observers show that for the "dark–dark" and "bright–bright" cases, the perceived depth is consistent with the physical disparities, e.g., when disk B is closer to the observer (positive disparity difference), the observer correctly selected disk B as closer in most of the cases.

For the "dark–bright" case, however, observers' answers were biased towards selecting the brighter disk B as being closer, even though the disparity difference provided the opposite cue. This is visible as a horizontal shift of the red line in Fig 5. This confirms that luminance contrast combines with binocular disparity when making depth judgements, similarly to other monocular depth cues [Burge et al. 2005]. The effect, however, is very weak (equivalent to a depth difference of 1.4 mm) and disappears when the disparity is sufficiently strong and well above the detection threshold.

5 Experiment 2: Increased Realism

Experiment 1 showed that the quantitative effect of luminance and contrast on depth perception is limited. In the absence of other depth cues, luminance is an effective ordinal cue [Rempel et al. 2011; Easa et al. 2013]. However, when stronger cues such as binocular disparity are present, the effect of luminance and contrast on perceived depth is comparatively very weak.

Such a weak effect does not explain the enhanced sense of realism and "3D-ness" when viewing content on an HDR display. We
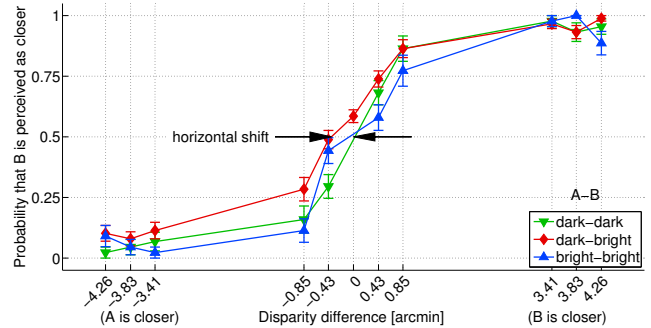

Figure 5: *Results of Experiment 1 averaged over all $N = 11$ observers. Error bars represent the standard error of the mean. The small horizontal shift of the red curve ("dark–bright") demonstrates near-threshold depth induction.*

therefore conducted another experiment to test our second hypothesis: HDR imagery is interpreted as a more realistic rendition of real-world scenes and therefore makes images appear more like actual objects rather than images drawn on a flat display plane.

5.1 Experiment

The goal of the experiment was to quantify any improvement in realism when scenes are rendered with enhanced contrast and with binocular depth cues. This was achieved by comparing displayed images with a physical scene acting as a reference.

Stimuli The effect of contrast and depth manipulations on realism can best be studied with photorealistic stimulus images. As described in Sec. 3 we used photographs of a real scene. This scene was a collection of diffuse, glossy, and transparent objects arranged inside the viewing frustum of the stereoscope. The distance from the observer to the objects ranged from 65 to 150 cm, which corresponds to the comfortable stereoscopic viewing zone that starts slightly in front of the virtual screen and extends quite far behind it [Shibata et al. 2011].

To allow manipulation of binocular disparity, the real scene (Fig. 3) was photographed for camera interaxial distances of 0, 3.25, 6.5, and 9.75 cm, which were controlled by sliding the position of the camera on a stereo photography rail. The stereo pair was displayed on the screen such that changing the interaxial distance resulted in a depth scaling towards or away from the screen, and objects at screen depth always remained at screen depth.

To manipulate contrast, the luminance values of the scene were altered using the formula:

$$Y_{\text{out}} = \left(\frac{Y_{\text{in}}}{Y_{\text{white}}} \right)^{\gamma} \cdot Y_{\text{white}}, \tag{1}$$

where Y_{in} and Y_{out} were the luminance values before and after the manipulation, Y_{white} was the luminance of the white patch on the color checker, and the contrast level γ was set to 0.63, 0.79, 1.00, 1.26 or 1.59. A pilot experiment showed that the contrast level perceived as the most realistic was typically within this range. To avoid color shifts, we used the color-preserving contrast correction formula from [Mantiuk et al. 2009].

Experimental procedure In this pairwise comparison experiment, observers toggled between a pair of stereo photographs with different contrast levels or interaxial distances, and were asked to

"select the picture that looks *more realistic*". There were in total 20 stereo images (5 contrasts × 4 interaxial distances). All 55 combinations of two images with adjacent contrast or interaxial distance levels were repeated 5 times in randomized order.

28 observers participated in the experiment (8 females and 20 males, aged 21–29 years, average of 24.3 years). They had normal (20 observers) or corrected-to-normal vision (8 wore glasses). All reported that they saw depth in stereo images. An experiment session took up to 30 min depending on the observer. All participants were naïve regarding the experiment's purpose.

Afterwards, observers were shown the real scene and told that this was the reference. They were allowed to look at the physical scene whenever they felt it necessary to recall how it appears during one of two randomly assigned follow-up tasks:

- Repeat exactly the same experiment, this time with access to the reference (13 observers).

- Complete a short open-ended questionnaire asking about any color shifts, shape distortions, or other differences they observed between the real scene and the most realistic displayed images (15 observers). See supplemental material for details.

5.2   Results

Without reference   To convert pairwise comparison results into just-noticeable difference (JND) units of realism, we employed Thurstonian scaling extended to incomplete paired comparison designs [Silverstein and Farrell 2001]. First, we analyze the data collected for the condition in which no reference scene was shown. Because of the large differences in the results between the observers, it would not be appropriate to analyze the data for an average observer. Instead, we observed that the contrast or interaxial distance (IAD) vs. realism curves could have four different shapes that correspond to the observer's selection of the most realistic IAD and contrast condition. To cluster observers according to these four curve shapes, we fitted each observer's realism scores separately for IAD and contrast as parabolas to use the peak (the most realistic condition) and the quadratic term (how clearly the peak stands out) as features for clustering. We also fitted straight lines to use the slope (increasing, decreasing, or flat) as another feature.

The observers fall into one of four clusters, in terms of the *interaxial distances* perceived as most realistic:

(a) *reduced stereo* or no stereo at all (6 observers)
(a decreasing straight line with its highest observed value at the low end of the examined range of IADs or contrasts)
(b) *moderate stereo* (6 observers)
(an inverted-U-shaped parabola with its peak inside the range)
(c) *exaggerated stereo* (8 observers)
(an increasing straight line with its highest observed value at the high end of the range)
(d) *indifferent* to the various settings of the IAD (8 observers)
(a flat straight line without any clear peak or slope)

The results for these four clusters are shown in Fig. 6. While it could be expected that stereo disparity improves realism (or "naturalness") [Kaptein et al. 2008], we were surprised to find a large number of observers who did not notice much difference in realism at all (cluster (d) – 8 observers), or those who found any non-zero binocular disparity unrealistic (cluster (a) – 6 observers). The Wheatstone stereoscope used in the experiment does not introduce artifacts such as ghosting due to cross-talk, so a decrease in realism for non-zero disparity cannot be attributed to such artifacts. The observers who found exaggerated stereo most realistic (cluster (c))

also tended to find higher contrast more realistic (see the steep increasing curve in Fig. 6(c), second column). At the same time, those who selected zero IAD as the most realistic (cluster (a)), were more moderate in their selection of contrast.

The observers fall into one of the three clusters, in terms of the *contrast* perceived as most realistic:

(a) *reduced contrast* (3 observers);
(b) *moderate contrast* close to neutral $\gamma = 1$ (14 observers); or
(c) *exaggerated contrast* (11 observers).

The results for these three clusters are shown in Fig. 7. We found no observers who were *indifferent* to the contrast level. The realism variation between contrast levels is much stronger than the realism differences between IAD conditions, as evident from the much steeper realism JND curves in Fig. 7, second column.

The results show that the notion of "realism" is clearly subjective and differs a lot between individuals. Such large individual variations cannot be explained by the noise in the data since the error bars are relatively small. The variations also cannot be explained by the limited sensitivity of the visual system to contrast and IAD, since the differences in the realism scores are mostly above 1 JND.

A strong variation in the preference for contrast has been observed before on HDR displays. Daly et al. [2013] found that while a 0.1–2500 cd m^{-2} display offered sufficient contrast to satisfy 50% of observers, a display operating in the range 0.005–20 000 cd m^{-2} is needed to be able to satisfy 90% of observers. Note, however, that their study measured subjective preference rather than realism and no reference was shown.

Our sample size is too small to estimate the size of each cluster in the overall population. But since similar numbers of observers were classified in each stereo cluster, those clusters are all likely to be well-represented in the population. There is a tendency to select a higher rather than lower contrast (11 vs. 3 observers). Such a trend was also observed in [Yoshida et al. 2006], where the observers selected higher than original contrast even when presented with a real-world reference scene. There could be multiple reasons for such enhanced contrast, including a tendency to emphasize the unique aspects of an HDR display compared to an LDR display, or a compensation for other missing cues and lower quality of displayed images compared to the real-world scene.

There is surprisingly little interaction between the effects of contrast and IAD on perceived realism. The curves in Figs. 6 and 7 are mostly parallel. Both dimensions therefore seem to make independent contributions to the sense of realism.

With reference   To analyze the change in the results for the experiment completed with and without seeing the reference scene, we looked at how the shape of the 13 observers' curves changed and whether observers were classified into a different cluster. With the reference we observed a trend towards lower contrast (6 observers moved into a lower contrast cluster) and more moderate stereo. See supplemental material for details.

We argue that the results collected when showing a reference scene are less relevant for our research question. Such an experiment effectively measures how accurately the observers can match the appearance of the reference, instead of measuring what they consider as realistic appearance.

Questionnaire   The questionnaire results showed that the main shortcomings of the displayed virtual images were inaccurate brightness, contrast, color, and insufficient resolution. Note that
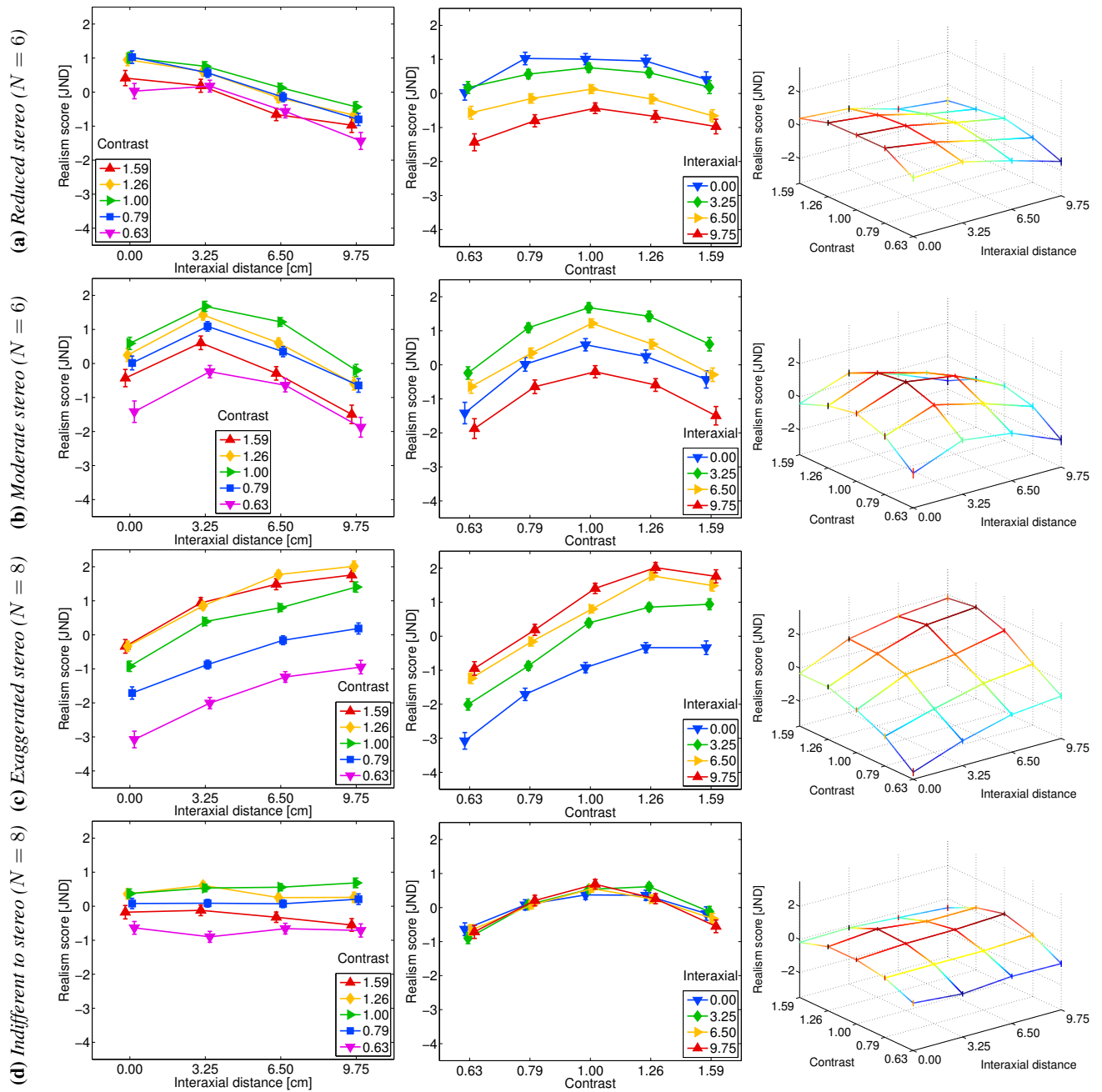
Figure 6: *Results averaged over observers, who were clustered according to their selection of the most realistic interaxial distance (stereo) (rows a–d). The number N in the sub-figure captions on the left indicates the number of observers that fall within each cluster. Error bars represent the standard error of the mean. Realism scores in JND units are interval scales with an arbitrary zero point and cannot be compared between clusters.*

these comments were made by non-experts so "color" could indicate problems with brightness and contrast. Some observers (5 out of 16) noticed shape distortions in the depicted images, though they all noted that these were small. Based on these comments, we can conclude that the accuracy of color and contrast seem to be the dominant attributes of realism.

6 Practical implications

Experiment 2 showed that the accuracy of contrast reproduction is a very important factor in producing realistic images both on 2D and stereo 3D displays. Note that many stereo 3D technologies reduce image brightness in order to produce 3D stereo pair and thus
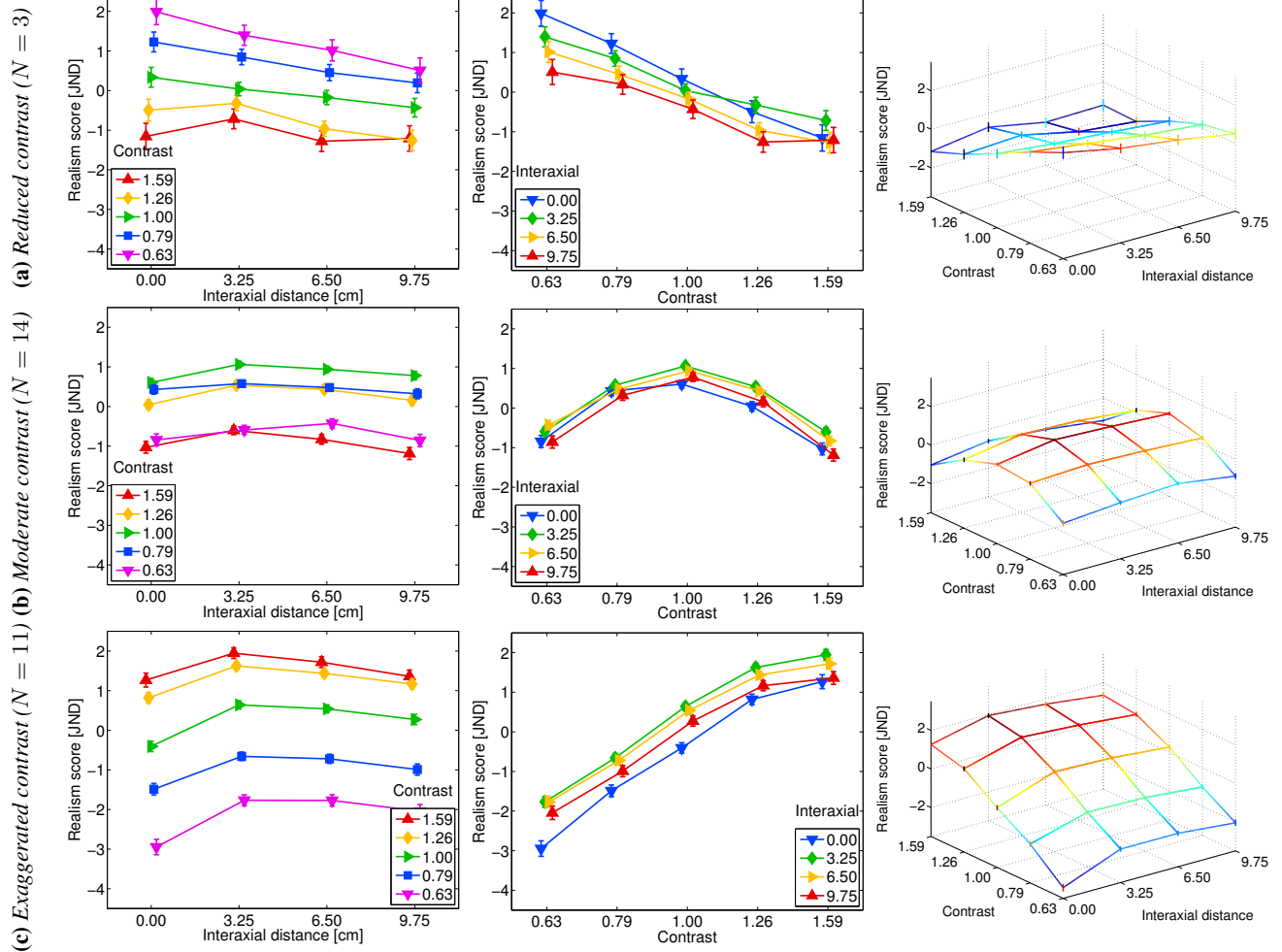
Figure 7: *Results averaged over observers, who were clustered according to their selection of the most realistic contrast (rows a–c). Error bars represent the standard error of the mean. Realism scores in JND units are interval scales with an arbitrary zero point and cannot be compared between clusters.*

decrease the effective image contrast. Although reducing display brightness does not change the physical contrast produced by a display, the contrast actually seen on the display is reduced by reflected ambient light, which elevates the black level. Therefore, when moving from 2D displays to stereo 3D, some gain in realism could be lost because of worse contrast reproduction. The majority of the observers consider as the most realistic the images reproduced with undistorted ($\gamma = 1$) or even enhanced contrast. Since the contrast shown on standard displays is usually distorted by some form of tone-mapping, such images may appear less realistic than their counterparts shown undistorted on HDR displays.

As shown in Experiment 1, luminance contrast has only a weak effect effect on depth perception. Therefore, the practical utility of luminance contrast as a depth enhancement cue in stereoscopic images may be limited. In the majority of complex images, the disparity difference will overpower the weak effect of luminance contrast. But luminance contrast can still be a useful depth-enhancement technique when the stereo signal is weak, for example when the depth contrast is away from the horopter and the sensitivity to binocular disparity is much lower.

The observation in Experiment 2 that many observers are not very sensitive to the disparity reduction can help in mapping a wide range of depths that can be found in the real-world to a much limited "depth budget" offered by most stereo 3D technologies. With smart counter-shading techniques [Didyk et al. 2012a; Didyk et al. 2012b], the range of disparities can be strongly reduced thus improving visual comfort, with little impact on perceived realism.

7 Conclusions

In this work we attempt to explain why images shown on an HDR display appear more three-dimensional than their counterparts on a standard display. To do so, we investigate two distinct qualities of a perceived image: depth perception (Experiment 1) and realism (Experiment 2). In both cases we modulate two visual cues: binocular disparity and contrast.

While the results of Experiment 1 confirm the findings from previous work [Easa et al. 2013; Rempel et al. 2011; Ichihara et al. 2007] that enhanced luminance contrast induces depth, they also show that it is a relatively weak depth cue, which is easily overpowered by (much stronger) binocular disparity. Therefore, the role of area or texture contrast in perceiving depth seems to be quite limited for most real-world scenes, as opposed to artificial isolated stimuli on a 2D display.

The luminance contrast, however, seems to have a bigger impact than binocular disparity on the assessment of realism. In Experiment 2 all observers indicated the preference for a certain contrast level. The preference for binocular disparity was more moderate, with some observers even unable to see much difference between different disparity settings. From that we can make another conclusion that depth perception is a required but not sufficient condition for achieving a high level of realism. If correct depth perception was the main factor in realism, the influence of binocular disparity on realism would be much stronger in Experiment 2.

This evidence supports the hypothesis that the "three-dimensional" appearance of images on an HDR display is more likely the result of images appearing closer to real scenes rather than providing additional or more accurate depth cues. The shapes shown on an HDR display do not "pop up" from the display plane, as happens in the case of stereo disparity. Nevertheless, they appear highly convincing. From that, we can also conclude that both factors, contrast and binocular disparity, are important and relevant for achieving realistic looking images but they contribute to different components of the overall impression of realism.

As future work, we plan to improve the quality of images produced by the display apparatus. From the questionnaire results and our own observation we can surmise that increasing the resolution and improving the accuracy of color reproduction should further improve the quality of realism shown on a display.

Acknowledgments

References

BERTHOLD, A. 1997. The influence of blur on the perceived quality and sensation of depth of 2D and stereo images. Tech. Rep. TR-H-232, ATR Human Inf. Process. Res. Lab., Kyoto, Japan.

BURGE, J., PETERSON, M. A., AND PALMER, S. E. 2005. Ordinal configural cues combine with metric disparity in depth perception. *Journal of Vision 5*, 6, 534–542.

CALABRIA, A. J., AND FAIRCHILD, M. D. 2003. Perceived image contrast and observer preference II: Empirical modeling of perceived image contrast and observer preference data. *The Journal of Imaging Science and Technology 47*, 494–508.

DALY, S., KUNKEL, T., SUN, X., FARRELL, S., AND CRUM, P. 2013. Viewer preferences for shadow, diffuse, specular, and emissive luminance limits of high dynamic range displays. *SID Symposium Digest of Technical Papers 44*, 1 (June), 563–566.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2012. Apparent stereo: The Cornsweet illusion can enhance perceived depth. In *Human Vision and Electronic Imaging XVII, Proc. Symposium on Electronic Imaging*, IS&T/SPIE, 1–12.

DIDYK, P., RITSCHEL, T., EISEMANN, E., MYSZKOWSKI, K., SEIDEL, H.-P., AND MATUSIK, W. 2012. A luminance-contrast-aware disparity model and applications. *ACM Trans. Graph. 31*, 6, 184:1–184:10.

DRAGO, F., AND MYSZKOWSKI, K. 2001. Validation proposal for global illumination and rendering techniques. *Computers & Graphics 25*, 3, 511–518.

EASA, H. K., MANTIUK, R. K., AND LIM, I. S. 2013. Evaluation of monocular depth cues on a high-dynamic-range display for visualisation. *ACM Trans. Appl. Percept. 10*, 3, 16.

ICHIHARA, S., KITAGAWA, N., AND AKUTSU, H. 2007. Contrast and depth perception: Effects of texture contrast and area contrast. *Perception 36*, 686–695.

IJSSELSTEIJN, W. A., DE RIDDER, H., HAMBERG, R., BOUWHUIS, D., AND FREEMAN, J. 1998. Perceived depth and the feeling of presence in 3DTV. *Displays 18*, 207–214.

KAPTEIN, R. G., KUIJSTERS, A., LAMBOOIJ, M. T. M., IJSSELSTEIJN, W. A., AND HEYNDERICKX, I. 2008. Performance evaluation of 3D-TV systems. In *Proc. Electron. Imaging Image Quality Syst. Perform. V*, vol. 6808, SPIE, 19:1–11.

MANTIUK, R., KRAWCZYK, G., MANTIUK, R., AND SEIDEL, H.-P. 2007. High dynamic range imaging pipeline: Perception-motivated representation of visual content. In *Proc. Human Vision and Electronic Imaging XII*, B. E. Rogowitz, T. N. Pappas, and S. J. Daly, Eds., vol. 6492, SPIE.

MANTIUK, R., MANTIUK, R., TOMASZEWSKA, A., AND HEIDRICH, W. 2009. Color correction for tone mapping. *Computer Graphics Forum 28*, 2, 193–202.

MCNAMARA, A. 2006. Exploring visual and automatic measures of perceptual fidelity in real and simulated imagery. *ACM Transactions on Applied Perception 3*, 3, 217–238.

MEYER, G. W., RUSHMEIER, H. E., COHEN, M. F., GREENBERG, D. P., AND TORRANCE, K. E. 1986. An experimental evaluation of computer graphics imagery. *ACM Transactions on Graphics 5*, 1, 30–50.

REMPEL, A. G., HEIDRICH, W., AND MANTIUK, R. 2011. The role of contrast in the perceived depth of monocular imagery. Tech. Rep. TR-2011-07, University of British Columbia.

SEUNTIËNS, P. J., HEYNDERICKX, I. E., IJSSELSTEIJN, W. A., VAN DEN AVOORT, P. M. J., BERENTSEN, J., DALM, I. J., LAMBOOIJ, M. T., AND OOSTING, W. 2005. Viewing experience and naturalness of 3D images. In *Optics East, Proc. 3-D TV, Video, Display IV*, vol. 6016, SPIE, 5–7.

SEUNTIËNS, P., MEESTERS, L., AND IJSSELSTEIJN, W. 2006. Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation. *ACM Trans. Appl. Percept. 3*, 2, 95–109.

SHIBATA, T., KIM, J., HOFFMAN, D. M., AND BANKS, M. S. 2011. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision 11*, 8, 11:1–29.

SILVERSTEIN, D., AND FARRELL, J. 2001. Efficient method for paired comparison. *Journal of Electronic Imaging 10*, 394–398.

WHEATSTONE, C. 1838. Contributions to the physiology of vision.–Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London 128*, 371–394.

YOSHIDA, A., MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2006. Analysis of Reproducing Real-World Appearance on Displays of Varying Dynamic Range. *Computer Graphics Forum (Proc. of Eurographics) 25*, 3 (Sept.), 415–426.