The effect of display capabilities on the gloss consistency between real and virtual objects

Bin Chen*
binchen@mpi-inf.mpg.de
Max-Planck-Institut für Informatik
Saarbrücken, Germany

Akshay Jindal*
akshay.jindal@intel.com
Intel Corporation
University of Cambridge
Bellevue, United States of America

Michal Piovarči
michal.piovarci@gmail.com
Institute of Science and Technology
Austria
Klosterneuburg, Austria

Chao Wang
chaowang@mpi-inf.mpg.de
Max-Planck-Institut für Informatik
Saarbrücken, Germany

Hans-Peter Seidel
hpseidel@mpi-inf.mpg.de
Max-Planck-Institut für Informatik
Saarbrücken, Germany

Piotr Didyk
piotr.didyk@usi.ch
Università della Svizzera Italiana
Lugano, Switzerland

Karol Myszkowski
karol@mpi-inf.mpg.de
Max-Planck-Institut für Informatik
Saarbrücken, Germany

Ana Serrano
anase@unizar.es
Universidad de Zaragoza, I3A
Zaragoza, Spain

Rafał K. Mantiuk
rafal.mantiuk@cl.cam.ac.uk
University of Cambridge
Cambridge, United Kingdom

ABSTRACT

The supplemental material provides additional details on:

- validation of signal correction pipeline, (Section 1)
- our experimental apparatus, (Section 2),
- MTF correction, (Section 3),
- camera-guided tone correction, (Section 4),
- tone mapping operators used in experiment 3, (Section 5),
- details of validation study in Sec. 6.3, (Section 6),
- details of our data analysis, (Section 7),
- all the stimuli used in main experiment (Sec. 4) and tone mapping experiment (Sec. 6), (Section 8).

CCS CONCEPTS

• **Computing methodologies** → *Appearance and texture representations*; *Perception*.

KEYWORDS

gloss perception, gloss reproduction, tone mapping, reproducing reality, HDR, computational display
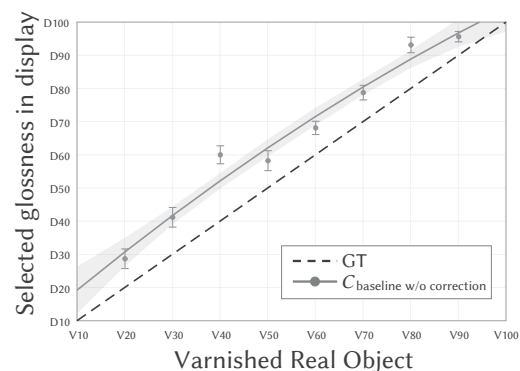
---

*Contribute equally

Figure 1: Results of our validation study on signal correction pipeline. The gloss level could not be faithfully reproduced without signal correction pipeline.

1 VALIDATION OF SIGNAL CORRECTION PIPELINE

A validation gloss-matching study was conducted without applying our signal correction pipeline. In total, 6 participants, age from 25 to 43, with 1 female and 5 males, volunteered in this user study. The experiment was approved by the department ethic board. The experiment procedure is the same as our C_{baseline} experiment. We show the result of this study in Fig. 1. A large error between the ground truth and selected gloss level has been observed. Our signal correction pipeline significantly improved the gloss reproduction accuracy, as shown in results of baseline experiment in Sec. 5 of the main paper.

2 EXPERIMENT APPARATUS

Our experimental setup consists of three main components: a high-dynamic range stereoscopic display, a *real-scene-box* (RSB), and a 2D camera gantry. The stereoscopic display uses two projector-based

9.7″ HDR displays [Seetzen et al. 2004], each with a resolution of 2048×1536. To prevent banding artifacts caused by insufficient bit-depth, the display algorithm utilizes spatiotemporal dithering. We align the LCD and projector planes, calibrate display colors, and align rendered viewpoints to user's eyes following the calibration procedure proposed by [Zhong et al. 2021]. We further optimize the display for psychophysical experiments as proposed by [Wuerger et al. 2020]. The apparatus also contains a RSB in front of the observer allowing a small-scale lightable physical scene seen through a pair of beam splitters. This enables a side-by-side presentation of a real object and its virtual counterpart (Fig. 2) positioned 65 cm distance away from the participants' eyes, producing a resolution of 120 pixels per visual degree. The box is made from black acrylic and its back wall is coated with Vantablack[1] to eliminate indirect illumination and improve the contrast of the captured images [Talvala et al. 2007]. A custom-built turntable was placed in the RSB (Fig. 2 (left)) that could hold and switch between four objects, thereby reducing the number of manual changes required during the experiment (Fig. 2 (middle)). Only the leftmost object and its virtual image are visible to the observer and the rest are outside the display's field of view (Fig. 2 (right)).

3 MTF CORRECTION

As discussed in Section 3.2 of the main paper, our captured images are prone to blur in sharp bright regions such as glossy specular highlights. We recover the lost high-frequency information by first estimating the camera's MTF and using it for deconvolution. To estimate the MTF we use the slanted edge technique as implemented in the SFRMAT4 MATLAB toolbox [Burns and Williams 2018]. To reduce measurement errors in the estimated MTF, we fit a sum of Gaussian to the data:

$$\text{MTF}(\rho) = a_1 \cdot e^{-\frac{(\rho - b_1)^2}{c_1^2}} + a_2 \cdot e^{-\frac{(\rho - b_2)^2}{c_2^2}}, \qquad (1)$$

where ρ is the spatial frequency in cycles/pixel, $a_1 = 0.935$, $b_1 = 0.066$, $c_1 = 0.266$, $a_2 = 9.679e + 08$, $b_2 = 27.115$, and $c_2 = 5.681$ are free parameters fitted to the noisy data from SFRMAT4 MTF estimation using least square minimisation. Note that before fitting, we clamped the MTF for spatial frequencies greater than 0.35 cy/px. This helps avoid enhancing noise and division by zero in the next deconvolution step. We can now use the smooth MTF filter to obtain a sharp image of the specular reflection by performing deconvolution:

$$I' = \mathfrak{F}^{-1}\left\{\frac{\mathfrak{F}\{I\}(\rho)}{\text{MTF}(\rho)}\right\}, \qquad (2)$$

where $\mathfrak{F}$ denotes the Fourier transform, I is the original blurry image and I' is the recovered sharp image.

4 CAMERA-GUIDED TONE CORRECTION

Since our MTF correction step is not sufficient to correct low-frequency aberrations, we adopt a camera-guided tone correction approach where we capture images of the real object and the rendered image side-by-side and use histogram matching to align their luminance histograms. Specifically, we find the intensity value

[1]Vantablack, Surrey Nanosystems

matching from the cumulative distribution function of displayed image CDF_d and the real object image CDF_r :

$$CDF_d(x_i) = CDF_r(x_j), i, j \in \{0, 1, ..., N\}, \qquad (3)$$

where N is the number of bins in the histogram and has been set to $1e6$. Then we formulate the mapping function as:

$$T(x_i) = CDF_d^{-1}(CDF_r(x_i)) = x_j, \qquad (4)$$

finally, the continuous version of the matching function can be further represented by a piecewise cubic Hermite polynomial:

$$T(x) = \{H_3^n(x)\}, x_n \le x \le x_{n+1}, n = 0, 1, ..., N - 1, \qquad (5)$$

where $H_3^n(x)$ is the nth cubic Hermite polynomial. To preserve the calibrated color, we perform the procedure for one color channel, which is green in our case as it covers the largest intensity range, and assume the same scaling applies to other channels. Note that we utilized the same T for all the gloss levels.

5 TONE MAPPING OPERATOR

We set three initial values that are related to the display capability: the starting value s, the ending value e, and the clipping luminance c. As shown in Fig. 4, the starting and ending points of cubic Bézier curve can then be represented as $P_0 = (s, s)$ and $P_3 = (e, c)$. A parameter $\rho \in [0, 1)$ is used to represent the steepness of a tone curve, from which the two controlling points can be calculated as $P_1 = (c - \rho(c - s), c - \rho(c - s))$ and $P_2 = (c + \rho(e - c), c)$. When $\rho = 0$, P_1 and P_2 will be overlapped to $P^* = (c, c)$ and the cubic Bézier curve shrink to a quadratic Bézier curve, while $\rho \to 1$ the cubic Bézier curve shrink to a linear Bézier curve and cause discontinuity at P_0 and P_3, which should be avoided. The value smaller than s is kept unchanged and the value larger than e is equal to the clipping luminance c, while the value in between could be compressed by tone mapping curve formulated explicitly as:

$$B(t) = (1 - t)^3 P_0 + 3(1 - t)^2 P_1 + 3(1 - t)^2 P_2 + t^3 P_3, 0 \le t \le 1 \quad (6)$$

There are two methods to control effect of tone curve: 1) adjusting ρ with pre-fixed starting and ending points; 2) adjusting the starting and ending points with ρ fixed. Either one could be used for tone curve generation. In the experiment, we want to examine the influence of clipping luminance and steepness of the tone mapping curve. We used the first method and set $s = 120, e = 1200$ for all designed four tone curves: $TM_{high-steep}, TM_{high-gradual}, TM_{low-steep}$, and $TM_{low-gradual}$, in which $high$ and low represent a large clipping luminance calculated as $c_{high} = 10^{0.5log(se)} = 379.47$ and a small clipping luminance calculated as $c_{low} = 10^{0.5log(sc_{high})} = 213.39$, $steep$ and $gradual$ correspond to $\rho = 0$ and $\rho = 0.3$ respectively. We show the results of four tone mapping operators together with HDR image in Fig. 8, 9, 10, 11, and 12.

6 DETAILS OF VALIDATION STUDY IN SEC. 6.3

Apparatus. A Dell UltraSharp 32 4K HDR Monitor (UP3221Q) with HDR 10/ Display HDR 1000 certificate was used in this experiment. We calibrate the display to BT. 2020 color space with D65 white point using gamma curve SMPTE ST.2084 PQ.
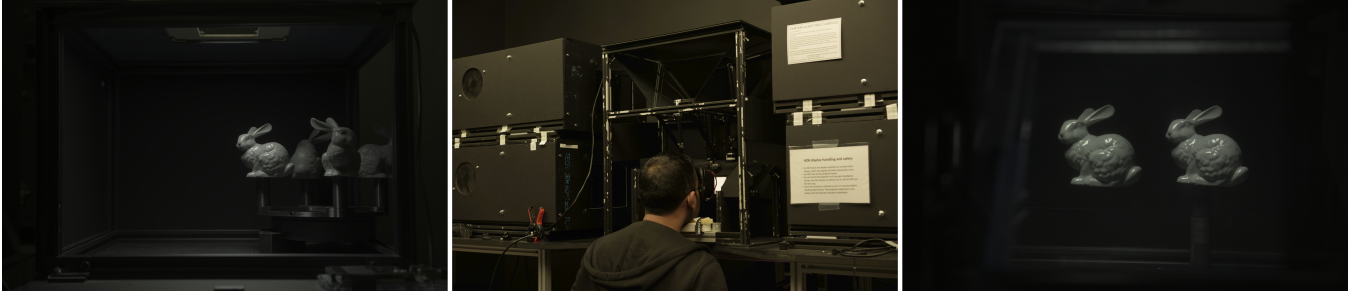
Figure 2: Experiment apparatus. The real scene box contains a turntable with 4 objects of different gloss levels. The observer sees only one real object in a given trial. A virtual counterpart of the object is rendered on an HDR stereoscope and shown side-by-side with the real object. The observer is asked to change the glossiness of the virtual object till it matches the real one.
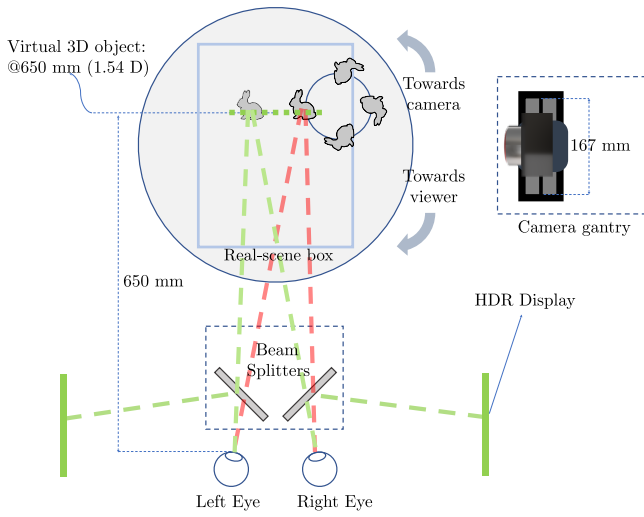

Figure 3: Illustration of our experimental apparatus that enables side-by-side comparison of HDR stereoscopic images with the real object.
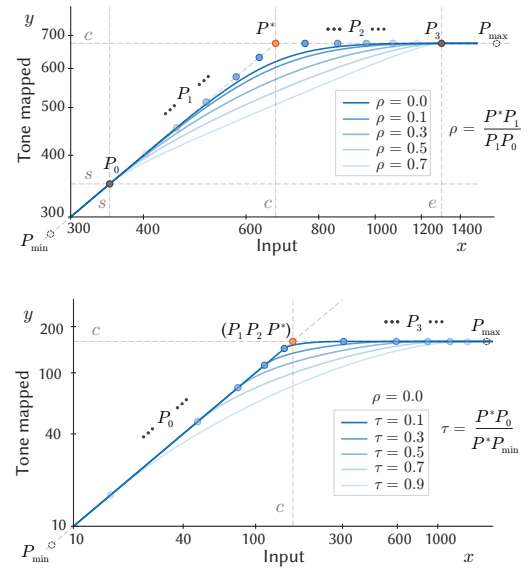

Figure 4: Tone mapping operator formulated by cubic Bézier curve.

Participants. Twelve volunteers participated in this experiment, aged from 25 to 48, with 2 females and 10 males, from the university. All participants have normal or corrected-to-normal vision, and were naïve to the experiment's goal. The experiment was approved by the department ethic board.

Procedure. During the experiment, participants adjusted their seats to make sure they were sitting around 80 cm away from the display (angular resolution of 80 pixels-per-degree (ppd)). A pair of stimuli was shown in the middle of the screen, and the participant's task was to: *"Select the image that has a higher gloss level.".* They could press *left* or *right* on the keyboard, indicating that left is glossier or right is glossier, respectively. This experiment consisted of 100 trials selected in total, which are actively sampled using [Mikhailiuk et al. 2021], for each participant and lasted for approximately 15 minutes.
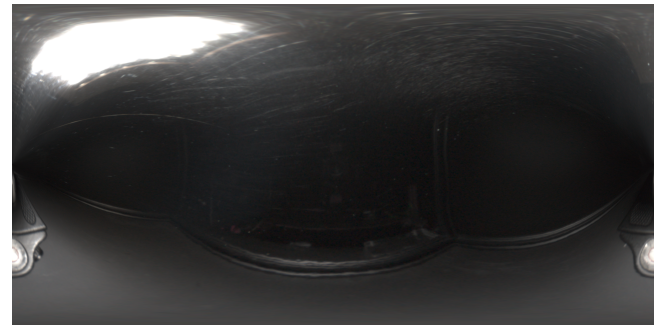

Figure 5: Captured latlong photograph of our illumination using reflective mirror ball.
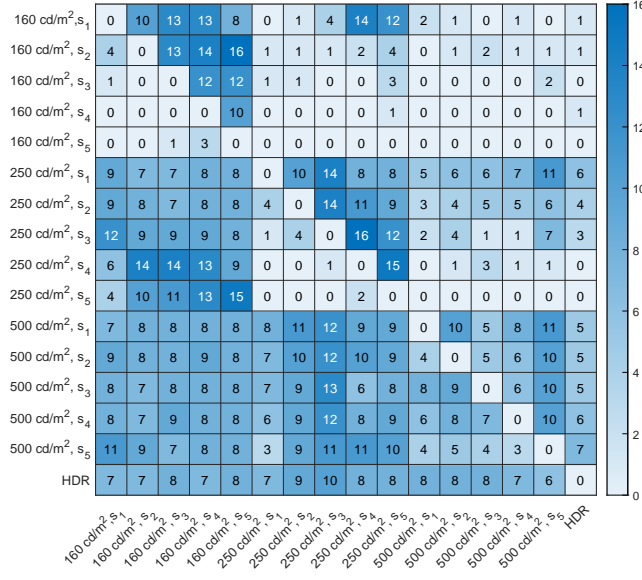
Figure 6: Heatmap of our pairwise comparison matrix M in tone mapping validation study.

Figure 7: Demonstration of captured light field stereo pair using 100% gloss level stimuli. View animation in Adobe Acrobat.

7 STATISTICAL ANALYSIS RESULTS

Tables 1, 2, 3, 4, and 5, and 6 show the detailed results of the statistical analysis for the factors affecting perceived gloss differences (Section 5) as described in the main paper. Tables 7, 8, 9, 10, and 11 show the detailed results of the statistical analysis for the effect of tone mapping (Section 6) as described in the main paper.

Table 1: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the *albedo* factor. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{baseline}$.

	Estimates	2.5% CI	97.5% CI	t	p
C_{black}	0.629	-0.301	1.157	1.131	0.188
Num. Obs	480				
R2 Nagelkerke	0.84				

Table 2: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the *stereo* factor. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{baseline}$.

	Estimates	2.5% CI	97.5% CI	t	p
C_{mono}	0.057	-0.869	0.983	0.121	0.904
Num. Obs	480				
R2 Nagelkerke	0.75				

Table 3: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the *dynamic range* factor. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{baseline}$.

	Estimates	2.5% CI	97.5% CI	t	p
C_{TM}	1.821	0.854	2.815	3.648	<0.001
Num. Obs	480				
R2 Nagelkerke	0.71				

Table 4: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the interaction of the *dynamic range* and *stereo* factors. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{baseline}$.

	Estimates	2.5% CI	97.5% CI	t	p
C_{mono}	0.050	-0.844	0.945	0.110	0.912
C_{TM}	1.763	0.821	2.761	3.653	<0.001
$C_{mono+TM}$	-0.496	-1.823	0.831	-0.734	0.463
Num. Obs	960				
R2 Nagelkerke	0.70				

Table 5: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the *peak luminance* factor. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{baseline}$.

	Estimates	2.5% CI	97.5% CI	t	p
$C_{middle:bright}$	2.747	1.728	3.766	5.293	<0.001
$C_{dark:bright}$	3.521	2.495	4.546	6.738	<0.001
Num. Obs	720				
R2 Nagelkerke	0.70				

Table 6: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the *peak luminance* factor. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is $C_{dark:bright}$.

	Estimates	2.5% CI	97.5% CI	t	p
$C_{baseline}$	-2.747	-3.766	-1.728	-5.293	<0.001
$C_{middle:bright}$	0.774	-0.173	1.720	1.605	0.109
Num. Obs	720				
R2 Nagelkerke	0.70				

Table 7: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the different tone mapping functions and the stereo condition. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is the *HDR* condition.

	Estimates	2.5% CI	97.5% CI	t	p
$TM_{high-steep}$	0.350	-0.632	1.332	1.698	0.485
$TM_{high-gradual}$	1.859	0.890	2.827	3.765	<0.001
$TM_{low-steep}$	2.641	1.627	3.600	5.196	<0.001
$TM_{low-gradual}$	4.838	3.797	5.878	9.122	<0.001
Stereoscopic	-0.027	-1.029	0.975	-0.052	0.958
Num. Obs	1200				
R2 Nagelkerke	0.78				

Table 8: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the different tone mapping functions and the stereo condition. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is the $TM_{high-steep}$ condition.

	Estimates	2.5% CI	97.5% CI	t	p
HDR	-0.350	-1.333	0.632	-0.699	0.485
$TM_{high-gradual}$	1.508	0.578	2.453	3.159	0.002
$TM_{low-steep}$	2.263	1.316	3.226	4.652	<0.001
$TM_{low-gradual}$	4.487	3.486	5.505	8.723	<0.001
Stereoscopic	-0.514	-1.469	0.439	-1.057	0.291
Num. Obs	1200				
R2 Nagelkerke	0.78				

Table 9: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the different tone mapping functions and the stereo condition. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is the $TM_{high-gradual}$ condition.

	Estimates	2.5% CI	97.5% CI	t	p
HDR	-1.860	-2.834	-0.897	-3.767	<0.001
$TM_{high-steep}$	-1.510	-2.453	-0.578	-3.162	0.002
$TM_{low-steep}$	0.754	-0.142	1.655	1.647	0.100
$TM_{low-gradual}$	2.978	2.034	3.932	6.157	<0.001
Stereoscopic	-0.635	-1.551	0.277	-1.364	0.173
Num. Obs	1200				
R2 Nagelkerke	0.78				

Table 10: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the different tone mapping functions and the stereo condition. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is the $TM_{low-steep}$ condition.

	Estimates	2.5% CI	97.5% CI	t	p
HDR	-2.613	-3.608	-1.634	-5.194	<0.001
$TM_{high-steep}$	-2.263	-3.226	-1.316	-4.652	<0.001
$TM_{high-gradual}$	-0.754	-1.655	0.142	-1.648	0.100
$TM_{low-gradual}$	2.225	1.283	3.171	4.625	<0.001
Stereoscopic	-1.657	-2.598	-0.733	-3.489	<0.001
Num. Obs	1200				
R2 Nagelkerke	0.78				

Table 11: Estimates of the regression coefficients, confidence intervals (2.5% and 97.5%), t-values (t), and p-values (p) for the ordinal logistic regression model modeling the effect of the different tone mapping functions and the stereo condition. Nagelkerke pseudo-R2 indicates the goodness-of-fit, i.e., how well the model explains the data. The baseline for this model is the $TM_{low-gradual}$ condition.

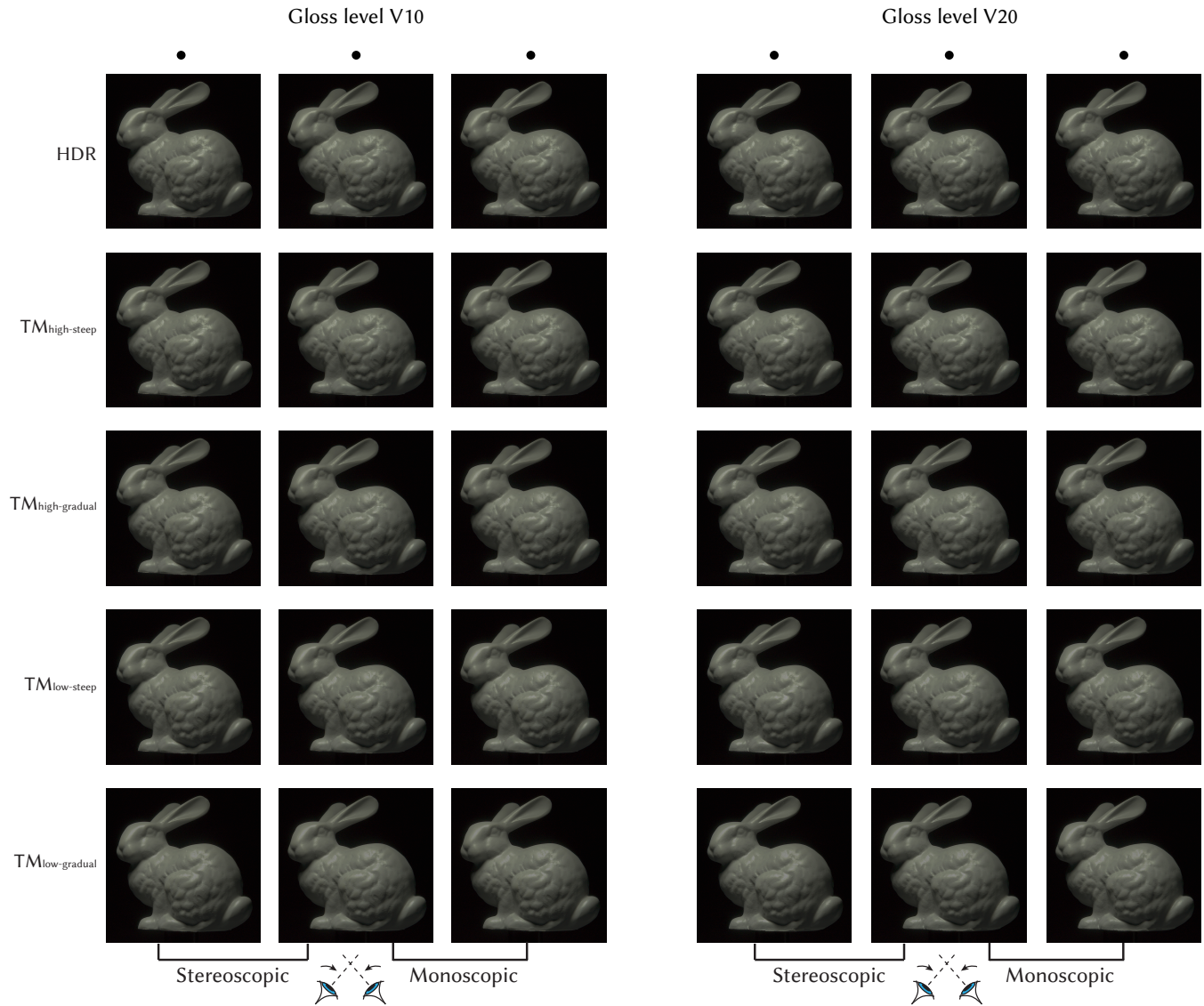|	Estimates	2.5% CI	97.5% CI	t	p
HDR	-4.837	-5.885	-3.805	-9.122	<0.001
$TM_{high-steep}$	-4.488	-5.505	-3.486	-8.724	<0.001
$TM_{high-gradual}$	-2.979	-3.932	-2.034	-6.159	<0.001
$TM_{low-steep}$	-2.224	-3.171	-1.283	-4.624	<0.001
Stereoscopic	-1.555	-2.576	-0.536	-2.990	0.003
Num. Obs	1200				
R2 Nagelkerke	0.78				

Figure 8: Stereoscopic stimuli demonstration. Note that all images are gamma corrected by $\gamma = \frac{1}{3.6}$ for better visualization on low dynamic range display.

8 ALL THE STIMULI USED IN MAIN EXPERIMENT (SEC. 4) AND TONE MAPPING EXPERIMENT (SEC. 6)

The stimuli used in our study are fabricated by dipping 3D printed objects in a varnish mixture. The mixture consists of a proportion of glossy and matte varnishes. For simplicity we refer to the samples by the percentage of glossy varnish. We manufactured stimuli with 10% and 20% gloss (Figure 8), 30% and 40% gloss (Figure 9), 50% and 60% gloss (Figure 10), 70% and 80% gloss (Figure 11), and 90% and 100% gloss (Figure 12).

REFERENCES

Peter D Burns and Don Williams. 2018. Camera resolution and distortion: Advanced edge fitting. *Electronic Imaging* 30 (2018), 1–5.

Aliaksei Mikhailiuk, Clifford Wilmot, Maria Perez-Ortiz, Dingcheng Yue, and Rafał K Mantiuk. 2021. Active sampling for pairwise comparisons via approximate message passing and information gain maximization. In *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2559–2566.

Helge Seetzen, Wolfgang Heidrich, Wolfgang Stuerzlinger, Greg Ward, Lorne Whitehead, Matthew Trentacoste, Abhijeet Ghosh, and Andrejs Vorozcovs. 2004. High dynamic range display systems. In *ACM SIGGRAPH 2004 Papers*. 760–768.

Eino-Ville Talvala, Andrew Adams, Mark Horowitz, and Marc Levoy. 2007. Veiling glare in high dynamic range imaging. *ACM Transactions on Graphics (TOG)* 26, 3 (2007), 37–es.

Sophie Wuerger, Maliha Ashraf, Minjung Kim, Jasna Martinovic, María Pérez-Ortiz, and Rafał K Mantiuk. 2020. Spatio-chromatic contrast sensitivity under mesopic and photopic light levels. *Journal of Vision* 20, 4 (2020), 23–23.

Fangcheng Zhong, Akshay Jindal, Özgür Yöntem, Param Hanji, Simon Watt, and Rafal Mantiuk. 2021. Reproducing reality with a high-dynamic-range multi-focal stereo display. *ACM Trans. Graph.* 40, 6 (2021), 241.
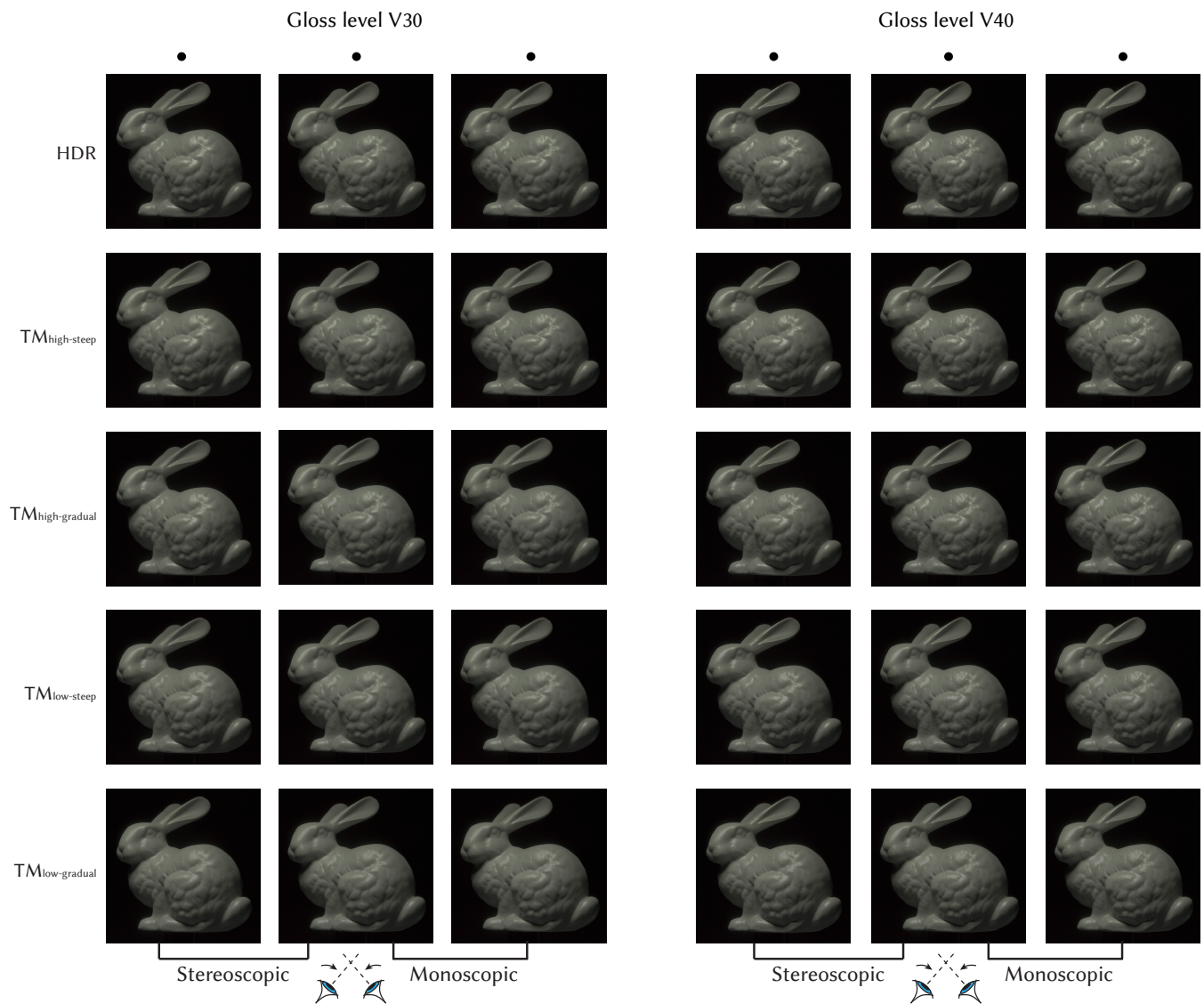
Figure 9: Stereoscopic stimuli demonstration. Note that all images are gamma corrected by $\gamma = \frac{1}{3.6}$ for better visualization on low dynamic range display.
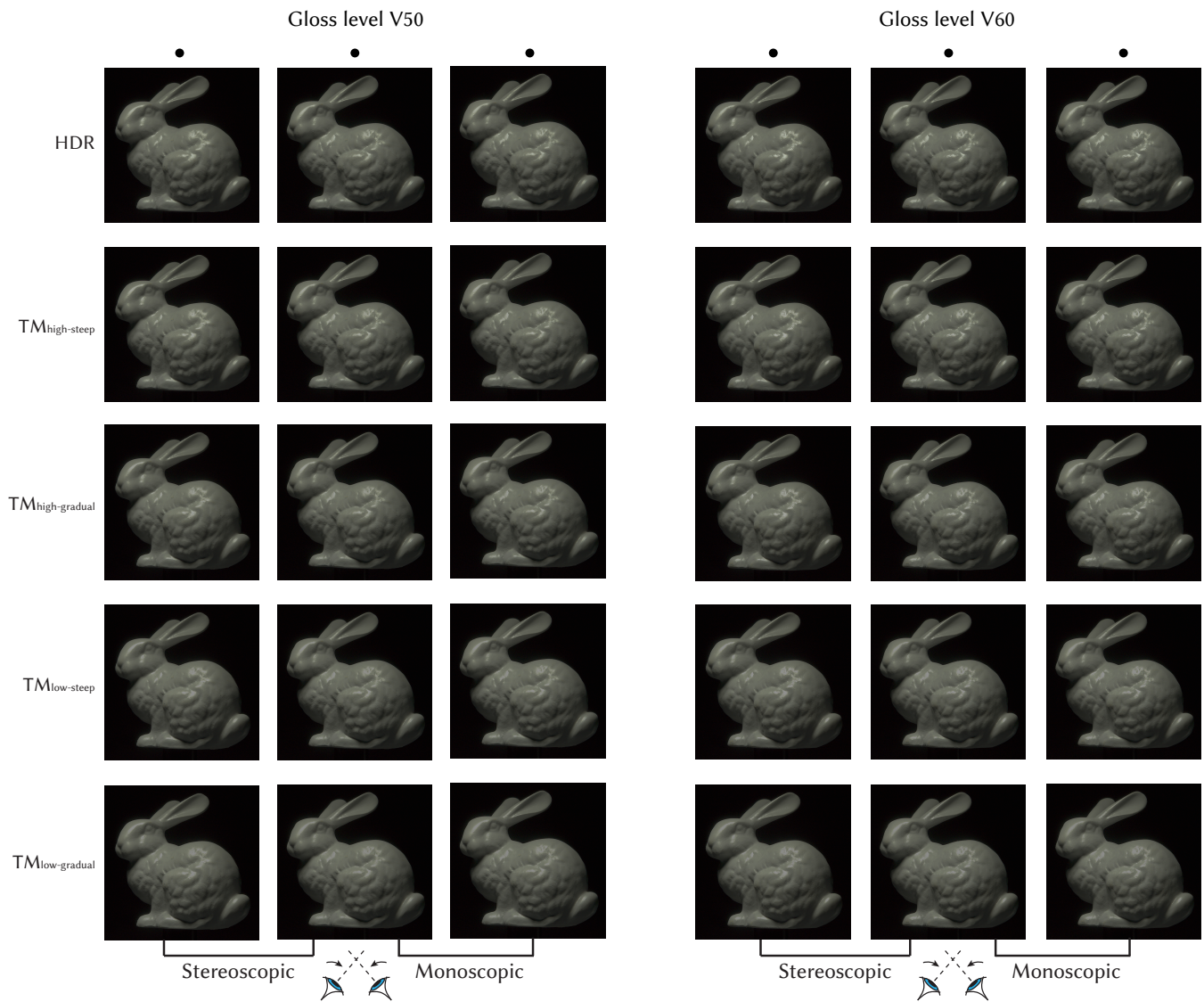
Figure 10: Stereoscopic stimuli demonstration. Note that all images are gamma corrected by $\gamma = \frac{1}{3.6}$ for better visualization on low dynamic range display.
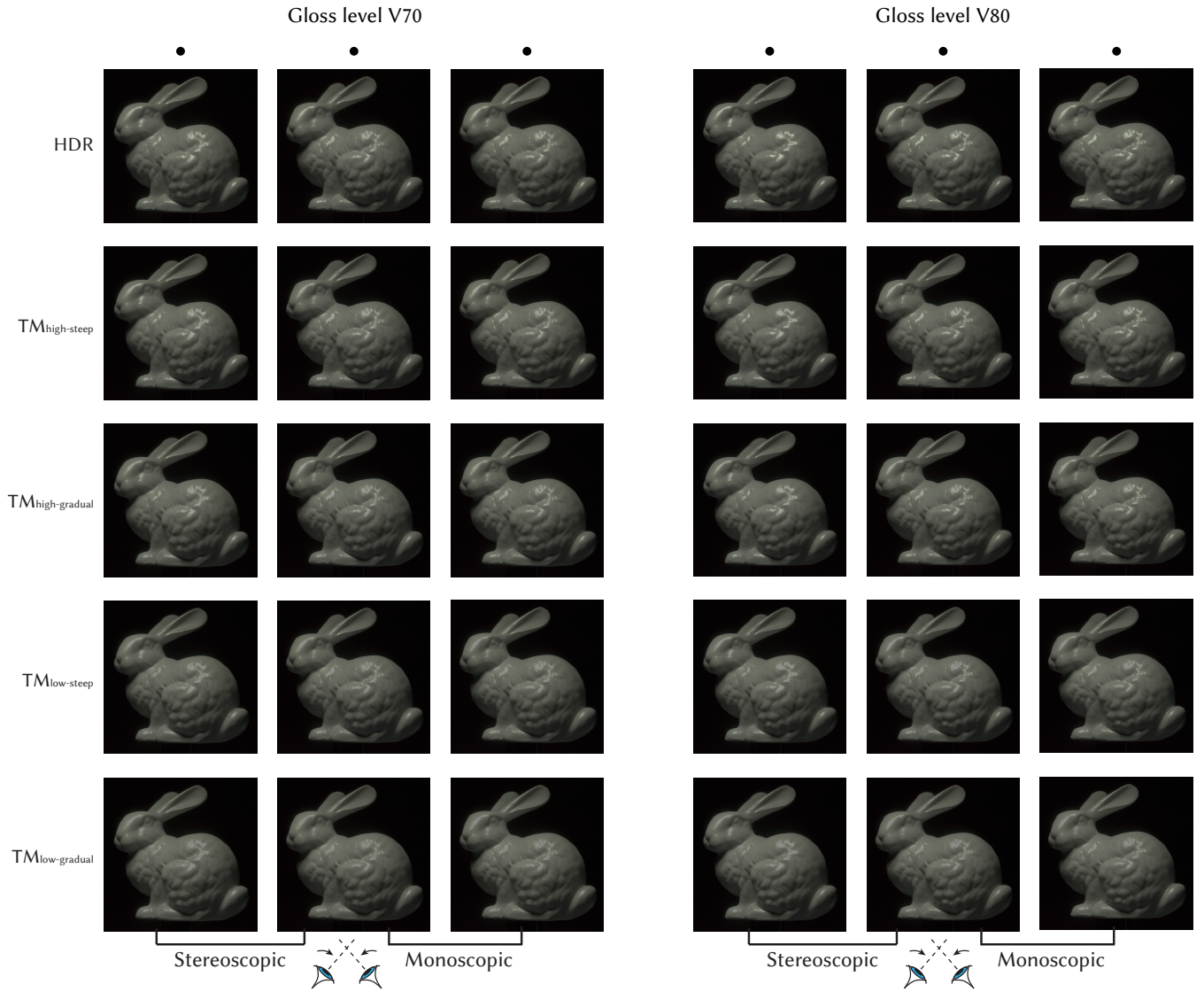
Figure 11: Stereoscopic stimuli demonstration. Note that all images are gamma corrected by $\gamma = \frac{1}{3.6}$ for better visualization on low dynamic range display.
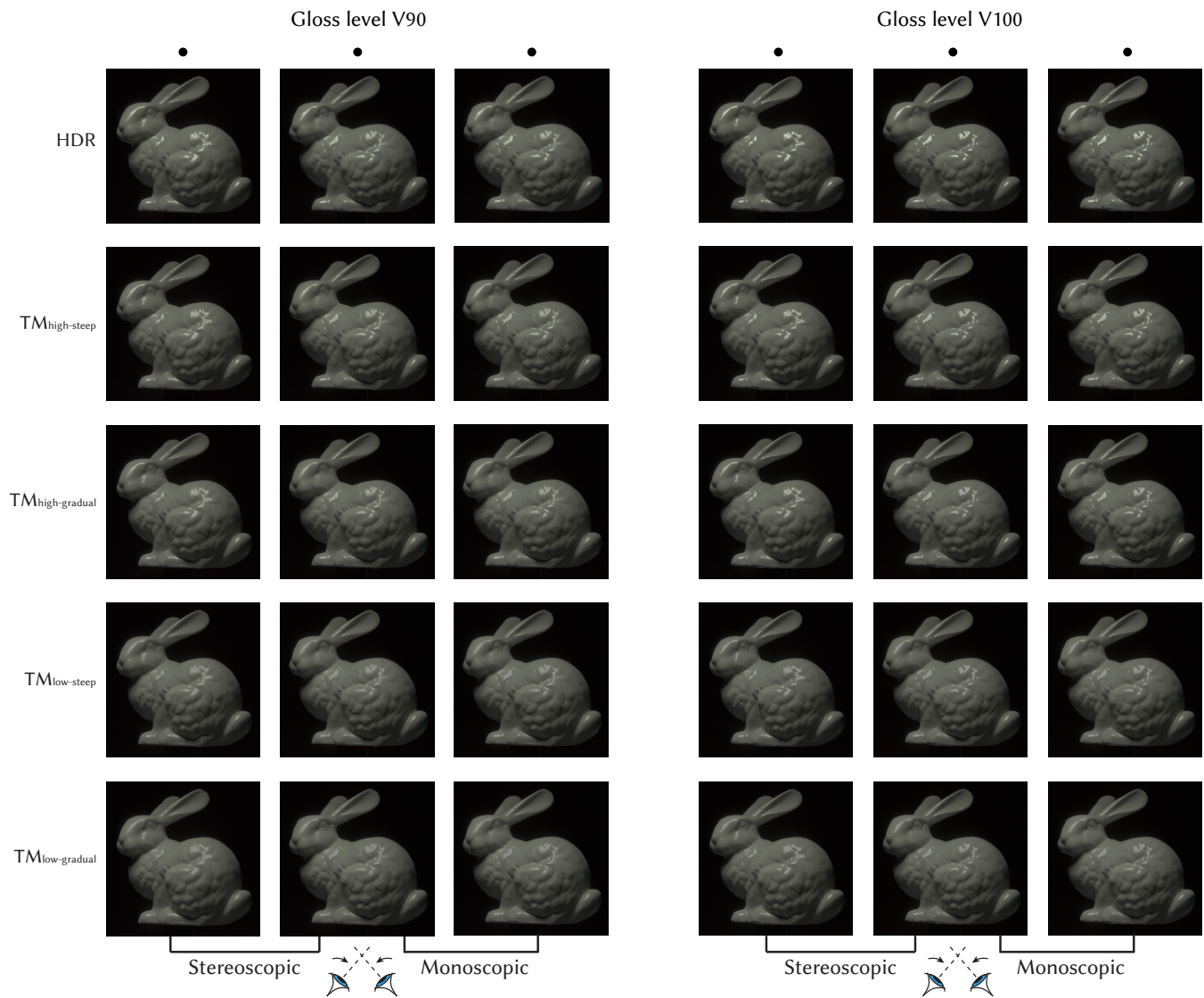
Figure 12: Stereoscopic stimuli demonstration. Note that all images are gamma corrected by $\gamma = \frac{1}{3.6}$ for better visualization on low dynamic range display.