Perception, 2014, volume 43, pages 368–380

doi:10.1068/p7616

Human discrimination of depth of field in stereoscopic and nonstereoscopic photographs

Tingting Zhang[1]§, Harold T Nefs[1], Ingrid Heynderickx[1,2]
[1] Interactive Intelligence Group, Department Intelligent System, Faculty of EEMCS, Delft University of Technology; [2] Department Human Technology Interaction, Faculty IE&IS, Eindhoven University of Technology; e-mail: t.zhang@tudelft.nl
Received 25 September 2013, in revised form 3 April 2014

Abstract. Depth of field (DOF) is defined as the distance range within which objects are perceived as sharp. Previous research has focused on blur discrimination in artificial stimuli and natural photographs. The discrimination of DOF, however, has received less attention. Since DOF introduces blur which is related to distance in depth, many levels of blur are simultaneously present. As a consequence, it is unclear whether discrimination thresholds for blur are appropriate for predicting discrimination thresholds for DOF. We therefore measured discrimination thresholds for DOF using a two-alternative forced-choice task. Ten participants were asked to observe two images and to select the one with the larger DOF. We manipulated the scale of the scene—that is, the actual depth in the scene. We conducted the experiment under stereoscopic and nonstereoscopic viewing conditions. We found that the threshold for a large DOF (39.1 mm) was higher than for a small DOF (10.1 mm), and the thresholds decreased when scale of scene increased. We also found that there was no significant difference between stereoscopic and nonstereoscopic conditions. We compared our results with thresholds predicted from the literature. We concluded that using blur discrimination thresholds to discriminate DOF may lead to erroneous conclusions because the depth in the scene significantly affects people's DOF discrimination ability.

Keywords: depth of field, discrimination, stereo photographs

1 Introduction

Depth of field (DOF) is the distance range within which objects are perceived as sharp. Objects that are outside of the DOF will appear blurred in an image. Figure 1 shows an example of small and large DOFs. DOF has various applications in enhancing the subjective quality of images. For example, firstly, it may be used to enhance depth perception in photographs (Marshall, Burbeck, Ariely, Rolland, & Martin, 1996; Pentland, 1987; Watt, Akeley, Ernst, & Banks, 2005). Secondly, it has been shown to contribute to the aesthetic appreciation of

Figure 1. [In color online, see http://dx.doi.org/10.1068/p7616] Depth of field effects: left: small depth of field; right: large depth of field.

§All correspondence should be addressed to Faculty of EEMCS, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.

photographs (Datta, Joshi, Li, & Wang, 2006), and to make images appear more natural and realistic (Joshi et al., 2011). Thirdly, DOF is believed to be closely related to visual attention—the focal point of the image can be highlighted by blurring the remainder, thus drawing viewers' attention to specific positions in the photograph (Cole et al., 2006; Steve, Caitlin, & James, 2010). To better understand the aesthetic and attention effects, it would be good to know the differences in DOF that can be perceived by the average viewer and whether they can be predicted from blur discrimination.

Because DOF is perceived as a change in blur in an image, it seems plausible that perceived differences in DOF are related to perceived differences in blur. Human blur detection and discrimination have been investigated extensively in the last few decades. For example, Hamerly and Dvorak (1981) investigated edge and line blur discrimination and found that observers could discriminate a blurred from a sharp high-contrast photograph when the edge-transition width was above 25 arcsec. Mather and Smith (2002) conducted an experiment to investigate blur discrimination of three kinds of blur: luminance border, texture border, and region blur. The results showed that the increment threshold of blur first decreased and then increased with increasing levels of blur in the reference blur circle, resulting in a parabolic shape of the relationship between the threshold and the reference blur with a peak sensitivity around one arcmin. Consistency in these results was shown across a variety of studies in spite of different stimuli and experimental methods (Hess, Pointer, & Watt, 1989; Mather, 1997; Mather & Smith, 2002; Watt & Morgan, 1983; Wuerger, Owens, & Westland, 2001). Assuming a peak ability to discriminate blur at about one arcmin, we may predict that this value is the limiting factor in discriminating DOF. If the image contains only regions with larger or smaller blur circles, the threshold will be larger than when the image contains blur circles around one arcmin.

There are, however, basic differences between blurred images and images with a limited DOF; in the latter case the level of blur is not homogeneously distributed over the whole image, but depends on the local distance of the imaged object with respect to the focal plane. DOF is generated in photographs as a result of optics of the imaging equipment, most often manipulated by varying the aperture size of the camera. In addition, most previous studies on blur discrimination used single blurred edges (Georgeson, 1998; Hamerly & Dvorak, 1981; Pääkkönen & Morgan, 1994), binary texture (Hoffman & Banks, 2010), or random dot stereograms (Mather & Smith, 2002), rendered by computer algorithms. In contrast, our stimuli contained a blur gradient over the figures in the scene that was affected by the scale of the scene. Even though the peak sensitivity is at a blur circle of one arcmin, it is possible that people would still benefit from the presence of blur at other (suboptimal) levels to discriminate DOF.

Stereoscopic and nonstereoscopic images are perceived differently in a number of important ways. Firstly, the optical state of the eyes may be different for stereoscopic photographs than for nonstereoscopic photographs because of the tight link between convergence and accommodation (Hoffman, Girshick, Akeley, & Banks, 2008; Otero, 1951). For stereoscopic viewing conditions the image on the retina may thus be more blurred because of the incorrect accommodation based on convergence rather than on the distance to the image plane. Perceived DOF may be influenced by the optical state of the eyes (Campbell, 1957), and therefore the discrimination in DOF may be different for stereoscopic and nonstereoscopic photographs. Secondly, the subjective experience of depth is qualitatively different in stereoscopic compared with nonstereoscopic photographs. In nonstereoscopic photographs pictorial space does not appear to occupy the same physical space as in stereoscopic images (Rogers, 1995). Further, stereoscopic images provide more depth cues than nonstereoscopic images, which could in principle be used to gain more complex information (Liu, Hua, & Cheng, 2010). It was found that more detail could be perceived in stereoscopic images than in nonstereoscopic images (Heynderickx & Kaptein, 2009). Therefore, there are more 'chances' to see differences in blur in these details. We may thus hypothesize that it may be easier to see the difference of DOF in stereoscopic images than in nonstereoscopic images.

In the current study we measured the just noticeable difference (JND) of two DOFs. To get more reliable results, we used two sets of photographs of similar scenes. Additionally, we also adjusted the absolute level of depth in the photographed scene, which directly influenced the blur gradient in the photographs. The JND for DOF was measured using both stereoscopic and nonstereoscopic photographs.

2 Experiment

2.1 *Methods*

2.1.1 *Participants.* Four female and six male observers, aged between 25 and 37 years old, with normal or corrected-to-normal visual acuity as measured with the Freiburg visual acuity test (Bach, 1996) and normal stereo acuity as measured with the TNO stereo test (Laméris Ootech BV), participated in our experiment. Informed consent forms were obtained from all participants. This research was approved by the Delft University of Technology and conducted according to the Declaration of Helsinki, Dutch Law, and common local ethical practice.

2.1.2 *Apparatus and stimuli.* A Wheatstone (1838) stereoscope with two 19" Iiyama CRT monitors (type MM904UT) and front-surface silver-plated mirrors was used in the experiment. The two monitors were set to a screen resolution of 1280 × 800 pixels and calibrated with a ColorMunki®, such that their luminance and color responses were identical. Figure 2 shows the diagram of the stereoscope. The path length from the eyes to the screen was 70 cm. The mirrors were orientated so that the convergence angle of the eyes was congruent with a viewing distance of 70 cm. Stimuli used in the experiment were generated with an Olympus E-330 d-SLR camera with a 50-mm Olympus Zuiko macro lens. The aperture of the camera lens could be set from F2.0 (the smallest DOF) to F22 (the largest DOF). The angle of view of the camera was 13.2 deg (horizontally) × 9.9 deg (vertically). The size of the stimuli displayed on the screens was constrained by the visual angle of the camera.

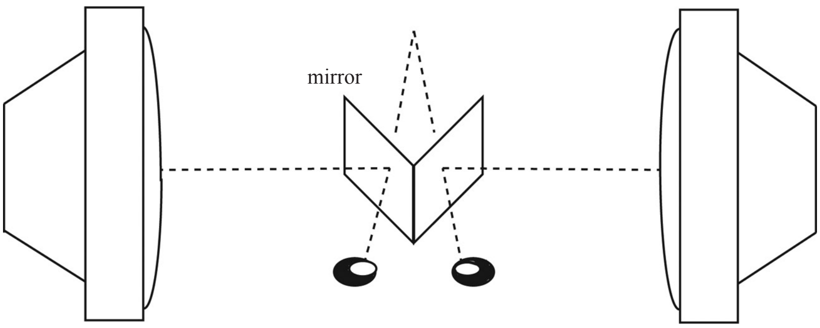

Figure 2. The stereoscope used in the experiment.

Figure 3 shows the stimuli for which the JND values were measured. The stimuli contained two scenes. The compositions of the two scenes were quite similar to each other: each consisted of six different objects standing at regular intervals on a white ground. The foremost object in the scene was always in focus, while the objects behind were gradually blurred depending on their distances to the front object and on the DOF of the camera lens. Each scene was named after its focal object: the 'Apple' scene and the 'Woody' scene, as can be seen in figure 3. The distance between the real-world objects, and so the physical depth structure in the scene, was also manipulated; this factor is referred to as 'scale of the scene' in the remainder of the text and was expressed as the maximum depth between the focal central object and the farthest background object in the scene. Three values were selected for this

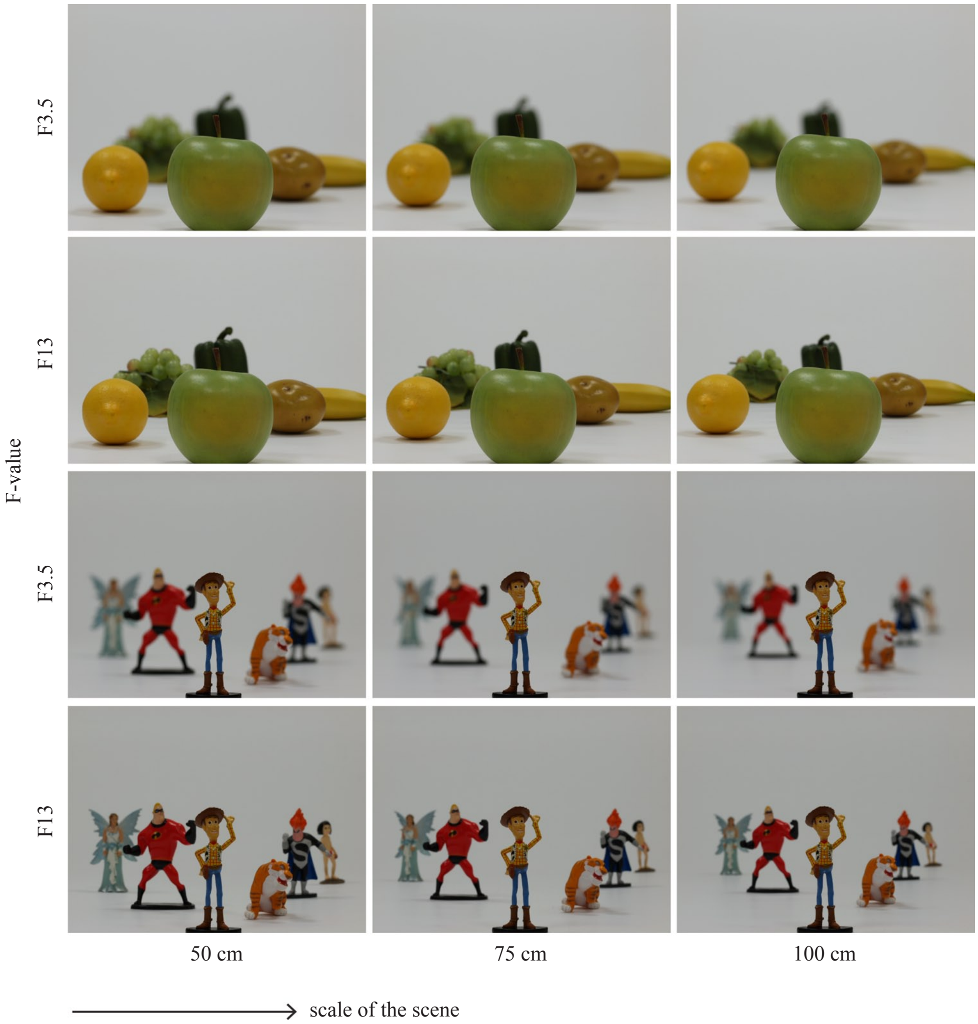

Figure 3. [In color online.] Stimuli used in the experiment.

scale of the scene factor—namely 50 cm, 75 cm, and 100 cm. Thus, the set of photographs that contained two scenes and three levels of scale of the scene were photographed with two different camera apertures—namely F3.5 and F13—which acted as reference values for the DOF. To measure the JND in DOF, we created ten additional pictures with the aperture of the camera being F2, F2.2, F2.5, F2.8, F3.2, F4, F4.5, F5, F5.6, and F6.3 for the reference DOF of F3.5, and with the aperture being F7.1, F8, F9, F10, F11, F14, F16, F18, F20, and F22 for the reference DOF of F13.

For the stereoscopic viewing conditions the left and right half images were taken sequentially using a metal slide bar. The distance over which the camera was displaced was called the stereo base, and it was 6.5 cm in our study. For the nonstereoscopic viewing conditions we set the camera in the middle of the slide bar. When taking the photographs, the camera was centered on the central figure. In our experiment the orientation of the mirrors of the stereoscope was set such that the distance to the virtual figurine specified by convergence was the same as the accommodation-defined distance to the screen. This calibration ensured that the two half images could be fused properly. A reference and a test image were always presented side-by-side (counterbalanced order) on the screens of the stereoscope. The angle of view of the photographs was the same as the angle of view of the camera. The figures were thus shown life-sized.

2.1.3 *Procedure.* The experiment was based on a within-subject design for the three independent variables, which were reference DOF, scale of the scene, and stereoscopic versus nonstereoscopic images. The observers were seated in a dark room in front of the stereoscope mirrors with the only direct light coming from the two monitors. The experiment was based on a two-alternative forced-choice (2AFC) procedure. On each trial a reference image and a test image were displayed simultaneously side-by-side in the middle of both monitors.

Observers were asked to decide which image appeared to have the larger DOF and then press the corresponding left or right arrow on the computer keyboard. Then, the next trial was presented automatically. The response time for each comparison was in principle unlimited, so the participants could take as long as they needed. The participants evaluated 240 trials per session [ie 2 scenes × 3 levels of scales of the scene × 2 reference DOFs × 10 comparisons presented twice (the reference was once on the left and once on the right half of the monitor)]. The full experiment consisted of 30 sessions, of which half used nonstereoscopic images and half stereoscopic images (and so we had 15 repetitions per viewing condition). The sessions with nonstereoscopic images alternated with the sessions with stereoscopic images. For each participant, the starting session (ie stereoscopic or nonstereoscopic) and the order of the comparisons within a session was random.

2.2 *Analysis*

DOF can be described in different ways such as by diopters, F-value, aperture size, distance range, or the diameter of the blur circle. The relationship between F-value and aperture size is described in equation (1), where L is the focal length of the camera and a is the aperture size. In our work the distance range in mm within which the blur circle is smaller than one arcmin is used as the value of DOF (Born & Wolf, 1999). Compared with diopters, F-value, or blur circle, distance is a visualized and intuitive parameter and easier to understand. The relationship between F-value and the distance range is shown in equation (2), with β indicating the angular size of the blur circle and D indicating the distance from the focal object to the lens. Figure 4 shows the geometrical relationship between aperture size, focus distance, and blur circle. Table 1 summarizes the values of DOF and the values of the aperture size in mm corresponding to all the F-values.

$$\mathrm{F}_{\text{value}} = \frac{L}{a}\,, \tag{1}$$

$$\delta = \frac{2D^2 \tan(\beta/2)}{(L/\mathrm{F}_{\text{value}}) - 2D\tan(\beta/2)}\,. \tag{2}$$

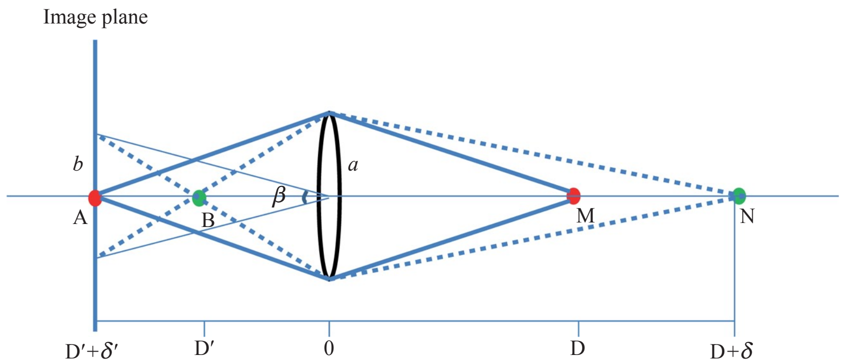

Figure 4. [In color online.] A lens at position 0 focuses on an object at position D. Object N is out of focus, and so its image is a disk in the image plane. The diameter b of the image of object N is defined as the blur circle. The angular size of the blur circle as seen from the center of the lens is indicated with β.

Table 1. The value of depth of field (DOF) and the aperture size in mm corresponding to the F-value of the camera lens.

	Reference	Test									
F-value	3.5	2	2.2	2.5	2.8	3.2	4	4.5	5	5.6	6.3
Aperture/mm	14.3	25	22.7	20	17.9	15.6	12.5	11.1	10	8.9	7.9
DOF/mm	10.1	5.7	6.3	7.2	8.1	9.2	11.6	13.1	14.5	16.3	18.4
F-value	13	7.1	8	9	10	11	14	16	18	20	22
Aperture/mm	3.8	7.0	6.3	5.6	5	4.5	3.6	3.1	2.8	2.5	2.3
DOF/mm	39.1	20.8	23.6	26.6	29.7	32.8	42.3	48.8	55.4	62.1	68.9

The proportion of trials where the participant chose the larger DOF from the combination of reference and test stimulus was fitted using a Gaussian cumulative function. The difference between the DOF at the point of subjective equality (probability of saying 'larger' = 0.5) and at a 0.75 probability of responding 'larger' was defined as the increment threshold (JND) of the reference DOF.

3 Results

We found that the JNDs for a DOF of 10.1 mm (ie F3.5) across all conditions and all participants ranged between 0.14 mm and 4.17 mm. For a reference DOF value of 39.1 mm (ie F13), the JNDs ranged between 0.6 mm and 57.16 mm. The data thus showed large individual differences, indicating that some people were sensitive to changes in DOF, whereas others could not really discriminate DOF well.

The JND values averaged across all ten participants are summarized in figure 5. Figure 5(a) shows that the JND for a reference DOF of 39.1 mm was much larger than the JND for a reference DOF of 10.1 mm. There was no big difference in JND between the 'Apple' scene and the 'Woody' scene in figure 5(b), while the JND in DOF was found to decrease with increasing scale of the scene, as shown in figure 5(c). Figure 5(d) demonstrates the discrimination thresholds observed under nonstereoscopic and stereoscopic viewing conditions.

We performed a 2 (reference DOFs) × 2 [replications (scene)] × 3 (scales of the scene) × 2 (viewing conditions) repeated-measures ANOVA. We found significant main effects of reference DOF ($F_{1,9} = 15.54$, $p < 0.003$) and scale of the scene ($F_{2,18} = 7.81$, $p < 0.004$). Additionally, a significant interaction between reference DOF and scale of the scene was found ($F_{2,18} = 7.52$, $p < 0.004$). Figure 5(e) shows that the change of the JND in DOF with Scale of the scene is larger when using a larger rather than a smaller reference DOF.

4 Modeling

4.1 *Predicting JNDs of DOF from blur discrimination*

We predicted the values of JNDs from blur discrimination studies in literature and compared them with our experimental results. Blur discrimination studies, however, typically have used only one level of blur in the stimulus—that is, the blur is uniform across the image. On the other hand, more levels of blur are available in photographs with limited DOFs. Therefore, our first step was to select a level of blur circle from our stimuli as the reference blur circle. Two different values for this blur circle were used: the minimum blur circle and the blur circle of one arcmin.

Since there was no other object between the focus object and the second object, other than a completely white ground floor, it was difficult to observe the blur circle located between the focus object and the second object. Therefore, the blur circle on the second object was

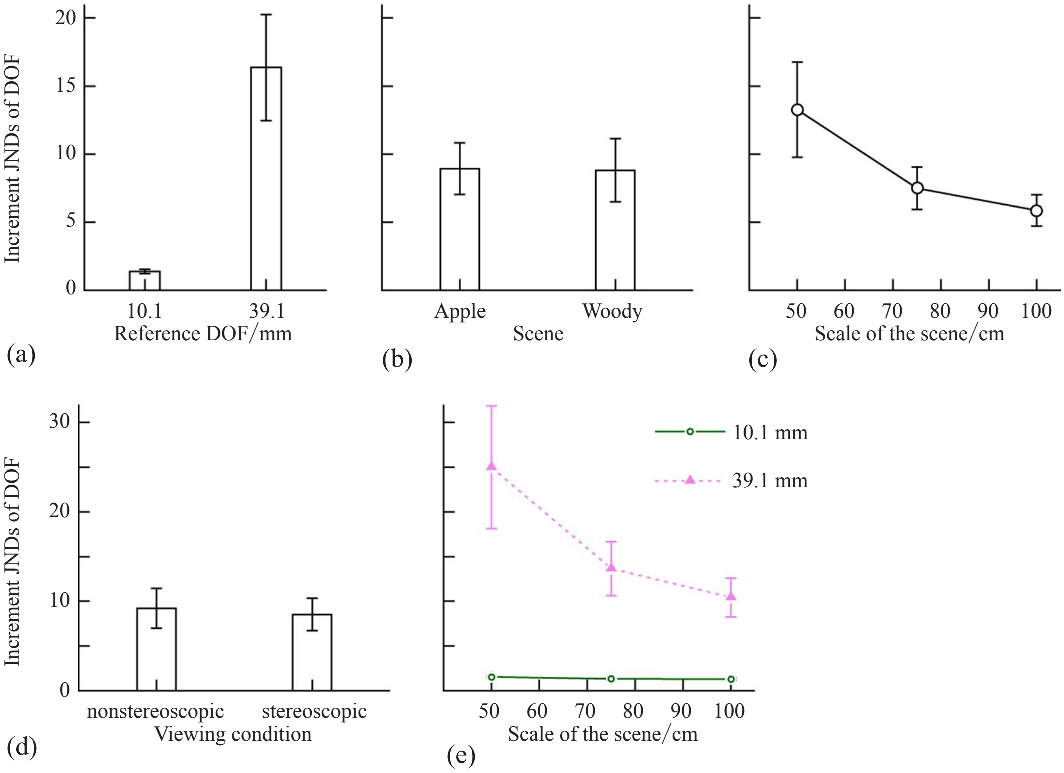

Figure 5. [In color online.] The averaged increment just noticeable differences (JNDs) across participants with the error bars represent ±1 standard error of the mean value. (a) JNDs in photographs with two reference depths of field (DOFs): 10.1 mm and 39.1 mm; (b) JNDs in photographs with different content: Apple and Woody; (c) JNDs in photographs with different scale of the scene: 50 cm, 75 cm, and 100 cm; (d) JNDs in photographs under different viewing conditions: nonstereoscopic and stereoscopic; (e) interaction between scale of the scene and viewing condition.

regarded as the minimum visible blur circle in the stimuli. Equation (3) was used to calculate this minimum blur circle b, with L being the focal length, d the depth between the second object and the focus object, D the focus distance, and F-value the aperture representing the reference DOF. In our experiment, d could be 10 cm, 15 cm, or 20 cm, depending on the scale of the scene:

$$\tan(b) = \frac{Ld}{D(D+d)\mathrm{F}_{\text{value}}}\,. \tag{3}$$

Although a blur circle of one arcmin was situated somewhere between the focal object and the second object on the white ground, participants may find it difficult to observe this blur circle. We nonetheless selected this value as a reference for two reasons. First, the definition of DOF in our paper was based on the blur circle of one arcmin. Second, the peak sensitivity for blur discrimination was found to be around one arcmin (Chen, Chen, Tseng, Kuo, & Wu, 2009; Hamerly & Dvorak, 1981; Hess, Pointer, Simmers, & Bex, 2003; Mather & Smith, 2002; Watt & Morgan, 1983).

The second step of the prediction was to use the reference blur circle b to calculate the JNDs. Watson and Ahumada (2011) summarized the previous studies in blur discrimination, and combined their data to build a universal model for the blur discrimination threshold. They assumed that a larger blur circle b_1 could be discriminated from a smaller blur circle b_2 when b_1 was a factor ω multiplied by b_2, raised to a power ρ. The resulting Weber model assumed that the blur discrimination threshold was determined by the total blur in the stimuli and the Weber fraction for blur discrimination. The total blur contained extrinsic blur and intrinsic blur. Extrinsic blur represented the image blur, and intrinsic blur represented the blur caused by the visual system. In our prediction the extrinsic blur was given by the blur circle values that we selected as reference blur (ie the minimum blur circle and the blur circle of one arcmin), while the intrinsic blur was obtained from literature. Finally, the equation for the blur discrimination threshold is shown as follows:

$$a = -r + \sqrt{\omega^2(\beta^2 + r)^\rho - \beta^2}, \tag{4}$$

with a the increment discrimination threshold of blur, r the extrinsic blur (ie the reference blur), ω the Weber fraction, β the intrinsic blur, and ρ the Weber exponent. These parameters varied across studies. An overview of the values of the parameters is given in table 2, taken from the paper of Watson and Ahumada (2011). Using equation (4) and the parameters in table 2, we could calculate the JNDs for DOF. Equation (3) was used to transform the JNDs for blur to the JNDs for DOF in mm.

Table 2. Weber model parameters for four studies and root-mean-square (RMS) error for the four studies. RMS values are in units of ln arcmin.

Study	β	ω	ρ	RMS
Chen et al. (2009)	1.20	1.18	1.05	0.043
Hess, Pointer, and Watt (1989)	1.71	1.06	1.03	0.034
Mather and Smith (2002)	1.54	1.13	1.04	0.040
Watson and Ahumada (2011)	1	1.15	1.02	0.055

Figure 6 shows the predicted DOF JND based on the blur JND for one arcmin blur circle, calculated from the data in table 2 and compared with our experimental data, taking into account one standard error of the mean. In order to compare our measured data with the predicted data from the model, we performed one-sample t-tests. Note that we did not take the variance of the predicted data into account as they were not available to us; but, strictly speaking, this may lead to some extra type-I error in the analysis. One-sample t-tests [comparing the mean predicted value for each model and each reference DOF with the experimental JND value

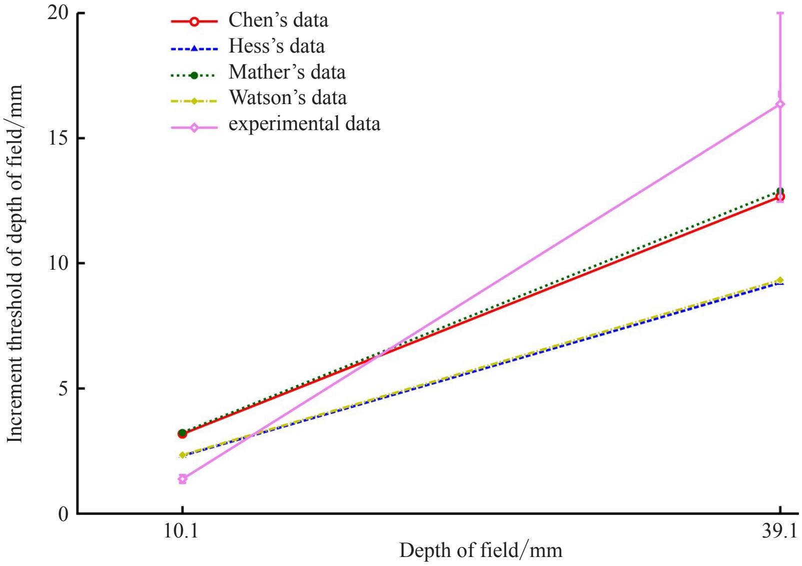

Figure 6. [In color online.] Comparing our measured just noticeable differences (JNDs) with the predicted JNDs from literature.

($N = 10$)] showed that for a reference DOF of 10.1 mm our experimentally determined JND in DOF was significantly smaller than what was predicted from the blur JND, independent of which dataset was used ($t_9 = -11.3$, $p < 0.001$; $t_9 = -5.8$, $p < 0.001$; $t_9 = -11.7$, $p < 0.001$; $t_9 = -6.0$, $p < 0.001$ for Chen's, Hess's, Mather's, and Watson's data, respectively). For a reference DOF of 39.1 mm, we found no significant difference between the experimentally determined JND in DOF and the predicted ones.

The predicted JNDs from Chen's, Hess's, Mather's, and Watson's data were quite consistent; therefore, we averaged the predicted JNDs. Figure 7 shows the mean predicted JNDs from the literature and our experimental data with ±1 standard error. The minimum blur circle on the second object in the scene varied with the value of the scale of the scene. We predicted the JND in DOF separately for the various scale of the scene values. Again, one-sample *t*-test analyses were performed, and we found that for a reference DOF of 10.1 mm the experimentally determined JND was significantly smaller than the predicted value, independent from the value of the scale of scene. For a reference DOF of 39.1 mm, we found something different. When the scale of the scene was 50 cm, there was no significant difference between the predicted JND and the experimental JNDs. However, when the scale of the scene value was 75 cm or 100 cm, the experimental JNDs were significantly smaller than the predicted JNDs. The results of the *t*-test analyses are summarized in table 3.

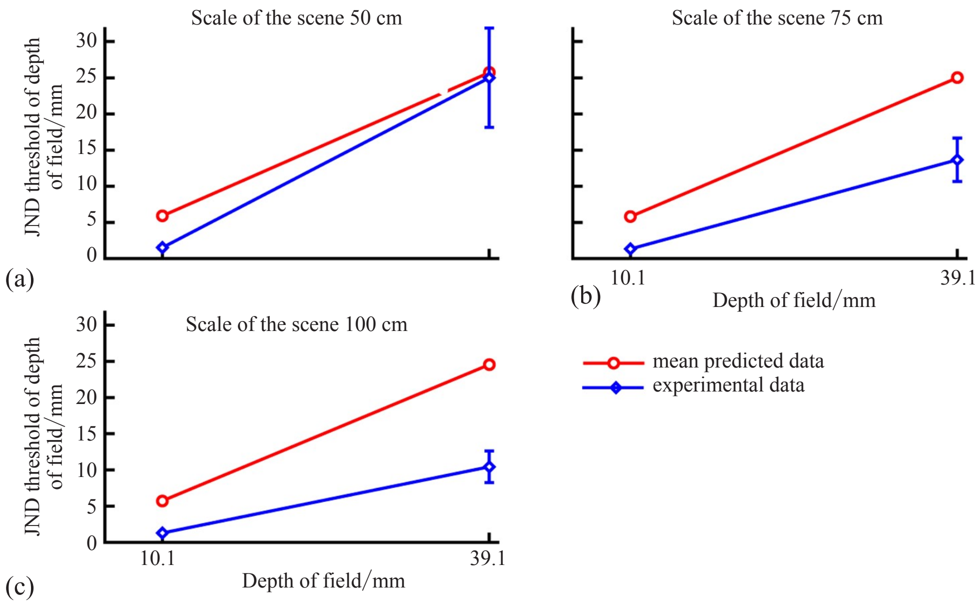

Figure 7. [In color online.] Comparison between the experimental JNDs (with ±1 standard error) and the mean predicted JNDs from the literature.

Table 3. *t*-values from one-sample *t*-test, comparing our experimentally measured just noticeable differences (JNDs) and the mean predicted JNDs.

Studies	*t*-value					
	10.1 mm (reference DOF)			39.1 mm (reference DOF)		
	50 cm	75 cm	100 cm	50 cm	75 cm	100 cm
Predicted JNDs	–24.89***	–27.62***	–29.51***	–	–3.77**	–6.49***

** $p < 0.05$; *** $p < 0.001$.
Note: DOF = depth of field.

4.2 *Fourier analysis*

In order to reveal the extent to which the power spectrum of images changed as the DOF blur changes, and to allow us to get a better insight into the similarities and differences between DOF blur and homogenous blur, we conducted a Fourier analysis. Further, we considered in this section how the visibility of differences in DOF related to differences in the power spectrum, taking into account the contrast sensitivity function.

The 'Apple' scene with the maximum depth of 75 cm was used as an example to show the results of Fourier analysis in figure 8. However, the analyses for the other images were similar. Figure 8(a) shows the changes in the power spectrum as a function of spatial frequency in the stimulus for a DOF of 10.1 mm and 39.1 mm, and also for Gaussian blur. In the latter case, a low and a high level of Gaussian blur were added to the sharpest photo in our experiment. The differences in the power spectrum between images with DOFs of 10.1 mm and 39.1 mm were similar to the differences between the low and high Gaussian blur levels. This might suggest that the DOF blur is in practice similar to uniform Gaussian blur, indicating that it may be possible to use blur discrimination thresholds to predict DOF discrimination.

The changes in contrast as a function of spatial frequency are shown in figure 8(b) together with the contrast sensitivity function (CSF) (Watson & Ahumada, 2011). The contrast difference between DOFs of 10.1 mm and 39.1 mm was above the CSF in the low-frequency area, indicating that the difference between the two DOFs should be visible. Similarly, the contrast difference between DOF of 10.1 mm and three of its test depths of field are presented

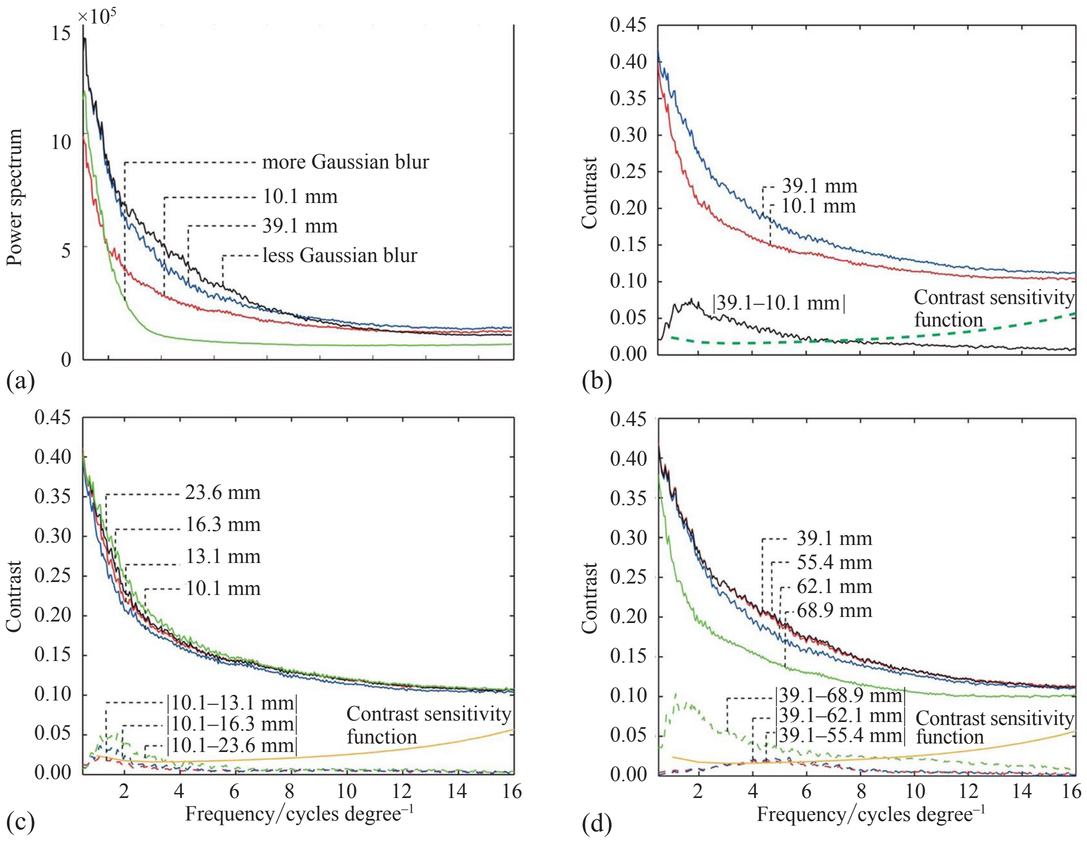

Figure 8. [In color online.] Fourier analysis on the 'Apple' scene with the maximum depth of 75 cm. (a) Power spectrum as a function of frequency; (b) contrast as a function of frequency for the reference depth of field, and the contrast difference between the reference depths of field; (c) the contrast difference between reference depth of field 10.1 mm and test depth of fields of 13.1 mm, 16.3 mm, and 23.6 mm; (d) the contrast difference between reference depths of field of 39.1 mm and test depths of field of 55.4 mm, 62.1 mm, and 68.9 mm.

in figure 8(c), and DOF of 39.1 mm with three of its test depths of field in 8(d). Figure 8(c) shows that only the difference between DOFs of 10.1 mm and 13.1 mm was below the CSF, suggesting that the difference between the two DOFs should not be visible, which was not in agreement with our experimental data. Figure 8(d) shows that the difference between DOFs of 39.1 mm and 62.1 mm was just below the CSF, which may indicate that we are possibly not able to discriminate them. However, our data showed that people could discriminate a DOF of 39.1 mm from a DOF of 62.1 mm. Thus, we could argue that the predictions from blur discrimination may underestimate people's ability to discriminate DOF blur.

5 Discussion

The increment threshold in DOF was measured for two reference DOFs (10.1 mm and 39.1 mm) using two scenes—namely, the 'Apple' scene and the 'Woody' scene. Additionally, the scale of the scene was manipulated, such that the maximum real depth in the scene was 50 cm, 75 cm, or 100 cm. The experiments were conducted under both stereoscopic and nonstereoscopic viewing conditions.

We compared the predicted DOF discrimination with the experimental data in the "Modeling" section. It showed that, for a reference DOF of 10.1 mm, the experimentally measured JND of DOF was smaller than the predicted values. Blur discrimination was investigated based on uniform blur, and we investigated DOF discrimination based on the changes in blur in the scene. For our stimuli we defined the minimum blur circle on the second object in the scene, but found that the predicted values based on this blur circle were much larger than the experimental values. This suggests that the observers may not use the minimum blur circle in the photographs to discriminate DOF when the reference DOF is 10.1 mm. It seems unlikely that observers have used a single higher level of blur, since humans are less sensitive at those higher blur levels and the predicted DOF JND would have been even higher. They may have used information from the combination of multiple blur levels to find the JND in DOF. The statistical analysis suggests that people's ability to discriminate DOF in a photograph is better than the discrimination of any single blur level included in the photograph when the reference DOF is 10.1 mm. However, this is not necessarily the case for a reference DOF of 39.1 mm. Also, when considering our results in the spatial frequency domain, we found that predictions based on blur discrimination may underestimate people's ability to discriminate DOF at 10.1 mm and 39.1 mm.

Our results showed that there was no significant difference in DOF JNDs between the 'Woody' scene and the 'Apple' scene, irrespective of the viewing conditions. Although there were obvious differences in size, color, and amount of spatial overlap, the difference between the two scenes was not big enough to affect the discrimination of JND in the scenes. Because the results for the two scenes were similar replications, we thus demonstrated the reliability of the estimated JNDs in our study.

The scale of the scene as a factor that significantly affects the DOF in the scene was found to significantly influence the JND of DOF. The scale of the scene directly changes the blur gradient visible in the stimuli. According to equations (1) and (2), we can calculate the blur circle on each object in the scene. The difference in blur between the reference DOF and the test values gets larger when the depth increases. So, when we enlarge the scale of the scene, we also enlarge the maximum depth in the scene, and thus increase the difference in blur between the reference and test stimulus, making differences in DOF more visible. The results suggest that photographers and movie directors could put less effort into choosing DOF when the scale of the scene is small, as people are unable to see the differences. However, when the scale of the scene is large, photographers could generate images with different effects by manipulating DOF.

Another interesting finding is that our results do not support the hypothesis that DOF would be easier to discriminate in stereoscopic compared with nonstereoscopic images. There was no difference found between the discrimination in stereoscopic and nonstereoscopic images. Although the stereoscopic DOF itself does not cause any discomfort (O'Hare, Zhang, Nefs, & Hibbard, 2013), the advantages of the stereoscopic images discussed in the introduction may be weakened by the drawbacks of stereoscopic displays such as the vergence–accommodation conflict. This conflict may cause fatigue (Hoffman et al., 2008; Lambooij, Marten, Heynderickx, & Ijsselsteijn, 2009), which in turn may decrease the ability to discriminate DOF. However, we did not find this effect, and neither did we find an increased sensitivity for stereoscopic conditions as predicted by the argument that more details can be seen in stereoscopic viewing than in nonstereoscopic viewing (Heynderickx & Kaptein, 2009), etc. Therefore, we conclude that all these factors are not relevant for discrimination of DOF and that thresholds are similar under stereoscopic and under nonstereoscopic viewing.

6 Conclusion

In summary, we conclude that the discrimination of blur caused by DOF differences is different from the discrimination of uniform Gaussian blur. In general, people are more sensitive to changes in DOF than what would be predicted from the known levels of blur discrimination. In accordance with what is known for blur discrimination, it is easier for observers to discriminate changes in DOF when the reference DOF is small, while people are not so sensitive to changes in DOF when the reference DOF is large. Our research also shows no significant difference between nonstereoscopic and stereoscopic viewing on DOF, indicating that the DOF characteristics of stereoscopic and nonstereoscopic photographs are comparable. Additionally, we conclude that the depth structure in the scene affects observers' ability to discriminate DOF as well.

Acknowledgment. The author Tingting Zhang is supported by a scholarship from the CSC program in China.

References

Bach, M. (1996). The Freiburg Visual Acuity Test: Automatic measurement of visual acuity. *Optometry & Vision Science*, **73**, 49–53.

Born, M., & Wolf, E. (1999). *Principles of optics* (7th ed.). Cambridge: Cambridge University Press.

Campbell, F. W. (1957). The depth of field of the human eye. *Optica Acta: International Journal of Optics*, **4**, 157–164.

Chen, C. C., Chen, K. P., Tseng, C. H., Kuo, S. T., & Wu, K. N. (2009). Constructing a metrics for blur perception with blur discrimination experiments. In S. P. Farnand, F. Gaykema (Eds.), *Proceedings of SPIE: Image Quality and System Performance VI*, **7242**. doi:10.1117/12.806107

Cole, F., Decarlo, D., Finkelstein, A., Kin, K., Morley, K., & Santella, A. (2006). Directing gaze in 3D models with stylized focus. In T. Akenine-Möller, W. Heidrich (Eds.), *Proceedings of the 17th Eurographics Conference on Rendering Techniques* (pp. 377–387). Aire-la-Ville, Switzerland: Eurographics Association. doi:10.2312/EGWR/EGSR06/377-387

Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. *Lecture Notes in Computer Science*, **3953**, 288–301.

Georgeson, M. A. (1998). Edge-finding in human vision: A multi-stage model based on the perceived structure of plaids. *Image and Vision Computing*, **16**, 389–405.

Hamerly, J. R., & Dvorak, C. A. (1981). Detection and discrimination of blur in edges and lines. *Journal of the Optical Society of America*, **71**, 448–452.

Hess, R. F., Pointer, J. S., Simmers, A., & Bex, P. (2003). Border distinctness in amblyopia. *Vision Research*, **43**, 2255–2264.

Hess, R. F., Pointer, J. S., & Watt, R. J. (1989). How are spatial filters used in fovea and parafovea? *Journal of the Optical Society of America A*, **6**, 329–339.

Heynderickx, I., & Kaptein, R. (2009). Perception of detail in 3D images. In S. P. Farnand, F. Gaykema (Eds.), *Proceedings of SPIE: Image Quality and System Performance VI*, **7242**. doi:10.1117/12.806186

Hoffman, D. M., & Banks, M. S. (2010). Focus information is used to interpret binocular images. *Journal of Vision*, **10**(5):13, 1–17.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, **8**(3):33, 1–30.

Joshi, D., Datta, R., Fedorovskaya, E., Luong, T., Wang, J. Z., Jia, L., & Luo J. (2011). Aesthetics and emotions in images. *Signal Processing Magazine*, **28**(5), 94–115.

Lambooij, M., Marten, F., Heynderickx, I., & Ijsselsteijn, W. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, **53**(3) 1–14.

Liu, S., Hua, H., & Cheng, D. (2010). A novel prototype for an optical see-through head-mounted display with addressable focus cues. *IEEE Transactions on Visualization and Computer Graphics*, **16**, 381–393.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., & Martin, K. E. (1996). Occlusion edge blur: A cue to relative visual depth. *Journal of the Optical Society of America A*, **13**, 681–688.

Mather, G. (1997). The use of image blur as a depth cue. *Perception*, **26**, 1147–1158.

Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception*, **31**, 1211–1219.

O'Hare, L., Zhang, T., Nefs, H. T., & Hibbard, P. B. (2013). Visual discomfort and depth-of-field. *i-Perception*, **4**, 156–169.

Otero, J. M. (1951). Influence of the state of accommodation on the visual performance of the human eye. *Journal of the Optical Society of America*, **41**, 942–947.

Pääkkönen, A. K., & Morgan, M. J. (1994). Effects of motion on blur discrimination. *Journal of the Optical Society of America A*, **11**, 992–1002.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **9**, 523–531.

Rogers, S. (1995). Perceiving pictorial space. In W. Epstein, S. Rogers (Eds.), *Perception of space and motion* (Vol. 5) (pp. 119–164). San Diego, CA: Academic Press.

Steve, D., Caitlin, R., & James, T. E. (2010). *Rembrandt's textural agency: A shared perspective in visual art and science*. Cambridge, MA: MIT Press.

Watson, A. B., & Ahumada, A. J. (2011). Blur clarified: A review and synthesis of blur discrimination. *Journal of Vision*, **11**(5):10, 1–23.

Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: Evidence for spatial primitives in human vision. *Vision Research*, **23**, 1465–1477.

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, **5**(10):7, 834–862.

Wheatstone, C. (1838). Contributions to the physiology of vision. *Philosophical Transactions of the Royal Society of London*, **128**, 371–394.

Wuerger, S. M., Owens, H., & Westland, S. (2001). Blur tolerance for luminance and chromatic stimuli. *Journal of the Optical Society of America A*, **18**, 1231–1239.