Perception and Cognition of Depth of Field

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universtiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op vrijdag 6 november 2015 om 15:00 uur

door

Tingting ZHANG

Master in Optical Engineering from Southeast University, China
Bachelor in Electronic Science and Technology from Southeast University, China
born in Liyang, China

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. I.E.J. Heynderickx

Copromotoren:

Dr. H.T. Nefs

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. I.E.J. Heynderickx,	Technische Universiteit Eindhoven, promotor
Dr. H.T. Nefs,	Technische Universiteit Delft, copromotor

Onafhankelijke leden:

Prof. dr. J. Harris, University of St Andrews
Prof. dr. S. F. te Pas, Utrecht University
Prof. dr. A. M. L. Kappers, VU University Amsterdam
Prof. dr. H. de Ridder, Delft University of Technology
Prof. dr. M. A. Neerincx, reserve Delft University of Technology

This study is supported by the Chinese Scholarship Council (No. 2010609040).

ISBN 978-94-6186-546-5

“The most beautiful thing we can experience is the mysterious. It is the source of all true art and science.”

Albert Einstein (1879-1955)

To my dearest family.

Summary

A common way to present 3D materials to human observers nowadays is by stereoscopic displaying on 3D TVs or head-mounted displays such as the Oculus Rift. However, not everyone can see three-dimensional solid shape from stereoscopic viewing and the three-dimensional images remain two-dimensional pictures to them rather than solid shapes. Scientists and artists have spent a lot of effort in finding ways to dissolve pictorial space into visual space. In other words, they want to figure out how to create a sense of stereopsis in an observer when he or she is looking at non-stereo and stereo pictures. There are many monocular depth cues in addition to binocular disparity that can enhance depth perception, and these cues can therefore also create an impression of stereopsis when the viewing conditions are right. One of these depth cues is *depth of field*. Depth of field is defined as the distance range in which objects are perceived as sharp. Depth of field is a popular photographic technique that effectively makes the main subject in a picture appear sharp and the foreground and background blurred.

The goal of this thesis is to understand the roles of depth of field in pictures from a relatively low perceptual level to a relatively high cognitive level. This thesis mainly reveals how depth of field influences the impression of stereopsis and how binocular disparity influences the perception of depth of field. I start the thesis with a study on the discrimination thresholds of depth of field. Then, a subjective study is reported in which I investigated the effects of depth of field on depth perception for binocular viewing. The next two studies address the roles of depth of field on a more cognitive level. The first of these two studies explored the effects of depth of field on change detection in pictures. The other one was conducted to evaluate how depth of field influences the aesthetic appeal and overall quality in photographs.

Four subjective studies were conducted to achieve our goals. The cornerstone of the work described in this thesis is that humans are much more sensitive to changes in small depth of field than in large depth of field. A second important finding is that stereoscopic viewing does not significantly affect discrimination thresholds. Based on the discrimination thresholds that were measured in this

study, I selected five levels of depth of field that can be well discriminated under stereo viewing conditions in the second study to explore the effects of depth of field on depth perception. I found that the presence of strong depth cues, in this case binocular disparity, weakens the effects of depth of field as a depth cue when depth of field is small. In contrast, when depth of field is not small, perceived depth decreases with increasing depth of field irrespective of whether binocular disparity is zero or not. With respect to the effects of depth of field on change detection, I found that depth of field directs viewers' attention similarly under both non-stereo and stereo viewing conditions. Depth of field does however weaken the effect of binocular disparity on change detection. In the final study, I found that there are no common rules on how to manipulate depth of field to make pictures more beautiful or have higher quality. The role of depth of field on aesthetic appeal and overall quality varies across content categories.

Samenvatting

Een gebruikelijke manier om 3D objecten tegenwoordig aan menselijke kijkers te laten zien is door gebruik te maken van 3D TV's of een op het hoofd gedragen beeldscherm zoals de Oculus Rift. Maar, niet iedereen kan een solide drie-dimensionele vorm zien op basis van een stereoscopisch beeld en de driedimensionale beelden blijven twee-dimensionele afbeeldingen in plaats van vaste vormen. Wetenschappers en kunstenaars hebben veel moeite gedaan om de afbeeldingsruimte in de visuele ruimte te integreren. Met andere woorden, zij wilden uitvinden hoe een gevoel van vaste vorm kon worden bewerkstelligd in een waarnemer wanneer hij of zij naar stereo en niet-stereo afbeeldingen kijkt. Er zijn vele monoculaire diepteaanwijzingen -naast binoculaire dispariteit- die de dieptewaarneming kunnen benvloeden en deze diepteaanwijzingen kunnen daarom ook een indruk van vaste vorm geven als de omstandigheden het toelaten. Scherptediepte is zo'n diepteaanwijzing. Scherptediepte is gedefinieerd als het dieptebereik waarin objecten als scherp worden waargenomen. Scherptediepte is een populaire techniek in de fotografie die de facto het hoofdonderwerp scherp afbeeldt in een foto, maar de voor- en achtergrond wazig maakt.

Het doel van dit proefschrift is het begrijpen van de verschillende rollen die scherptediepte speelt van een relatief laag perceptueel tot een relatief hoog cognitief niveau. Dit proefschrift laat zien hoe scherptediepte de indruk van vaste vorm kan benvloeden en hoe binoculaire dispariteit de waarneming van scherptediepte benvloedt. Ik begin dit proefschrift met een studie over de discriminatiedrempels voor scherptediepte. Vervolgens wordt een studie gerapporteerd waarin ik de effecten van scherptediepte op de dieptewaarneming onderzocht heb. De volgende twee studies beschouwen scherptediepte op een meer cognitief niveau. De eerste van deze twee studies bekijkt het effect van scherptediepte om veranderingsblindheid in afbeeldingen. De andere studie werd uitgevoerd om te evalueren hoe scherptediepte de esthetische aantrekkelijkheid en de overall beeldkwaliteit bepaalt.

Er werden vier subjectieve studies gedaan om onze doelen te bereiken. De hoeksteen van het werk dat beschreven is in dit proefschrift is dat mensen gevoeliger

zijn voor veranderingen in een kleine scherptediepte dan in een grote scherptediepte. Een tweede belangrijke bevinding is dat stereoscopisch zien de discriminatiedrempels niet significant verandert. Gebaseerd op de discriminatiedrempels die in deze studie werden gemeten, heb ik vijf scherptediepteniveaus geselecteerd die makkelijk van elkaar konden worden onderscheiden, om de effecten van scherptediepte op de dieptewaarneming te bestuderen. Ik vond dat de aanwezigheid van sterke diepteaanwijzingen, in dit geval binoculaire dispariteit, de effecten van scherptediepte als een diepteaanwijzing afzwakt, wanneer de scherptediepte klein is. Als de scherptediepte daarentegen niet klein is, neemt de waargenomen diepte toe met toenemende scherptediepte onafhankelijk of de binoculaire dispariteit nul is of niet. Met betrekking tot de effecten van scherptediepte op veranderingsblindheid heb ik gevonden dat scherptediepte aandacht op dezelfde manier benvloedt in stereo als in niet-stereo kijkcondities. Scherptediepte zwakt echter het effect van binoculaire dispariteit op veranderingsdetectie af. In de laatste studie heb ik gevonden dat er geen universele regels zijn over hoe scherptediepte moet worden gemanipuleerd om afbeeldingen mooier te maken of een hogere beeldkwaliteit te geven. De rol van scherptediepte op esthetische aantrekkelijkheid en overall kwaliteit verschilt met verschillende onderwerpscategorien.

Contents

Chapter 1

Introduction

1.1 Pictures and Stereopsis

A picture can be regarded as a surface that consists of an optic array of arrested structures (Gibson, 2013), which can be created by painting, drawing, photography or by computer rendering. A picture can also be considered as a record of perception, recording what the picture maker was seeing at that specific time and location. When looking into a picture, people will become aware of what is called “pictorial space”. The term pictorial does not refer to the physical matter of the space, and neither to the physical substrate of the image, but rather to the imaginary world depicted on the canvas, display or other medium (Pettersson, 2011). Pictorial space is very different from physical space. Physical space has, at least on an ecologically relevant scale, a Euclidean geometry and observers can move in it, whereas pictorial space exists in the mental domain rather than in the physical domain. Simultaneously, the viewer is aware that the picture itself is a physical object (usually a two dimensional surface of some sort), and that there is a three-dimensional scene depicted by the picture.

Artists and other picture-makers have spent a lot of energy in finding ways to dissolve pictorial space into visual space. That is, “tricking the eye” to believe that what they see is not a two-dimensional surface but part of the real world; or, in other words, to create “stereopsis”. “Stereopsis” comes from the Greek root “stereo” which means “solid” (Howard and Rogers, 2002). Stereopsis refers to the impression that objects or scenes are “solid” rather than depicted: that is, not pictorial. A strong way to create stereopsis is to present the viewer with slightly disparate images of the world to the two eyes. Here, stereopsis had been meant to refer to “binocular stereopsis”- the kind of percept you have when looking with two eyes. Not until the early twentieth century was “monocular stereopsis” acknowledged. At that time, the optical industry produced two major types of systems to create stereopsis from single pictures: Zeiss “Verant” (designed by Rohr and Gullstrand,1904) and Zeiss “Synopter” (designed by van Rohr, 1920). These inventions show that “monocular stereopsis” is perfectly possible. In fact, just by closing one eye and looking into the world one still has the impression that the world is solid even though the world is projected onto a curved, but otherwise two-dimensional surface, namely the retina of the eye.

Enhancing depth cues in images is one of the important ways to increase stereopsis. Both monocular depth cues such as perspective, occlusion (Howard and Rogers, 2002) and binocular depth cues such as binocular disparity have already been investigated a lot in the past several decades to explore the relation between these cues and depth perception (Braunstein, 1962; Howard, 1995; Marshall et al., 1996; Rogers and Graham, 1982). However, it remains unclear what exactly happens when pictorial perception becomes stereopsis. It remains for example unclear how depth cues that are inconsistent with each other or with the viewer may affect perception and attention. Some models for visual perception pose that perceptual conflicts are vetoed or that some kind of Bayesian weighted average is taken (Buelthoff and Yuille, 1991; Landy et al., 1995), but many of these studies are based on simplified stimuli of geometrical shapes. It is important to understand what happens where picture and stereopsis meet with the effects of depth cues. In this thesis we investigate how a particularly interesting depth/distance cue, namely "depth of field" behaves in this respect.

In 1941, cinematographer Gregg Toland and director Orson Welles used a very small aperture to make details sharp in both the foreground and the background in the movie "Citizen Kane"; this is known as "deep focus". Since then, photographers and cinematographers have used depth of field more and more often not only to create deep-focus effect but also shallow-focus effect. Depth of field is defined as the distance range within which objects are perceived as sharp.

So far, we have already known that when we focus on one object in the real world, this object will be projected in the fovea on the retina and hence be perceived as sharp. The images of other objects are more blurred if the distance from the other objects to the focus plane is farther away. It is possible to create an image with proper blur in it to make it the same as the image of the real scene on the retina. It suggests that proper depth of field in an image may create a realistic pictorial space. Hence, a picture can be perceived as an "Alberti" window (Alberti, 2005), through which observers seem to see things in real world. The size of the depth of field in the picture can be controlled by the picture-makers to manipulate how much observers can see. In other words, picture makers can manipulate stereopsis by controlling depth of field under non-stereo viewing conditions (Mather and Smith, 2000). However, it is not clear to what extent depth of field can influence stereopsis under stereo viewing conditions when binocular stereopsis is a strong depth cue and hence observers may feel that they are looking into a real world. In this case, stereopsis may influence the perception of depth of field. Although the relation between depth of field, stereopsis and other cues are not very well documented, depth of field has already been applied widely.

From a practical point of view, there are several reasons why depth of field effect is popular in cinema and photography industry. First, it can be controlled to create miniaturization (diorama illusion) or magnification effect (Held et al., 2012). Second, it is believed that depth of field effect can enhance the realism and aesthetics of the photographs or videos (Datta et al., 2006; Hillaire et al., 2007). Third, shallow depth of field may help improve the viewing experience after image or video compression by blurring the background.

From a theoretical point of view, depth of field effect has also attracted many researchers' attention for several reasons. First, depth of field has been found to be a pictorial depth cue (Pentland, 1987). Secondly, depth of field can be used to concentrate viewers' attention at a specific point in the image and thereby emphasize that part of the image (Cole et al., 2006; DiPaola, 2009; Kingslake, 1992). Thirdly, it also has been shown that restricting depth of field in virtual reality improves people's performance and preference (Hillaire et al., 2007).

Before looking into the details about the research interest of this thesis, the technical knowledge of depth of field was first introduced in the following subsection 1.1.1. With the background knowledge of depth of field, it would be easier to understand the research described in this thesis.

1.1.1 The optical characteristics of depth of field

Before investigating depth of field in psychological area, the optical characteristics of depth of field will first be introduced briefly since it is essential for the following studies. Images of the real scenes often contain regions that are gradually blurred because of distance variations in the scene and limitation of the eye, camera, or other optical systems. That is, if an object at a certain distance were brought into focus by an ideal optical system, its image will by definition appear to be sharp. In other words, the image of the object which is in the focal plane of the lens will be sharp. However, objects that are closer or farther away than the focal plane will be "out of focus" and appear blurred in the image. The amount of blur increases as a function of the distance away from the focal plane. Because the eye can only resolve details down to a certain level, a certain distance range in the object space will be perceived as sharp in the (retinal) image. Figure 1.1 describes the relationship between the blur of the object and the distance how far it is from the focus plane. The distance range between the closest and the farthest objects in the scene that appears to be sharp is called the "depth of field", which can be calculated based on Equation 1.1. In Equation 1.1, F_value represents the aperture size, L is the focal length, D is the focal distance.

$$\delta = \frac{2\frac{L}{F\_value}^{2}\tan\frac{\beta}{2}}{D - 2\frac{L}{F\_value}\tan\frac{\beta}{2}} \tag{1.1}$$

The two main methods to adjust the size of depth of field is to manipulate the aperture size of the camera or the distance between the camera and the focal plane.

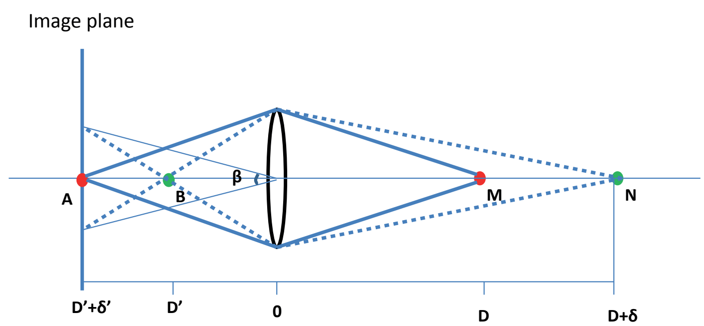

Figure 1.1: The lens at position 0 is focusing on an object at position D. Object N is out of focus whose image is a disk on image plane. The diameter of the image of object N is defined as blur circle. The angular size of the blur circle as seen from the center of the lens, is indicated with β.

1.2 Research questions

The goal of this thesis is to understand the roles of depth of field in pictures from a low perception level to a high cognitive level. Hence, we start the thesis from finding the discrimination thresholds of depth of field to provide information on how to control depth of field in the following studies, followed by a study on investigating the relation between depth of field and depth perception. After that, we conduct a study to explore the effects of depth of field on change detection, which can reflect the relation between depth of field and attention and can also give suggestions on how to inhibit or facilitate change blindness. In the end, a study to investigate the influence of depth of field on the beauty and visual quality of the pictures is carried out. The evaluation of aesthetics and overall quality is up to the cognitive understanding of depth of field, which can help applications in practical industry such as improving the viewing experience of 3D TV from the point view of the end-users.

According to the definition of depth of field, it is obviously difficult to give out an accurate value of depth of field because of the phrase “appears to be sharp”. It largely depends on the limitation of human visual system whose ability to measure changes in depth of field is not infinite. The size of depth of field can change gradually in a photograph, and it remains unclear how many different levels of depth of field humans can distinguish. The data on the discrimination of depth of field may improve the efficiency to generate pictures through which observers can absolutely perceive different worlds. As we have already known, depth of field in non-stereo pictures may create stereopsis. Hence, the discrimination of depth of field in non-stereo pictures may not only reflect the perception of depth of field in pictorial space but also give suggestions on how to control the stereopsis created by depth of field. There has already been a lot of research about the discrimination of depth of focus (Campbell and Green, 1965) and image blur (Hamerly and Dvorak, 1981; Liu et al., 2008; Mather and Smith, 2002; Wuerger et al., 2001) in non-stereo viewing conditions, which may give suggestions on how to conduct studies to find out the thresholds of depth of field and build proper hypotheses (Campbell and Green, 1965; Marcos et al., 1999; Wang and Ciuffreda, 2004).

However, the discrimination of depth of field in stereo pictures may be not the same as in non-stereo pictures. With stereo pictures, observers can perceive a solid space. We have no idea whether the stereopsis will affect the perception of depth of field or not. Hence, it is worthwhile to investigate the thresholds of depth of field in stereo pictures and compare them with that in non-stereo pictures. In this context, the first research question that is formulated and addressed in Chapter 2 is:

1. What is the discrimination threshold of depth of field in (stereo) photographs and will factors such as viewing condition, image content, or scale of the scene influence the threshold of the depth of field?

Enhancing the depth perception can increase the sensation of presence or immersion. In other words, depth perception can enhance the perception of stereopsis. Hence, investigating the effects of depth of field on depth perception may provide a simple and effective method to manipulate perceived depth and change the impression of stereopsis.

So far, the effects of depth of field on depth perception under non-stereo viewing conditions have been directly investigated (Nefs, 2011), it is still unclear whether the conclusions from non-stereo viewing can be generalized for stereo viewing conditions. Binocular disparity as an important depth cue may be the dominant cue when there are multiple depth cues, which may affect the effects of depth of field on depth perception. On one hand, depth of field can create the impression of stereopsis, which may facilitate the stereopsis together with binocular disparity. On the other hand, the stereopsis created by depth of field may be in conflict with the stereopsis created by binocular disparity and hence introduce viewing discomfort. Additionally, the investigation on the relation between depth of field and perceived depth in stereo pictures can also make it clear whether binocular disparity will weaken or enhance the effects of depth of field on perceived depth. Hence, chapter 3 addresses the following research question:

2. How does depth of field influence perceived depth when there are other strong depth cues such as binocular disparity in pictures?

So far, the influence of depth of field on depth perception under either non-stereo or stereo viewing conditions is clear. This low level feature depth of field helps understand the effects of depth of field at a relatively high level. For example, researchers suggest that depth perception plays an important role on attentive behavior when viewing 3D content (Huynh-Thu et al., 2011). From this point of view, it is reasonable to indicate that depth of field as a depth cue may influence visual attention. However, this is not the only reason that depth of field can influence attention, it has also been found that depth of field can direct attention because the blur separates the foreground and background from the focus objects (Cole et al., 2006; Khan et al., 2010; Shepherd et al., 1986).

Change detection not only requires direct attention to be paid to the area where change happens but also the attention should be held on the change. Hence, depth of field may influence the change detection since it influences visual attention. In the studies of visual attention, researchers always use eye-tracking apparatus to record eye movements to get the saliency maps of images. The conclusions on the effects of depth of field on visual attention were drawn mainly based on the saliency maps of the images (Khan et al., 2010; Shepherd et al., 1986). However, saliency maps only provide limited insight in how visual attention is distributed over the scene as a function of time. Saliency maps are also not the most effective and accurate way to represent the allocation of attention. Hence, the flicker paradigm (Rensink et al., 1997) used in change detection may be another method to double check the effects of depth of field on visual attention.

Previous studies also show that a change happening closer to the observer may be detected faster than a change happening farther away from the observer (Jansen et al., 2009; Mazza et al., 2005). Hence, if there is a very strong depth cue in the image such as binocular disparity, the effect of depth of field on change detection may be influenced. To explore more about the effects of depth of field in attention related area, Chapter 4 addresses the following research question:

3. How does depth of field influence the change detection in pictorial scenes and whether stereopsis will influence such effects since binocular disparity may introduce cue conflicts such as vergence-accommodation conflicts?

Apart from the interest from researchers in psychology, it is also important to explore depth of field from the point of view of cinematographers, photographers, or customers of cameras. Aesthetics is always an important aspect that photographers try to achieve when they take photos. Depth of field is used by photographers for both aesthetic and realistic reasons. Low depth of field indicates that the focus object is sharp while surroundings are blurred which may highlight the main subject of an image and cover the useless objects. Hence it may improve the evaluation of aesthetics of an image (Datta et al., 2006; Dhar et al., 2011). Although researchers reported statements relating depth of field blur to aesthetics (Datta et al., 2006; Luo and Tang, 2008), these statements are only marginally substantiated by empirical data.

Since aesthetic appeal is assumed to affect the overall quality positively (Fedorovskaya, 2002; Loui et al., 2008) and depth of field is assumed to have a positive effect on aesthetic appeal, it is reasonable to expect a positive effect of depth of field on overall quality. However, there are also studies on the negative influence of distortion blur on overall quality (Crete et al., 2007; de Ridder, 1998). Hence, the effects of depth of field on overall quality still need to be examined. In Chapter 5, the fourth research question is formulated as follows:

4. How does depth of field influence the subjective evaluation of aesthetic appeal and overall quality in photographs without considering other features such as color, contrast, and so on?

1.3 Approach and thesis structure

The thesis is structured in six chapters, including Introduction and Conclusions. Chapter 2, 3, 4, 5 are based on four empirical studies to address the research questions raised in section 1.2.

To address the first research question, we conducted a subjective experiment. We took pictures of a scene created in the lab, and then generated the non-stereo and stereo stimuli separately. A few observers participated in the experiment and the experiment was based on a two-alternative forced-choice (2AFC) procedure. The thresholds of selected depth of field were calculated by using a Gaussian cumulative psychometric function. Then, statistical analysis was performed on the experimental data to investigate the effects of scene content and scale of the scene on discrimination of depth of field. After that, we used a model from the literature to predict the thresholds of the selected depth of field and then compared the predicted values with the experimental values. In the end, we also performed Fourier analysis on the stimuli to get a better insight about the discrimination of depth of field in pictures. The details of the empirical study and its results are demonstrated in Chapter 2.

In order to explore the influence of stereopsis on the relation between depth of field and perceived depth, we designed an empirical study. We created controllable scenes including flowers as background and two puppets standing on a floor. Non-stereo and stereo stimuli were generated with a d-SLR camera separately. Two stereo systems were used to present the stimuli: Wheatstone Stereoscope and 3D TV. Participants were asked to estimate the position of one puppet in the stimulus. After collecting the data of the estimated positions, statistical analysis was performed to investigate the effects of depth of field and physical depth on perceived depth. In addition, the results from the stereoscope were compared with that from 3D TV to check whether the effects of stereopsis on the relation between depth of field and perceived depth were robust or not. Chapter 3 presented all the details of the study.

In Chapter 4, a within-subjects experiment and its results were described to address the third research question how depth of field influences the change detection in pictorial and solid scenes. A flicker paradigm was used to show two scenes alternatively with a gray scene in between. Stimuli were taken with a d-SLR camera based on a scene created in the lab. There were two sessions in the experiment, one was using pictorial scenes (two-dimensional) and the other was using solid scenes (three-dimensional). Participants participated in both sessions and were required to response as soon as they found the change in the flicker paradigm. In such experiments, the hit rate and response time were recorded to quantify the change detection.

In Chapter 5, we investigated how depth of field influenced the evaluation of the aesthetic appeal and overall quality of photographs covering several categories. In this experiment, most of the photographs used were obtained from Internet sources and a few were from personal collections or previous studies. A preliminary experiment was done before the main experiment to first quantify the size of the depth of field of all the photographs. After all the photographs have been evaluated, another group of participants observed all these photographs in the lab on a screen to give a score for the aesthetic appeal and overall quality separately in different sessions. The method of Single Stimuli was used in this study, so participants only need to score one image for either aesthetic appeal or overall quality in one trial.

In Chapter 6, the main conclusions of all the four studies are discussed. Additionally, the contribution and the limitation together with the possible future work are also presented.

Bibliography

Alberti, L. (2005). *On painting*. Penguin UK.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: effects of numerosity and perspective. *Journal of experimental psychology*, 64(4):415.

Buelthoff, H. H. and Yuille, A. L. (1991). Shape-from-x: Psychophysics and computation. In *Fibers' 91, Boston, MA*, pages 235–246. International Society for Optics and Photonics.

Campbell, F. and Green, D. (1965). Optical and retinal factors affecting visual resolution. *The Journal of Physiology*, 181(3):576.

Cole, F., DeCarlo, D., Finkelstein, A., Kin, K., Morley, K., and Santella, A. (2006). Directing gaze in 3d models with stylized focus. *Proceedings of the 17th Eurographics conference on Rendering Techniques*, pages 377–387.

Crete, F., Dolmiere, T., Ladret, P., and Nicolas, M. (2007). The blur effect: perception and estimation with a new no-reference perceptual blur metric. *Electronic Imaging*, pages 64920I–64920I–11.

Datta, R., Joshi, D., Li, J., and Wang, J. Z. (2006). Studying aesthetics in photographic images using a computational approach. *Computer VisionECCV*, pages 288–301.

de Ridder, H. (1998). Psychophysical evaluation of image quality: from judgment to impression. *Electronic Imaging*, pages 252–263.

Dhar, S., Ordonez, V., and Berg, T. L. (2011). High level describable attributes for predicting aesthetics and interestingness. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1657–1664.

DiPaola, S. (2009). Exploring a parameterised portrait painting space. *International Journal of Arts and Technology*, 2(1):82–93.

Fedorovskaya, E. A. (2002). Perceived overall contrast and quality of the tone scale rendering for natural images. *Electronic Imaging*, pages 119–128.

Gibson, J. J. (2013). *The ecological approach to visual perception*. Psychology Press.

Hamerly, J. R. and Dvorak, C. A. (1981). Detection and discrimination of blur in edges and lines. *JOSA*, 71(4):448–452.

Held, R. T., Cooper, E. A., and Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology*, 22(5):426–431.

Hillaire, S., Lcuyer, A., Cozot, R., and Casiez, G. (2007). Depth-of-field blur effects for first-person navigation in virtual environments. *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 203–206.

Howard, I. P. (1995). *Binocular vision and stereopsis*. Oxford University Press.

Howard, I. P. and Rogers, B. (2002). *Seeing in depth, Vol. 1: Basic mechanisms*. University of Toronto Press.

Huynh-Thu, Q., Barkowsky, M., and Le Callet, P. (2011). The importance of visual attention in improving the 3d-tv viewing experience: overview and new perspectives. *IEEE Transactions on Broadcasting*, 57(2):421–431.

Jansen, L., Onat, S., and Knig, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1):29.

Khan, R. A., Konik, H., and Dinet, E. (2010). Enhanced image saliency model based on blur identification. *25th International Conference of Image and Vision Computing New Zealand (IVCNZ),*, pages 1–7.

Kingslake, R. (1992). *Optics in photography*, volume 6. SPIE Press.

Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research*, 35(3):389–412.

Liu, R., Li, Z., and Jia, J. (2008). Image partial blur detection and classification. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8.

Loui, A., Wood, M. D., Scalise, A., and Birkelund, J. (2008). Multidimensional image value assessment and rating for automated albuming and retrieval. *15th IEEE International Conference on Image Processing*, pages 97–100.

Luo, Y. and Tang, X. (2008). Photo and video quality evaluation: Focusing on the subject. *Computer VisionECCV*, pages 386–399.

Marcos, S., Moreno, E., and Navarro, R. (1999). The depth-of-field of the human eye from objective and subjective measurements. *Vision research*, 39(12):2039–2049.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., and Martin, K. E. (1996). Occlusion edge blur: a cue to relative visual depth. *JOSA*, 13(4):681–688.

Mather, G. and Smith, D. R. (2000). Depth cue integration: stereopsis and image blur. *Vision research*, 40(25):3501–3506.

Mather, G. and Smith, D. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception*, 31(10):1211–1220.

Mazza, V., Turatto, M., and Umilt, C. (2005). Foregroundbackground segmentation and attention: A change blindness study. *Psychological Research*, 69(3):201–210.

Nefs, H. T. (2011). Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing and perceiving*, 25(6):577–595.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):523–531.

Pettersson, M. (2011). Seeing what is not there: Pictorial experience, imagination and non-localization. *British Journal of Aesthetics*, 51(3):279–294.

Rensink, R. A., O'Regan, J. K., and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological science*, 8(5):368–373.

Rogers, B. and Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision research*, 22(2):261–270.

Shepherd, M., Findlay, J. M., and Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, 38(3):475–491.

Wang, B. and Ciuffreda, K. J. (2004). Depth-of-focus of the human eye in the near retinal periphery. *Vision research*, 44(11):1115–1125.

Wuerger, S. M., Owens, H., and Westland, S. (2001). Blur tolerance for luminance and chromatic stimuli. *JOSA A*, 18(6):1231–1239.

Chapter 2

Discrimination

Human discrimination of depth of field in stereoscopic and non-stereoscopic photographs

This chapter has been published as:

Human discrimination of depth of field in stereoscopic and nonstereoscopic photographs (2014) *Tingting Zhang, Harold T. Nefs, Ingrid Heynderickx,* Perception, *43(5), 368-380.*

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 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 to predict discrimination thresholds for DOF. We therefore measured discrimination thresholds for DOF using a 2AFC-task. Ten participants were asked to observe two images and select the one with larger DOF. We manipulated the scale of the scene, i.e. the actual depth in the scene. We did the experiment under stereoscopic and non-stereoscopic viewing conditions. We found that the threshold for large DOF (39.1mm) was higher than for small DOF (10.1mm), and the threshold decreased when scale of scene increased. We also found that there was no significant difference between stereoscopic and non-stereoscopic 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.

2.1 Introduction

Depth of field is the distance range within which objects are perceived as sharp. Objects that are outside of the range of depth of field will appear blurred in an image. Figure 2.1 shows an example of a small and a large depth of field respectively. Depth of field has various applications in enhancing the subjective quality of images. For example, depth of field may be used to enhance depth perception in photographs (Marshall et al., 1996; Pentland, 1987; Watt et al., 2005). Secondly, it has been shown to contribute to the aesthetic appreciation of photographs (Datta et al., 2006), and to make images appear more natural and realistic (Joshi et al., 2011). Thirdly, depth of field 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; DiPaola, 2010). To better understand the aesthetic and attention effects, it would be good to know the differences in depth of field that can be perceived by the average viewer and whether they can be predicted from blur discrimination.

Because depth of field is perceived as a change in blur in an image, it seems plausible that perceived differences in depth of field 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 sec of arc. 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 minute of arc. Consistency in these results were shown across a variety of studies in spite of different stimuli and experimental methods

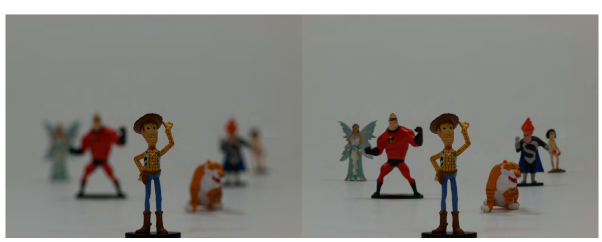

Figure 2.1: Depth of field effects. Left: small depth of field. Right: large depth of field.

(Hess et al., 1989; Mather, 1997; Mather and Smith, 2002; Watt and Morgan, 1983; Wuerger et al., 2001). Assuming a peak ability to discriminate blur at about one minute of arc, we may predict that this value is the limiting factor in discriminating depth of field. If the image only contains regions with larger or smaller blur circles, the threshold will be larger than when the image does contain blur circles around one minute of arc.

There are, however, basic differences between blurred images and images with a limited depth of field; i.e., in the latter case the level of blur is not homogeneously distributed over the whole image, but depends on the distance locally of the imaged object with respect to the focal plane. Depth of field 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, 2011; Hamerly and Dvorak, 1981; Paakkonen and Morgan, 1994), binary texture (Hoffman and Banks, 2010) or random dot stereograms (Mather and Smith, 2002), rendered by computer algorithms. In contrast, our stimuli contained a blur gradient over the figurines in the scene that was affected by the scale of the scene. Even though the peak sensitivity is at a blur circle of one minute of arc, it is possible that people still benefit from the presence of blur at other (suboptimal) levels to discriminate depth of field.

Stereoscopic and non-stereoscopic 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 non-stereoscopic photographs because of the tight link between convergence and accommodation (Hoffman et al., 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 depth of field may be influenced by the optical state of the eyes (Campbell, 1957), and therefore, the discrimination in depth of field may be different for stereoscopic and non-stereoscopic photographs. Secondly, the subjective experience of depth is qualitatively different in stereoscopic compared to non-stereoscopic photographs. In non-stereoscopic 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 non-stereoscopic images, which could in principle be used to gain more complex information (Liu et al., 2010). It was found that more details could be perceived in stereoscopic images than in non-stereoscopic images (Heynderickx and Kaptein, 2009). Since more details can be observed in stereoscopic images, 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 depth of field in stereoscopic images than in non-stereoscopic images.

In the current study, we measured the just noticeable difference (JND) of two depths of field. 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 depth of field was measured using both stereoscopic and non-stereoscopic photographs.

2.2 Experiment

2.2.1 Methods

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 (Lamris Ootech BV), participated in our experiment. The informed consent forms were obtained from all participants. This research was approved by the Delft University of Technology and done according to the Declaration of Helsinki, Dutch Law and common local ethical practice.

Apparatus and stimuli. A Wheatstone stereoscope (Wheatstone, 1838) with two 19 inches 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 2.2 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 depth of field) to F22 (the largest depth of field). The angle of view of the camera was 13.2°(horizontally)×9.9°(vertically). The size of the stimuli displayed on the screens was constrained by the visual angle of the camera.

Figure 2.3 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 depth of field 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 2.3. The distance between the real-world objects, and so the physical depth structure in the scene, was also manipulated; this

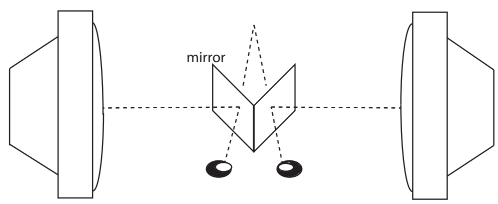

Figure 2.2: Diagram of the stereoscope used in the experiment.

factor was 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 "scale of the scene" factor, namely 50cm, 75cm and 100cm. 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 & F13, which acted as reference values for the depth of field. To measure the JND in depth of field, 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 depth of field of F3.5, and with the aperture of the camera being F7.1, F8, F9, F10, F11, F14, F16, F18, F20 and F22 for the reference depth of field 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 6.5cm in our study. For the non-stereoscopic viewing conditions, we set the camera in the middle of the slide-bar. When taking the photographs, the camera was centered on the central figurine. 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 figurines were thus shown life-sized.

Procedure. The experiment was based on a within-subject design for the three independent variables, being reference depth of field, scale of the scene and stereoscopic vs. non-stereoscopic 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

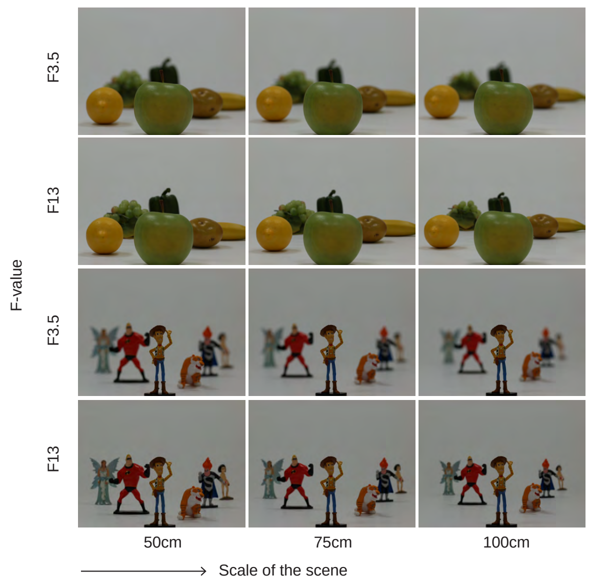

Figure 2.3: Stimuli used in the experiment.

displayed simultaneously side-by-side in the middle of both monitors. Observers were asked to decide which image appeared to have the larger depth of field 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, and the participants could take as long as they needed. The participants evaluated 240 trials per session (i.e., 2 Scenes $\times$ 3 levels of Scale of the scene $\times$ 2 Reference Depth of Field $\times$ 10 Comparisons presented twice (reference once on the left and once on the right half of the monitor)). The full experiment consisted of 30 sessions, of which half used non-stereoscopic images and half stereoscopic images (and so, we had 15 repetitions per viewing condition). The sessions with non-stereoscopic images alternated with the sessions with stereoscopic images. For each participant, the starting session (i.e., stereoscopic or non-stereoscopic) and the order of the comparisons within a session was random.

2.2.2 Analysis

Depth of field can be described in different ways such as 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 2.1, where L is the focal length of the camera and a is the aperture size. In our work, the distance range in millimeters within which the blur circle is smaller than one minute of

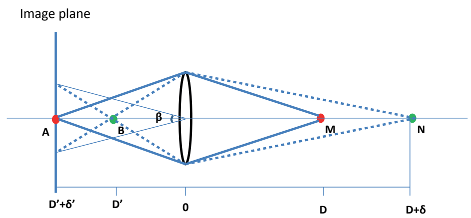

Figure 2.4: The lens at position 0 is focusing on an object at position D. Object N is out of focus whose image is a disk on image plane. The diameter of the image of object N is defined as blur circle. The angular size of the blur circle as seen from the center of the lens, is indicated with β.

arc is used as the value of depth of field (Born and 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 2.2, with β indicating the angular size of the blur circle. Figure 2.4 shows the geometrical relationship between aperture size, focus distance and blur circle. Table 2.1 summarizes the values of depth of field and the values of the aperture size in units of mm corresponding to all the F-values.

$$F\_value = \frac{L}{a} \tag{2.1}$$

$$\delta = \frac{2\frac{L}{F\_value}^2 \tan\frac{\beta}{2}}{D - 2\frac{L}{F\_value}\tan\frac{\beta}{2}} \tag{2.2}$$

The proportion of trials where the participant chose the test as having the larger depth of field from the combination of reference and test stimulus was fitted using a Gaussian cumulative function. The difference between the depth of field 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 (Just Noticeable Difference) of the reference depth of field.

2.3 Results

We found that the JNDs for a depth of field of 10.1mm (i.e., F3.5) across all conditions and all participants ranged between 0.14mm and 4.17mm. For a reference depth of field value of 39.1mm (i.e., F13), the JNDs ranged between 0.6mm and 57.16mm. The data thus showed large individual differences, in-

Table 2.1: The value of depth of field and the aperture size in millimeter 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

dicating that some people were sensitive to changes in depth of field, whereas others could not really discriminate depth of field well.

The JND values averaged across all ten participants are summarized in Figure 2.5. Figure 2.5(a) shows that the JND for a reference depth of field of 39.1mm was much larger than the JND for a reference depth of field of 10.1mm. There was no big difference in JND between the "Apple" scene and the "Woody" scene in Figure 2.5(b) while the JND in depth of field was found to decrease with increasing scale of the scene, as shown in Figure 2.5(c). Figure 2.5(d) demonstrates the discrimination thresholds observed under non-stereoscopic and stereoscopic viewing conditions.

We performed a 2 (Reference DOF) $\times$ 2 (Replications (Scene)) $\times$ 3 (Scale of the scene) $\times$ 2 (Viewing condition) repeated-measures ANOVA. We found significant main effects of Reference DOF ($F(1,9) = 15.54$, $p = .003$) and Scale of the scene ($F(2, 18) = 7.81$, $p = .004$). Additionally, a significant interaction between Reference DOF and Scale of the scene was found ($F(2, 18) = 7.52$, $p = .004$). Figure 2.5(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.

2.4 Modeling

2.4.1 Predicting JNDs of DOF from blur discrimination

In this section, we predicted the values of JNDs from blur discrimination studies in literature and compared them with our experimental results. Blur discrimination studies, however, typically only have used 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 depths of field. 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

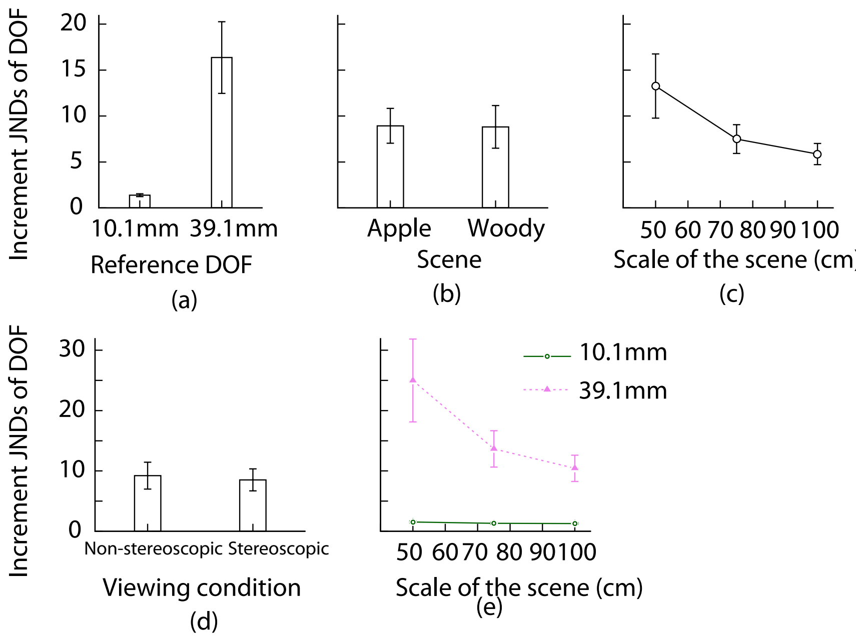

Figure 2.5: The averaged increment JNDs across participants with the error bars represent ±1 standard error of the mean value. (a) JNDs in photographs with two reference DOF: 10.1mm and 39.1mm; (b) JNDs in photographs with different content: Apple and Woody; (c) JNDs in photographs with different scale of the scene: 50cm, 75cm, and 100cm; (d) JNDs in photographs under different viewing conditions: non-stereoscopic and stereoscopic; (e) Interaction between Scale of the scene and viewing condition.

minimum blur circle and the blur circle of one minute of arc.

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 regarded as the minimum visible blur circle in the stimuli. Equation 2.3 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 depth of field. In our experiment, d could be 10cm, 15cm, or 20cm, depending on the scale of the scene.

$$\tan b = \frac{L \times d}{D \times (D + d) \times F\_value} \tag{2.3}$$

Although a blur circle of one minute of arc 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 reference for two reasons. First, the definition of depth of field in our paper was based on the blur circle of one minute of arc. Second, the peak sensitivity for blur discrimination was found to be around one minute of arc (Chen et al., 2009; Hamerly and Dvorak, 1981; Hess et al., 1989; Mather and Smith, 2002; Watt and 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 b1 could be discriminated from a smaller blur circle b2 when b1 was a factor ω times b2, 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 (i.e., the minimum blur circle and the blur circle of one minute of arc) , while the intrinsic blur was obtained from literature. Finally, the equation for the blur discrimination threshold is shown as follows:

$$a = -r + \sqrt{\omega^2(\beta^2 + r)^\rho - \beta^2} \tag{2.4}$$

Table 2.2: Weber model parameters for four studies and root-mean-square (RMS) error for the four studies. RMS values are in units of log10 arcmin.

Study	β	ω	ρ	RMS
Chen(2009)	1.20	1.18	1.05	0.043
Hess (1989)	1.71	1.06	1.03	0.034
Mather (2002)	1.54	1.13	1.04	0.040
Watson (2011)	1	1.15	1.02	0.055

Figure 2.6 shows the predicted depth of field JND based on the blur JND for one arcmin blur circle, calculated from the data in Table 2.2 and compared to 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 it was 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 depth of

field to the experimental JND value (N=10)) showed that for a reference depth of field of 10.1mm our experimentally determined JND in depth of field was significantly smaller than what was predicted from the blur JND, independent of which data set was used (t (9) = -11.3, p <.001; t(9) = -5.8, p < .001; t (9) = -11.7, p < .001; t (9) = -6.0, p < .001 for Chens, Hess, Mathers, and Watsons data, respectively). For a reference depth of field of 39.1mm, we found no significant difference between the experimentally determined JND in depth of field and the predicted ones.

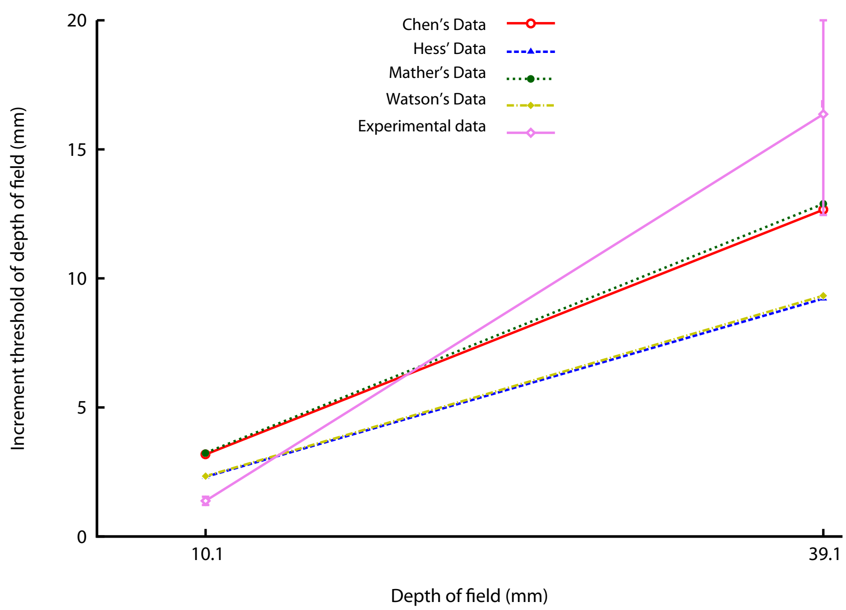

Figure 2.6: Comparing our measured JNDs with the predicted JNDs from literature.

The predicted JNDs from Chen's, Hess', Mather's, and Watson's data were quite consistent; therefore, we averaged the predicted JNDs. Figure 2.7 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 depth of field separately for the various scales of the scene values. Again, one-sample t-test analyses were performed, and we found that for a reference depth of field of 10.1mm the experimentally determined JND was significantly smaller than the predicted value, independent from the value of the scale of scene. For a reference depth of field of 39.1mm, we found something different. When the scale of the scene was 50cm, there was no significant difference between the predicted JND and the experimental JNDs. However, when the scale of the scene value was

75cm or 100cm, the experimental JNDs were significantly smaller than the predicted JNDs. The results of the t-test analyses are summarized in Table 2.3.

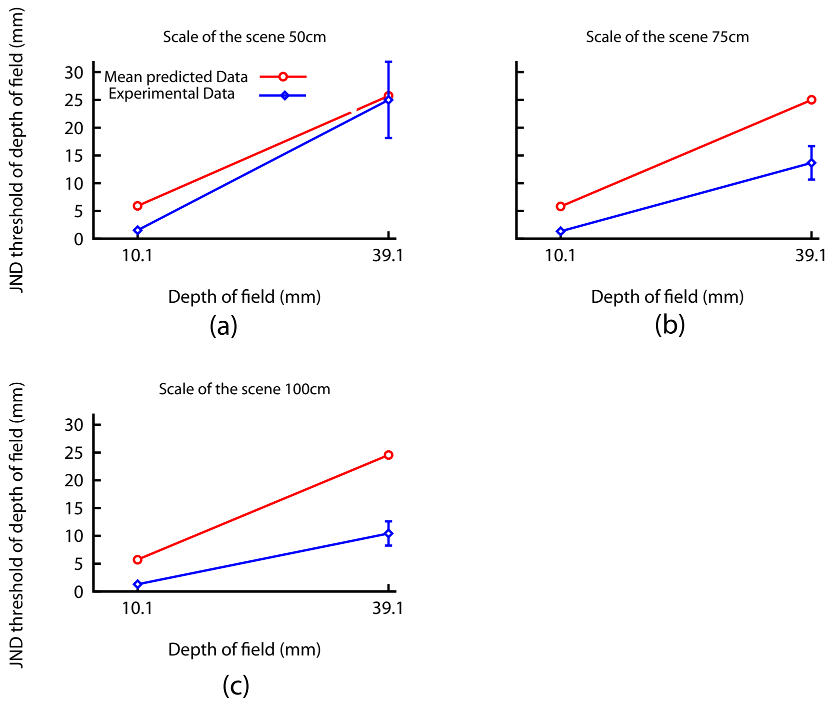

Figure 2.7: Comparison between the experimental JNDs (with ±1 standard error) and the mean predicted JNDs from the literature.

Table 2.3: T-values from one-sample t-test, comparing our experimentally measured JNDs and the mean predicted JNDs. *** refers to a significant difference at a value of $p < 0.001$, ** refers to a significant difference at a value of $p < 0.05$.

	t-value					
Reference DOF (mm)	10.1mm			39.1mm		
Scale of the Scene	50cm	75cm	100cm	50cm	75cm	100cm
Predicted JNDs	-24.89***	-27.62***	-29.51***	-	-3.77**	-6.49***

Note: $^{**}p < 0.05$; $^{***}p < .001$.

2.4.2 Fourier Analysis

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

The "Apple" scene with the maximum depth of 75cm was used as an example to show the results of Fourier analysis in Figure 2.8. The analyses for the other images were however similar. Figure 2.8(a) shows the changes in the power spectrum as a function of spatial frequency in the stimulus for a depth of field of 10.1mm and 39.1mm, 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 depths of field of 10.1mm and 39.1mm were similar to the differences between the low and high Gaussian blur levels. This might suggest that the depth of field blur is in practice similar to uniform Gaussian blur, indicating that it may be possible to use blur discrimination thresholds to predict depth of field discrimination.

The changes in contrast as a function of spatial frequency are shown in Figure 2.8(b) together with the contrast sensitivity function (CSF) (Watson and Ahumada, 2011). The contrast difference between depths of field 10.1mm and 39.1mm was above the CSF in the low frequency area, indicating that the difference between the two depths of field should be visible. Similarly, the contrast difference between depth of field 10.1mm and three of its test depths of field are presented in Figure 2.8(c), and depth of field 39.1mm with three of its test depths of field in 2.8(d). Figure 2.8(c) shows that only the difference between depths of field 10.1mm and 13.1mm was below the CSF, suggesting that the difference between the two depths of field should not be visible, which was not in agreement with our experimental data. Figure 2.8(d) shows that the difference between depth of fields 39.1mm and 62.1mm 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 depth of field of 39.1mm from depth of field of 62mm. Thus, we could argue that the predictions from blur discrimination may underestimate people's ability to discriminate depth of field blur.

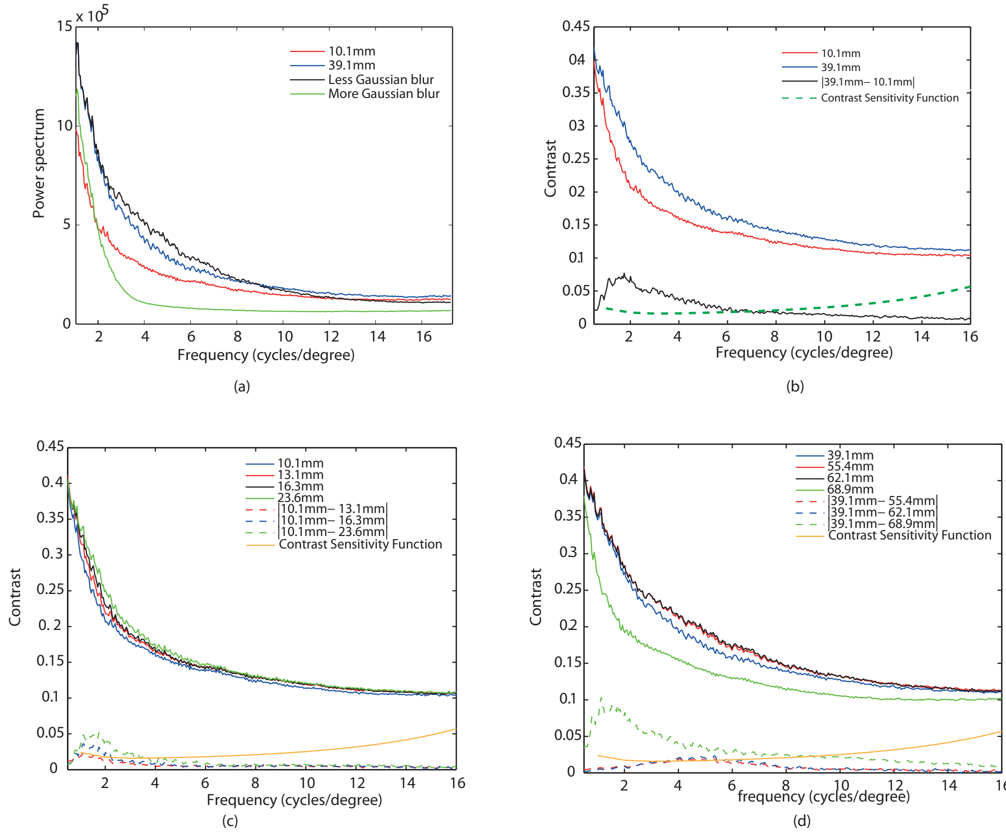

Figure 2.8: Fourier analysis on the Apple scene with the maximum depth 75cm. (a) Power spectrum as a function of frequency; (b) Contrast as a function of frequency for the reference depth of field, and the contrast difference between the reference depths of field; (c) the contrast difference between reference depth of field 10.1mm and test depth of field 13.1mm, 16.3mm, and 23.6mm respectively; (d) the contrast difference between reference depth of field 39.1mm and test depth of field 55.4mm, 62.1mm, and 68.9mm.

2.5 Discussion

The increment threshold in depth of field was measured for two reference depths of field (10.1mm and 39.1mm) 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 50cm, 75cm, or 100cm. The experiments were done under both stereoscopic and non-stereoscopic viewing conditions.

We compared the predicted depth of field discrimination with the experimental data in the modelling section. It showed that for a reference depth of field of 10.1mm, the experimentally measured JND of depth of field was smaller than the predicted values. Blur discrimination was investigated based on uniform blur, while we investigated depth of field 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 depth of field when the reference depth of field is 10.1mm. 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 depth of field JND would have been even higher. They may have used information from the combination of multiple blur levels to find the JND in depth of field. The statistical analysis suggests that people's ability to discriminate in depth of field in a photograph is better than the discrimination of any single blur level included in the photograph when the reference depth of field is 10.1mm. However this is not necessarily the case for a reference depth of field of 39.1mm. 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 depth of field at 10.1mm and 39.1mm.

Our results showed that there was no significant difference in depth of field 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 was found to significantly influence the JND of depth of field. The scale of the scene directly changes the blur gradient visible in the stimuli. According to equations 2.1 and 2.2, we can calculate the blur circle on each object in the scene. The difference in blur between the reference depth of field 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 depth of field more visible. The results suggest that photographers and movie directors may put less effort on choosing depth of field when the scale of the scene is small since people are unable to see the differences. However, when the scale of the scene is large, photographers can generate pictures with different impressions by manipulating depth of field.

Another interesting finding is that our results do not support the hypothesis that depth of field would be easier to discriminate in stereoscopic compared to non-stereoscopic images. There was no difference found between the discrimination in stereoscopic and non-stereoscopic images. Although the stereoscopic depth of field itself does not cause any discomfort (O'Hare et al., 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 et al., 2009) which in turn may decrease the ability to discriminate depth of field. 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 non-stereoscopic viewing (Heynderickx and Kaptein, 2009), etc. Therefore, we conclude that all these factors are not relevant for discrimination of depth of field and that thresholds are similar under stereoscopic viewing and under non-stereoscopic viewing.

2.6 Conclusion

In summary, we conclude that the discrimination of blur caused by depth of field differences is different from the discrimination of uniform Gaussian blur. In general, people are more sensitive to changes in depth of field 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 depth of field when the reference depth of field is small, while people are not so sensitive to changes in depth of field when the reference depth of field is large. Our research also shows no significant difference between non-stereoscopic viewing and stereoscopic-viewing on depth of field, indicating that the depth of field characteristics of stereoscopic and non-stereoscopic photographs are comparable. Additionally, we conclude that the depth structure in the scene affects observers' ability to discriminate depth of field as well.

Bibliography

Bach, M. (1996). The freiburg visual acuity test-automatic measurement of visual acuity. *Optometry & Vision Science*, 73(1):49–53.

Born, M. and Wolf, E. (1999). Principles of optics, 7th. *Cambridge Univ Pub.*

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

Chen, C. C., Chen, K. P., Tseng, C. H., Kuo, S. T., and Wu, K. N. (2009). Constructing a metrics for blur perception with blur discrimination experiments. *SPIE Proceedings of the International Society for Optical Engineering.*

Cole, F., Decarlo, D., Finkelstein, A., Kin, K., Morley, K., and Santella, A. (2006). Directing gaze in 3d models with stylized focus. *Proceedings of the 17th Eurographics conference on Rendering Techniques*, pages 377–387.

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

DiPaola, S., R. C. . E. J. (2010). Rembrandt's textural agency: A shared perspective in visual art and science.

Georgeson, M. A. (2011). Edge-finding in human vision: A multi-stage model based on the perceived structure of plaids. *Image and Vision Computing*, 16(6-7):389–405.

Hamerly, J. R. and Dvorak, C. A. (1981). Detection and discrimination of blur in edges and lines. *Journal of the Optical Society of America*, 71(4):448–452.

Hess, R. F., Pointer, J. S., and Watt, R. J. (1989). How are spatial filters used in fovea and parafovea? *J. Opt. Soc. Am. A*, 6(2):329–339.

Heynderickx, I. and Kaptein, R. (2009). Perception of detail in 3d images. *SPIE Electronic Imaging.*

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

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

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

Lambooij, M., Marten, F., Heynderickx, I., and Ijsselsteijn, W. (2009). Visual discomfort in stereoscopic displays : A review. *Journal of Imaging Science and Technology*, 53.

Liu, S., Hua, H., and 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(3):381–393.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., and Martin, K. E. (1996). Occlusion edge blur: a cue to relative visual depth. *J. Opt. Soc. Am. A*, 13(4):681–688.

Mather, G. (1997). The use of image blur as a depth cue. *Perception*, 26(9):1147–1158.

Mather, G. and Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception*, 31(10):1211–1219.

O'Hare, L., Zhang, T., Nefs, H. T., and Hibbard, P. B. (2013). Visual discomfort and depth-of-field. *i-perception*, 4(3)(3):156–169.

Otero, J. M. (1951). Influence of the state of accommodation on the visual performance of the human eye. *J. Opt. Soc. Am.*, 41(12):942–947.

Paakkonen, A. K. and Morgan, M. J. (1994). Effects of motion on blur discrimination. *J. Opt. Soc. Am. A*, 11(3):992–1002.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9 (4)(4):523–531.

Rogers, S. (1995). Perceiving pictorial space. *Perception of space and motion*, 5.

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

Watt, R. J. and Morgan, M. J. (1983). The recognition and representation of edge blur: Evidence for spatial primitives in human vision. *Vision Research*, 23(12):1465–1477.

Watt, S. J., Akeley, K., Ernst, M. O., and Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, 5(10):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., and Westland, S. (2001). Blur tolerance for luminance and chromatic stimuli. *J. Opt. Soc. Am. A*, 18(6):1231–1239.

Chapter 3

Depth Perception

Depth of field affects perceived depth in stereographs

This chapter has been published as:

Depth of field affects perceived depth in stereographs (2015) *Tingting Zhang, Louise O'Hare, Paul B. Hibbard, Harold T. Nefs, Ingrid Heynderickx,* ACM Transactions on Applied Perception, *11(4), No.18*

Although it has been reported that depth of field influences depth perception in non-stereo photographs, it remains unclear how depth of field affects depth perception under stereo viewing conditions. We showed participants stereo photographs with different depths of field using a Wheatstone stereoscope and a commercially available 3D TV. The depicted scene contained a floor, a background, and a measuring probe at different locations. Participants drew a floor plan of the depicted scene to scale. We found that perceived depth decreased with decreasing depth of field for shallow depths of field in scenes containing a height-in-the-field cue. For larger depths of field, different effects were found depending on the display system and the viewing distance. There was no effect on perceived depth using the 3D TV, but perceived depth decreased with increasing depth of field using the Wheatstone stereoscope. However, in the 3D TV case, we found that the perceived depth decreased with increasing depth of field in scenes in which the height-in-the-field cue was removed. This indicates that the effect of depth of field on perceived depth may be influenced by other depth cues in the scene, such as height-in-the-field cues.

3.1 Introduction

Depth of field (DOF) is the depth range around the focal plane that is perceived as sharp. All optical lens systems, including the human eye and digital photo cameras, generate images with a limited depth of field, as long as they are focused nearer than the hyper focal distance. Depth of field depends on the size of the lens aperture relative to the focal distance of the imaging devices. The amount of blur of an object in the scene depends on the distance of that object to the focal plane and can be expressed as:

$$b = A\frac{s_0}{d_0}|1 - \frac{s_0}{d_1}| \quad (3.1)$$

where b is the diameter of the blur circle, i.e., the diameter of the area on the projection plane over which a point is spread out, A is the diameter of the lens aperture, s_0 is the distance from the lens to the image plane, d_0 is the distance from the lens to the focal plane, and d_1 is the distance from the focal plane to an object in the scene. A different form of this equation can be found in (Nefs, 2012). Hence, depth of field can be considered as the distance range in which the blur circles remain smaller than the smallest amount of perceptible blur. Equation 3.1 shows that the blur circle b at a certain depth distance d_1 increases as the focal distance d_0 decreases on condition that the relative distance d_0/d_1 is fixed. In other words, when the focal distance decreases, depth of field becomes smaller as well. However, if the focal distance d_0 is fixed, the depth distance d_1 can be directly estimated from the blur circle based on Equation 3.1, assuming that A, d_0 and s_0 are known. When an optical system has a small depth of field, the blur circle b will rapidly increase with increasing distance from the focal plane, and so, people will perceive a large blur gradient in the resulting images.

Depth of field is an important aspect of vision, and there has been significant research on the effects of depth of field in non-stereo photographs over the last few years. For example, it has been shown that depth of field can be used to direct viewers' attention (Baveye et al., 2012; Hamerly and Dvorak, 1981; Hillaire et al., 2008; Khan et al., 2010), and it may also influence the aesthetic appreciation of images (Datta et al., 2006). Other research has directly assessed its effects on the perceived distance from the observer to objects and the perceived distances between objects in the scene (Held et al., 2010; Mather, 1996; Mather and Smith, 2000; Pentland, 1987; Vishwanath and Blaser, 2010; Watt et al., 2005).

In the current study, we assess the effect of depth of field on the perception of depth in stereo photographs. A strong perception of depth can be created using stereoscopic presentation. Stereoscopic images and videos are widely used in many fields such as cinematographic industries, virtual reality, visualization and computer assisted design (Hubona et al., 1999). However, the challenge of delivering a high-quality, comfortable experience of depth perception using current display technology remains an active and important area of research (Banks et al., 2012; Shibata et al., 2011). We investigated the extent to which depth of field can be used as a cue to depth that complements that provided by binocular disparity. To do this, we investigated the effects of depth of field on depth perception in stereoscopically presented photographs, using two different presentation methods.

3.1.1 Depth of Field in Non-Stereo Photographs

Previous research has already investigated the effect of depth of field on perceived depth in non-stereo photographs. Before considering the results of these experiments, it is important to appreciate the different ways in which depth of field may affect the perception of depth. In the first instance, let us assume that the viewing distance is known, and that the parameter d_0 in equation 3.1 is therefore fixed and known. In this case, the depth of field could be calculated, and the amount of blur could be used as a cue to depth. Thus, a small blur circle would be interpreted as the result of viewing an object with a small deviation in distance from the focal plane (i.e. d_1 is similar to d_0). The consequence of decreasing the depth of field, and increasing the amount of blur in the scene, would in this case be to increase the magnitude of apparent depth. In this paper, the term "depth" refers to the distance between two objects in the visual direction, whereas "distance" refers to the egocentric distance (i.e., the distance from the observer to the object).

We can also consider the situation in which the viewing distance is not known, and d_0 in equation 3.1 is a free parameter. In this case, depth-of-field blur can be used to estimate the viewing distance. In equation 3.1, d_0/d_1 can be obtained from other perspective cues such as relative size. This implies that the blur circle is inversely proportional to the focal distance d_0 (see Equation 3.1). Hence, a large blur gradient caused by a small depth of field is consistent with a short viewing distance. Note that, if depth-of-field blur is used to estimate viewing distance in this way, it may also be expected to affect apparent size. Thus, a large blur gradient, indicating a close-by viewing distance, would be expected to create a miniaturization effect through scaling of retinal information to estimate object size. Depth-of-field blur could thus be used to provide a cue both to the relative depth of objects in the scene, and as a way of estimating the egocentric distance to the scene, and thus as a way of scaling relative cues so as to provide absolute depth information. Previous research has shown both effects of depth of field in non-stereo photographs, depending on the availability of other cues to egocentric distance.

Nefs (2012) demonstrated the influence of depth of field on the perceived depth/width relationship in non-stereo photographs of natural scenes. In his study, the photographs contained a central figurine with a flanking figurine at each side. The observers were asked to move the flanking figurines to such a position that the distance between them and the depth to the central figurine were the same. Five levels of depth of field were used by manipulating the camera aperture. He found that perceived depth, relative to perceived width, increased as the depth of field became smaller. In these experiments, the viewing distance was well-defined, since the focal plane was always on the computer screen. Thus, the magnification of relative depth with decreasing depth of field found was consistent with the use of depth-of-field blur as a relative depth cue. Held et al. (2012) have also argued that blur might be used as a cue to depth in this way, and showed that under some circumstances it can be more reliable than binocular disparity.

Depth-of-field blur can also affect observers' estimates of egocentric distance. For example, Held et al. (2010) asked participants to estimate the distance from a target in the scene to the camera. They found that the reported camera distance was smaller for photographs with larger blur gradients than for photographs with smaller blur gradients. Vishwanath and Blaser (2010) found that perceived egocentric distances were smaller for larger blur gradients. This effect was observed both when observers were estimating pictorial depth in a scene, and when they were explicitly asked to judge the distance to the display screen. Although neither of these studies directly assessed the 'miniaturization' effects that can occur when the depth of field is reduced (e.g. in tilt-shift photography), both argue that such effects can be attributed to the reduction in apparent egocentric distance.

3.1.2 Stereoscopic Viewing

Binocular disparity provides an important depth cue in stereo displays. Disparity provides precise depth information for points that are close to where one is fixating (Held et al., 2012). However, using disparity as a cue to depth in current displays can create a conflict between distance information signaled by convergence and accommodation (Hoffman et al., 2008). This conflict usually causes visual discomfort, particularly when there are frequent, large motor responses of the eyes as a consequence of a large disparity (Lambooij et al., 2009). Limiting the size of disparity in order to decrease discomfort, has the inevitable consequence that the magnitude of depth that can be signaled will be limited. Previous investigations have suggested that depth-of-field blur may be used to compensate for the loss of perceived depth on 3D displays that is caused by restricting the range of disparities presented (Held et al., 2012; Wang et al., 2011). Wang et al. (2011) investigated the effect of combining blur and disparity on perceived depth in 3D images. They found that image blur (created using a Gaussian filter) in stereo images contributed to the impression of perceived depth. These results suggest that depth-of-field blur may have the same effect on perceived depth in stereo photographs as it does in non-stereo photographs (Nefs, 2012).

Wang et al. (2011) used a two-alternative forced-choice task to investigate the apparent depth separation between a foreground target object and a background fronto-parallel plane. The target object was a photograph of a butterfly, and the background was a photograph of a flowerbed. They showed that the perceived depth between the target and the blurred background matched that of a sharp target and a sharp background when the former had a smaller disparity than the latter. Thus, the blur in the background contributed to apparent depth. Our current study assesses whether depth-of-field blur in scenes with complex depth information has similar effects on perceived depth. In Wang et al. (2011), depth information from both disparity and blur were rather simplified. For both the background and the target, the disparity and blur were constant, indicating a fronto-parallel structure. The only disparity and blur differences present were those between the background and target. This means that neither disparity nor blur depicted a complex depth structure of a type that might be expected in natural scenes. Thus, by using a uniform Gaussian blurring of the entire background, this study did not replicate the depth-of-field blur gradients found in natural images.

In the current work, we investigate the effect of optically created depth of field of different levels on perceived depth in stereo images. There are a number of reasons to believe that the effect of depth of field on perceived depth is not necessarily the same for stereo depth as for non-stereo depth. Firstly, the subjective experience of depth in stereo photographs is qualitatively different from that in non-stereo photographs. In non-stereo photographs, pictorial space does not appear to occupy the same physical space as in stereo images (Rogers, 1995). By contrast, stereo images typically create an experience of stereopsis, which is the sense of a 3D space rather than a flat pictorial space. Importantly, the perceived space appears to be physically embedded in the real world (Vishwanath and Hibbard, 2013). Secondly, the optical state of the eyes might be more tightly linked to depth of field in stereo photographs than in non-stereo photographs. In the latter case, optical accommodation and binocular convergence typically provide cues to the location of the pictorial surface, but not to the objects in the scene. Indeed, when the link between accommodation and convergence is weakened, a greater sense of stereopsis is experienced even when viewing non-stereo images (Vishwanath and Hibbard, 2013) and the sense of spatial presence is enhanced (Ling et al., 2013). This suggests that the effect of depth of field on perceived egocentric distance is likely to be minimal because stereo photographs contain relatively strong depth cues. This means that the depth miniaturization effects discussed by Held et al. (2010) and Vishwanath and Blaser (2010) are unlikely to occur in stereo photographs.

By contrast, if depth-of-field blur provides a relative depth cue which is then scaled by an estimate of egocentric distance, we may expect the degree of apparent depth to increase with decreasing depth of field, as found by Nefs (2012). At the same time, we acknowledge that the processing of binocular disparity may be impaired with a small depth of field, which may lead to somewhat unpredictable results. That is, the introduction of blur by a small depth of field will tend to reduce the high spatial frequency content in part of the image, and so, will reduce the quality of depth information available from disparity. It is well known that the human binocular stereo system is relatively insensitive at low spatial frequencies (Arndt et al., 1995). This means that an increase in blur, while in itself acting to increase apparent depth, may have the effect of reducing the magnitude of depth perceived from binocular disparity. This loss of binocular disparity may either lead to an increase or a decrease in perceived distance depending on the distances specified by other depth cues.

3.1.3 Current Study

In the current work, we measured the perceived position of a probe in a photograph of a scene whilst varying the degree of depth of field. Observers were asked to adjust the position of a small rectangle in an experimental interface corresponding to the position of the probe in the scene. A reference distance was provided to match the distance in the real world with the distances in the experimental interface. Two experiments were performed, in which the stereoscopic images were presented using firstly a Wheatstone stereoscope, and secondly a commercially available 3D television with passive shutter glasses. The Wheatstone stereoscope is a traditional stereo display which can be used to present high-quality stereo stimuli and it is a very tightly controlled presentation method that allows for comparison with previous laboratory research. The 3D TV, on the other hand, is a new technology which is becoming more common in home entertainment systems, and considered important to investigate the effects of depth of field in a real application. Two display devices were used, firstly to assess how easily potential effects of depth of field could be replicated across different viewing systems, and secondly to allow us to vary the viewing distance: as the 3D TV was much larger than the Wheatstone stereoscope, the viewing distance could also be much larger. According to geometrical considerations and previous research (Gooding et al., 1991; Ritter, 1977; Watt et al., 2005), we predicted that the perceived depth on the 3D TV would be larger than on the stereoscope, due to the larger viewing distance combined with the same angular disparities in the two cases. Since the stimuli in these two experiments contained another important depth cue, that of height-in-the-field cue (He and Ooi, 2000), we conducted an extra experiment on the 3D TV, using stimuli without this height-in-the-field cue to investigate the influence of other depth cues on the effect of depth of field on perceived depth.

3.2 Experiment

3.2.1 Participants

Sixty participants, aged between 21 and 35 years, with normal or corrected-to-normal visual acuity (measured with the Freiburg visual acuity test) and normal or better-than-normal stereoacuity (measured with the TNO stereo test), took part in the experiments. All participants were naive to the hypotheses of the experiments. This research was approved by the Ethics Committee of the Delft University of Technology, and was in accordance with the 1969 Declaration of Helsinki.

3.2.2 Apparatus

A Wheatstone stereoscope with two 19 inches MM904UT Iiyama CRT monitors and front surface silver-plated mirrors was used in the first experiment to present the stereo photographs. The path length between the eyes and the screens was 70 cm. The two monitors were calibrated with a ColorMunki spectrophotometer such that their luminance and color responses were identical. The second experiment was a replication of the first, using a TX-P65VT30E Panasonic 65 inches plasma 3D TV with passive shutter glasses. We used the sidebyside 3D mode as the stereo signal format. All stimuli were presented using a Mac Pro computer running OSX 10.6. An IBM ThinkPad was used as an interface to collect the participants' responses. The experimental interface was presented using MATLAB R2011b.

3.2.3 Stimulus

Stereo photographs were taken with an Olympus E-440 d-SLR camera with an Olympus Zuiko 50-mm macro lens. An example of the stereo photographs is shown in Figure 3.1. The aperture of the lens could be adjusted from F2 to F22. In our experiments, F2, F3.5, F6.3, F11 and F22 (lens aperture = focal length / F_value = 25, 14.3, 7.9, 4.5, and 2.3mm respectively) were used to create different depths of field. The visual angle of the camera was 13.2° (horizontally) × 9.9° (vertically). The size of the stimuli was constrained by the visual angle of the camera and the dimensions of the display systems. The visual angle of the stimuli on the stereoscope and the 3D TV was adjusted to the visual angle of the camera. In order to keep the visual angle identical for the Wheatstone stereoscope (with a screen diameter of 19 inches) and the 3D TV (with a screen diameter of 65"), the observers were seated at the center of projection and the viewing distance was 70cm for the stereoscope and 300cm for the 3D TV (accounting for a factor of 4.3 difference in horizontal size). We took stereo photographs of the scene with a stereobase of 6.5 cm (i.e., averaged horizontal distance between the eyes in adults) for presentation on the Wheatstone stereoscope. For presentation on the 3D TV the stereobase of the cameras was reduced by the same scaling factor (i.e., 4.3) with which the viewing distance was increased in order to display the same angular disparity. As such, the stereobase of the cameras became 1.5 cm for presentation on the 3D TV. In our stimuli, the focus objects were on the same level as the camera and the camera was toed-in. As a consequence, incorrect vertical disparities might have been introduced when the stimuli were projected, which might have affected the perception of depth (Banks et al., 2012). Although the target and flanker objects themselves were close to fixation, and therefore unlikely to provide usable vertical disparity information in themselves, it was important to also consider the fact that vertical disparities in the surround might also have affected the perceived distance of the targets (O'Kane and Hibbard, 2007). We calculated the maximum error of the vertical disparities between the real scene and the images and found it to be relatively small (less than 0.5 arc min). Moreover, since the overall extent of the projected images was considerably less than 20 degrees, vertical disparities were unlikely to be effective (Bradshaw et al., 1996). Although we cannot rule out any such effects, the vertical disparities present in our stimuli were not affected by the depth of field.

The stereo stimuli used in our experiments consisted of a colorful background containing a broad range of spatial frequencies, made up of flowers and leaves, a

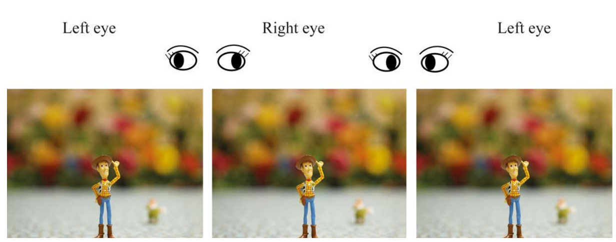

Figure 3.1: Examples of stereo photographs. Readers capable of free fusion of stereo photographs may use the left and centre photographs for uncrossed free-fusion, and the centre and right photographs for crossed free-fusion.

ground plane and two objects standing on the ground plane. One of the objects was a reference (Woody from the movie "Toy Story"), positioned in the center of the scene at zero parallax. The other object was used as the probe, which was positioned at one of the fourteen positions behind Woody on the ground plane (see Figure 3.2). When making the stereo photographs, the camera was always aimed at Woody's nose for both the left and right half-images. The distance between the camera and Woody was 70cm. Three probes were used, namely a Snow White (SW), a Dwarf (DW) and a Green Pushpin (GP) figurine.

3.2.4 Procedure

We used a between-group design in this experiment. Thirty observers participated in the experiment using the stereoscope, and another thirty observers participated in the experiment with the 3D TV. The experimental procedure was identical for both groups. Observers were seated in a room with the only direct light coming from the monitors of the Wheatstone stereoscope/the 3D TV and the IBM ThinkPad. For the Wheatstone stereoscope, participants used a chin rest in front of the mirrors. For the 3D TV, observers wore shutter glasses but their head movements were not constrained. Next to the participants was the IBM ThinkPad on which they could indicate the position of the probe as they perceived it, relative to their own sitting position and the reference object. In the experiments, we did not give explicit instructions to observers on how to

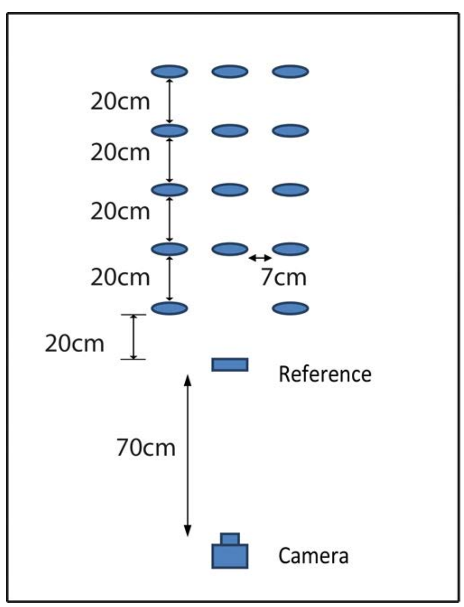

Figure 3.2: Layout of 14 locations.

scale the Z dimension to a setting in the Y dimension, except that the positions of the observer and of the reference object were given in the interface. As such, the settings in the Y dimension were assumed to be proportional to the scale in the Z dimension. Figure 3.3 shows the experimental interface. It contained a small image of a cartoon human and of "Woody" to represent the positions of participant and reference object respectively in the real world. Observers were asked to drag the small yellow square representing the probe into the position where they thought it should be, using the distance between the observer and the reference object as a guide. The observers had two degrees of freedom in moving the target object when doing the task. They could move the target object either in the vertical direction to represent the depth or in the horizontal direction to represent the width. They were also free to look at wherever they wanted in the stereo photographs, and the viewing time was unlimited.

When depth of field was small, it was sometimes difficult to discern the probe in the stereo photographs. Furthermore, the probe was also occluded by the reference object in some positions. This latter situation was more often the case for the 3D TV, because the stereobase of the camera was smaller for the stereo photographs displayed on the 3D TV than on the stereoscope. For this reason, the interface incorporated a "cannot see" button. There were 210 trials in total, including five levels of depth of field, fourteen probe positions and three probes. Each trial was presented only once. The order of the trials was randomized anew for each participant. The experiment took around 30 minutes for each participant.

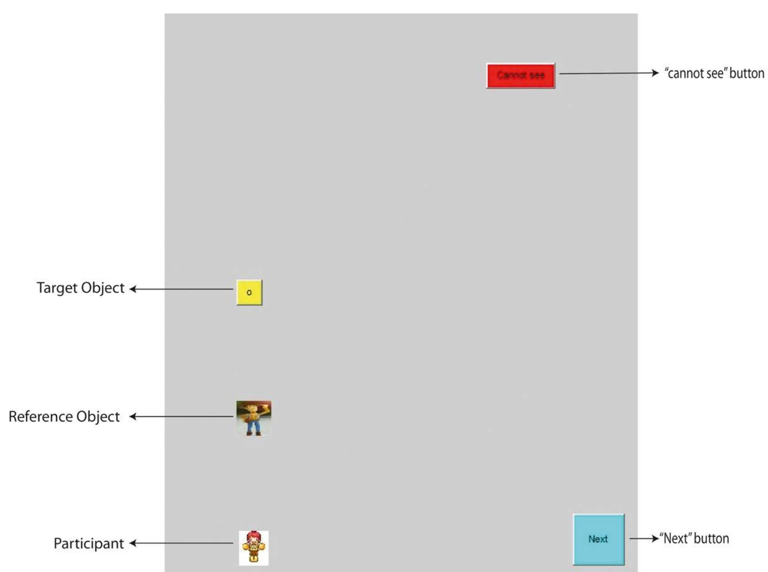

Figure 3.3: The experimental interface.

3.3 Comparison Between the Stereoscope and the 3D TV

Since two stereo systems were used in our study, it is necessary to be clear about the differences between the systems. Table 3.1 summarizes the parameters for the two setups. Under these circumstances, the horizontal angular disparity between the flanker objects and the focus object was the same when observed through the stereoscope and the 3D TV (e.g. the angular disparity between the first right flanker object and Snow White was 0.279 degrees on both the stereoscope and the 3D TV). As shown in Table 3.1, the visual angle subtended by the stimuli was the same on the two displays. Hence, the rendered volume in the visual space was constant on the two displays, which means that not only the size of the stimuli was scaled up, but also the stereoscopically presented depth was scaled up proportionally. Thus, the combination of a greater viewing distance with the same angular size of images may be expected to lead to a larger perceived depth on the 3D TV.

Table 3.1: Comparison Between the Stereoscope and the 3D TV

	Stereoscope	3D TV
Stimulus Stereobase	6.5cm	1.5cm
Stimulus Size (on the screen)	16.2cm × 12.1cm	69.4cm × 52cm
Viewing Distance	70cm	300cm
Visual angle	13.2 × 9.9	13.2 × 9.9

3.4 Results

Before proceeding with the main analyses, we calculated how well each probe could be seen at each position. If more than half of the participants indicated they could not see the probe in a particular position, then that particular condition was defined as "could not be seen" and was excluded from the analysis. In the experimental interface, the vertical/horizontal distance between the yellow square and the reference object was recorded in units of pixels. Given the distance between the participant and reference object, we could calculate the equivalent relationship between perceived depth/width in the experimental interface in pixels and perceived depth/width in the real world in centimeters. Therefore, perceived depth/width refers to the vertical/horizontal distance between the yellow square and the reference object in centimeters in the interface, while physical depth/width refers to the distance in depth/width between the probe and the reference figurine in the real world in centimeters.

Figures 3.4 and 3.5 show the stereoscopically presented positions (red dots, full lines) and the corresponding perceived positions (blue dots, dashed lines, averaged across 30 participants) of all three probes in the images with five levels of depth of field presented on the stereoscope and on the 3D TV. The black dots represent the stereoscopically presented positions where the probes could not be observed. The stereoscopically presented positions on the stereoscope were the same as the physical positions in the real world. The stereoscopically presented positions on the 3D TV, however, were scaled (with a factor of 4.3) to account for the larger viewing distance, keeping the angular dimensions of the rendered volume in visual space equal for both display systems. Note that the red dots in Figure 3.5 are not exactly a factor 4.3 more apart than the red dots in Figure 3.4, because of scaling to fit the graphs on one page. This scaling, however, did not affect the relative positions between dots.

It is obvious from Figures 3.4 and 3.5 that the positions of the probes observed through the stereoscope were scaled up in both the horizontal (width) and the vertical (depth) direction, whereas on the 3D TV the perceived positions were scaled up in the horizontal direction only and very close to the stereoscopically

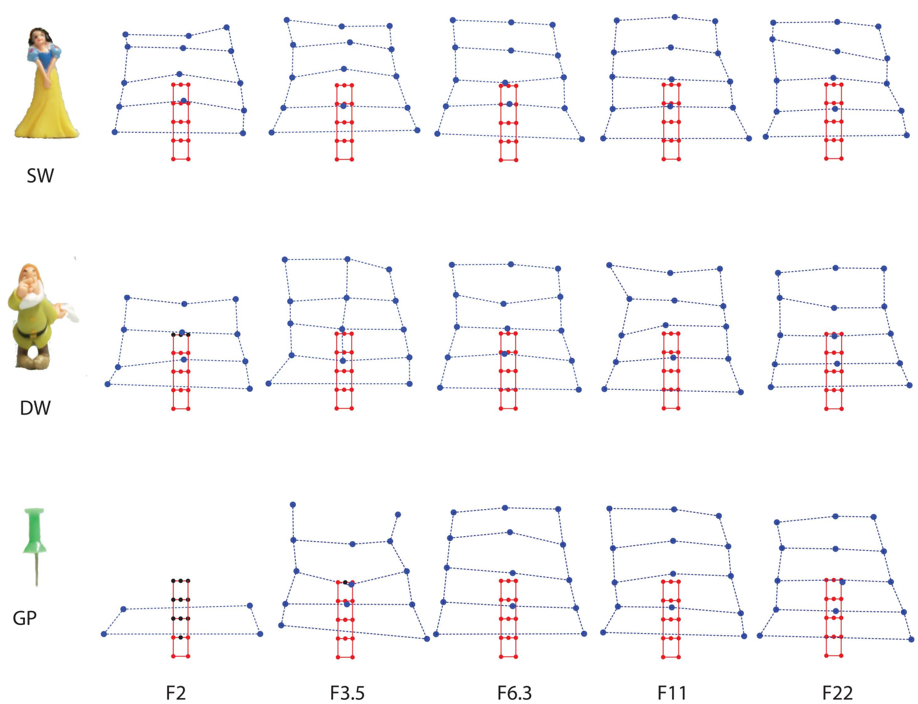

Figure 3.4: Layout of the mean perceived positions per figurine on the stereoscope. The red lines indicate the relationship of real world objects used to make the photographs, and the blue lines the average estimates of observers for the three probes.

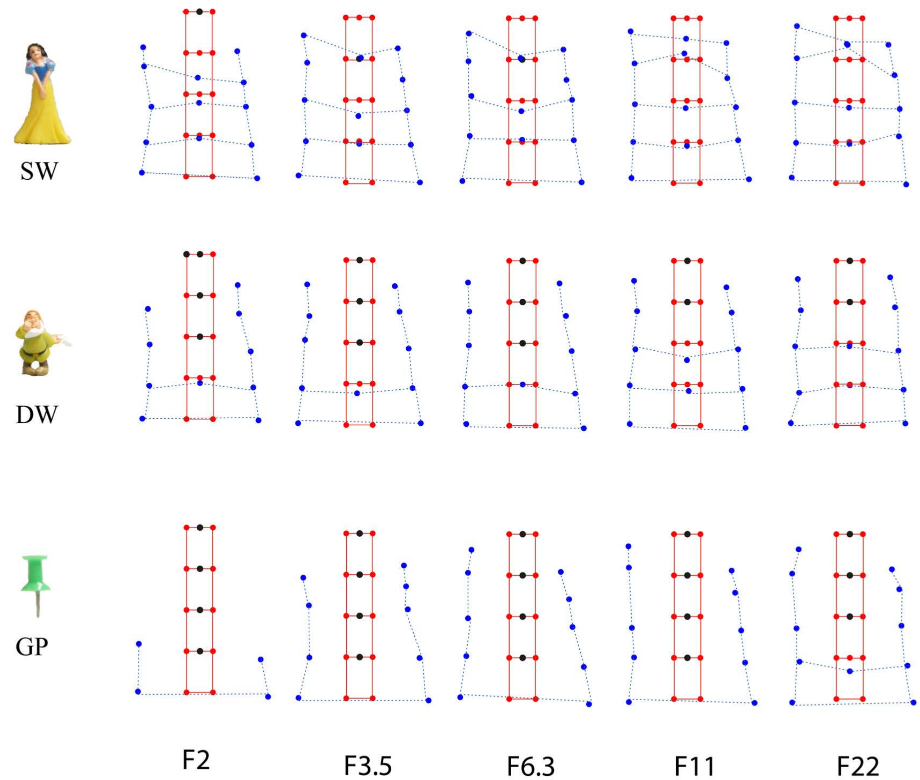

Figure 3.5: Layout of the mean perceived positions per figurine on the 3D TV. The red lines indicate the relationship of real world objects used to make the photographs, and the blue lines the average estimates of observers for the three probes.

presented positions in the vertical direction. To better compare the results on both the stereoscope and the 3D TV in terms of perceived depth, Figure 3.6 shows the relationship between stereoscopically presented depth and perceived depth. The figure clearly illustrates that the perceived depth on both the stereoscope and the 3D TV increases with increasing stereoscopically presented depth. Additionally, perceived depth is very close to the stereoscopically presented depth on the 3D TV (i.e. the data fall very close to the diagonal of the graph), whereas observers overestimate the presented depth on the stereoscope.

In the following analyses, we use the ratio of perceived depth over physical (presented) depth to describe the effect of depth of field and physical depth on depth perception. Here, physical depth represents the depth between the objects in the real world. Since the perceived depth is inclined to increase with increasing physical depth, it is better to use the ratio of perceived depth over physical depth to describe the effect of depth of field on perceived depth compared with just the value of perceived depth. The ratio of perceived depth/physical depth is averaged over all three probe conditions and also across positions on the same

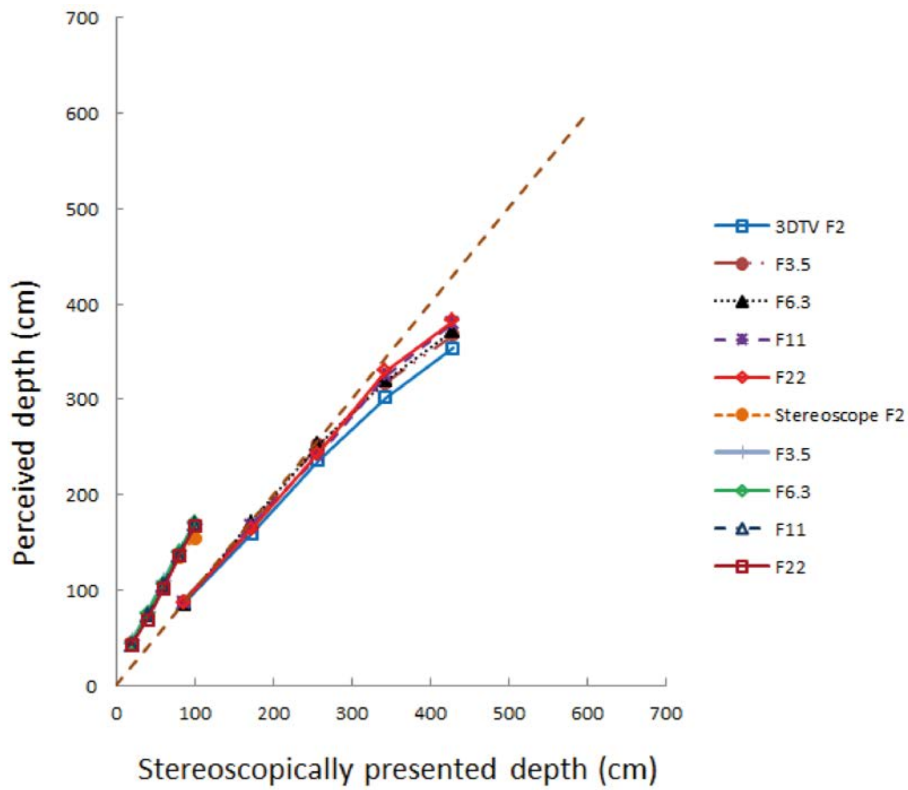

Figure 3.6: The relationship between perceived depth and stereoscopically presented depth for all depths of field on both the stereoscope and the 3D TV. Relative depth is the stereoscopically-indicated distance in depth between the focus object and the flanker. Perceived depth is the average estimated depth made by all the observers.

row excluding the missing data. In the rest of the paper we refer to this average, when using the ratio of perceived depth/physical depth.

For stimuli presented on the stereoscope, the ratio of perceived depth/physical depth across all conditions for all participants ranged from 0.56 to 5.53. Figure 3.7(a) shows that, for stimuli viewed on the stereoscope, perceived depth decreased with increasing depth of field from F3.5 to F11, while perceived depth increased from F2 to F3.5. Figure 3.7(b) illustrates that the ratio of perceived depth/physical depth decreased with increasing physical depth from 20cm to 100cm. Figure 3.7(c) demonstrates that the ratio of perceived depth/physical depth decreased with increasing physical depth for all the five levels of depth of field. We performed a 5 (depth of field) × 5 (physical depth) repeated measures ANOVA and found main effects of depth of field ($F(4, 116) = 4.44$, $p = .002$), and physical depth ($F(4, 116) = 39.35$, $p < .001$). A significant interaction between depth of field and physical depth was also found ($F(16, 464) = 2.98$, p < .001). The results of post hoc pairwise comparison t-tests are also shown in Figure 3.7, with the asterisk indicating whenever there was a significant difference between a pair of depth of field levels.

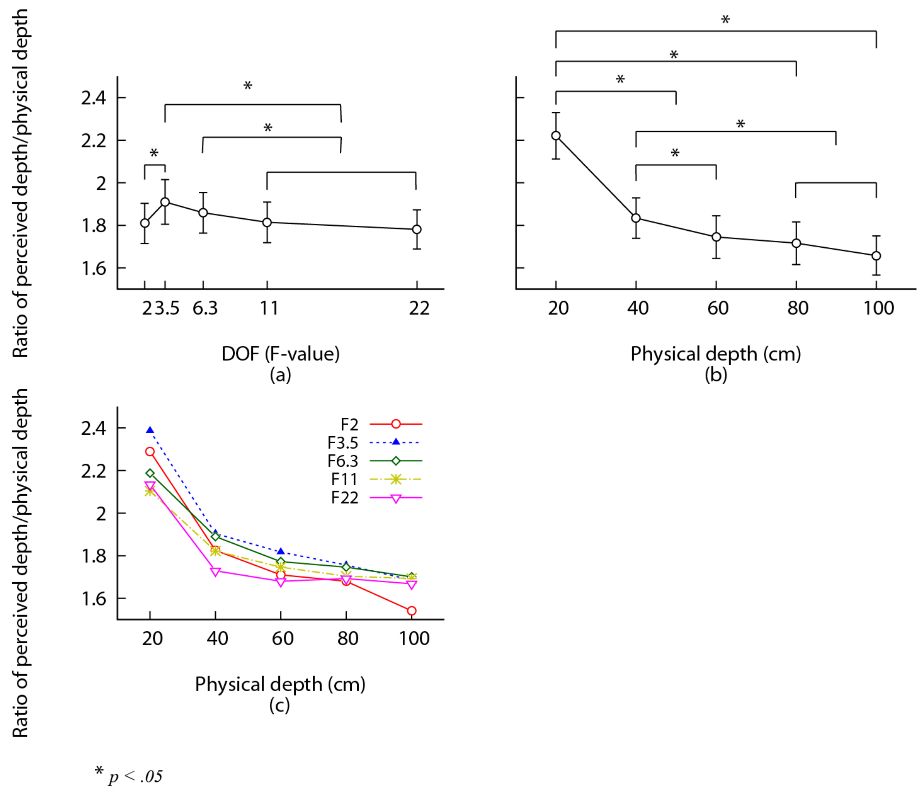

Figure 3.7: Mean values of ratio of perceived depth/physical depth as measured on the stereoscope based on a 5 (Depth of field) x 5 (Physical depth) experimental design. Error bars show ±1 standard error of the mean. The ratio of perceived depth/physical depth as a function of (a) depth of field collapsed across two probes and four physical depths. (b) physical depth collapsed across two probes and five depths of field. (c) physical depth collapsed across probes, but for the various depths of field separately.

For stimuli presented on the 3D TV, the mean ratio of perceived depth/physical depth across all conditions ranged from 1.74 to 7.74. The data are summarized in Figure 3.8. On visual inspection, Figure 3.8 (a) shows that perceived depth at depth of field F2 was smaller than at all other depths of field, and that perceived depth was not affected by depths of field larger than F3.5. Figure 3.8(b) illustrates that perceived depth, relative to width, decreased with increasing physical depth. Finally Figure 3.8(c) shows that the ratio of perceived depth/physical depth generally decreased with increasing physical depth. We performed a 5 (depth of field) × 5 (physical depth) repeated measures ANOVA, and found main effects of depth of field (F (4, 116) = 10.87, p < .001), and physical depth (F (4,116) = 8.77, p < .001). A significant interaction between the depth of field and physical depth was also found (F (16,464) = 2.55, p < .001). Figure 3.8 also shows the results of post hoc pairwise comparison t-tests, the asterisk representing that there was significant difference between each of the pairs of depth of field values.

3.4.1 Additional Control Experiment:Removing the height-in-the-field cue in the stimuli on the 3D TV

The results reported above showed that depth of field did not influence perceived depth in the range between F3.5 and F22. When preparing the stimuli, the cameras were focused and converged on the central figurine from a greater

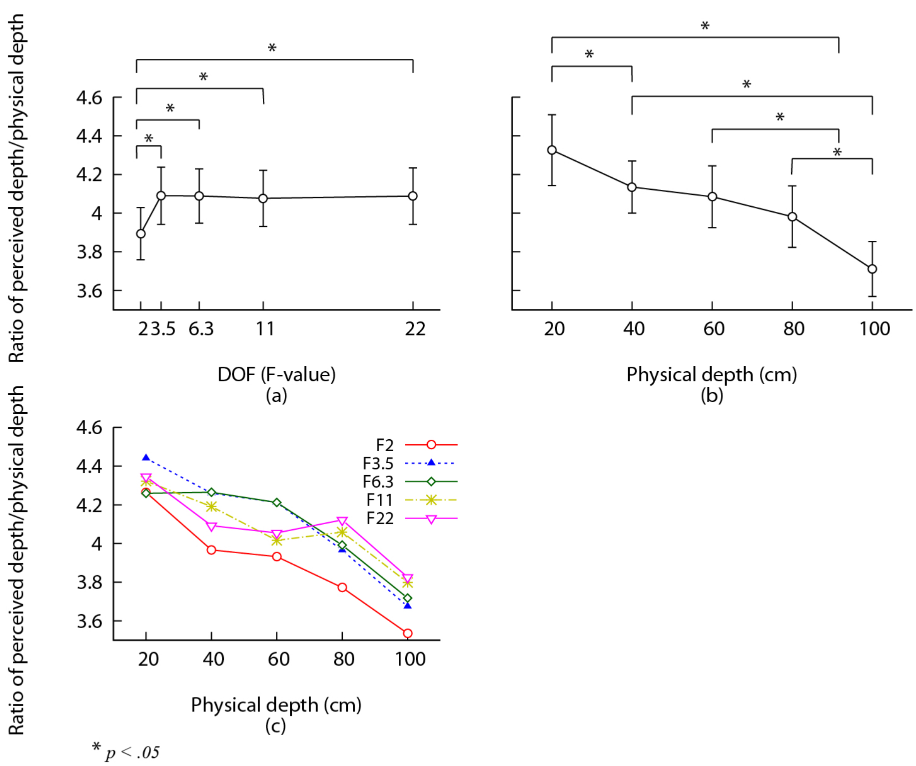

Figure 3.8: Mean values of ratio of perceived depth/physical depth as measured on the 3D TV based on a 5 (Depth of field) x 5 (Physical depth) experimental design. Error bars show ±1 standard error of the mean. Ratio of perceived depth/physical depth as a function of (a) depth of field collapsed across all probes and physical depths. (b) physical depth collapsed across all probes and depths of field. (c) physical depth collapsed across depth of field, but for the two probes separately.

height above the ground surface than the nose. As a consequence, the distance from the probe to Woody could to some degree be estimated from its height in the visual field (see Figure 3.9(a)). Literature shows that the surface ground can help observers to make accurate distance judgments in stereoscopic viewing conditions (He and Ooi, 2000). It is possible therefore that the effect of depth of field on perceived depth was weakened in our experiment because of the height-in-the-field cue, or surface ground information. Consequently, we hypothesized that depth of field may have a stronger influence on perceived depth in stereoscopic photographs without a height-in-the-field cue (e.g. Figure 3.9(b)) than in photographs with a height cue (e.g. Figure 3.9(a)).

Figure 3.9: (a) An example stimulus used in the main experiment containing the height-in-the-field cue, (b) the same stimulus used in the control experiment but now without a height-in-the-field cue.

To investigate this hypothesis, we created some new stimuli, as shown in Figure 3.9(b) with the influence of the surface ground occluded. The experimental setup was the same as for the experiment on the 3D TV described in Section 3.2.2 and 3.2.3. In this additional control experiment, we only used the "Snow White" probe, but still for the 14 positions and 5 different levels of depth-of-field blur. To remove the height cue, we set the camera at the height of the ground surface, taking care that "Woody", "Snow White" and the background were at the same height in the photograph (see Figure 3.9(b)). To make the stimuli consistent with the previous ones, we cropped all the images so that the differences in height in the visual field between objects were removed. 17 naive observers participated in the experiment with a mean age of 27.6 years old (SD = 4.5). The experimental procedure was again the same as described for the main experiment on the 3D TV.

Figure 3.10 shows that the ratio of perceived depth/physical depth decreases with increasing depth of field. This means that objects in the image with more blur are perceived as further away than the same objects in images with less blur, showing that depth-of-field blur can act as a depth cue in stereoscopic photographs. We performed a 5 (depth of field) $\times$ 5 (physical depth) repeated measures ANOVA and found a significant effect of depth of field ($F(4,64)=3.76$, $p= .008$) and of physical depth ($F(4, 64)=36.49$, $p < .001$) on perceived depth. The effect of depth of field from F2 to F3.5 in this experiment is opposite to what we found in the stimuli with the height cue in the visual field. The effect of depth of field from F3.5 to F22 turns out to be significant and consistent with what we found on the stereoscope, for photographs with a height- in-the-field cue in the visual field.

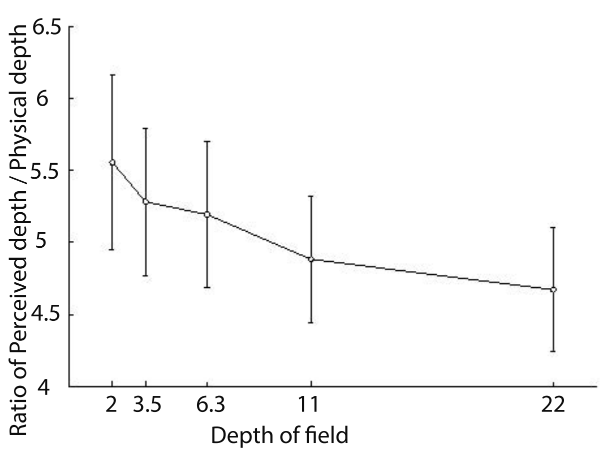

Figure 3.10: Ratio of perceived depth/physical depth as a function of depth of field in the stimuli without height cue collapsed across physical depths.

3.5 Discussion

We have shown that depth of field may affect perceived depth in stereo photographs, using a stereoscope at a closer viewing distance, or a 3D TV at a larger viewing distance. For both systems, the individual range of the ratio of perceived depth/physical depth for all participants under all conditions was large, i.e., from 0.56 to 5.53 on the stereoscope and from 1.74 to 7.74 on the 3D TV. Secondly, perceived depth increased with increasing depth of field from F2 to F3.5 on both systems, when the height-in-the-field cue in the images was congruent with the depth of field. Thirdly, perceived depth, as a proportion of physical depth, decreased with increasing physical depth on both systems.

We predicted that, if blur is used as a depth cue, then decreasing the depth of field would result in an increase in perceived depth. In the main experiments, our results were partly consistent with this hypothesis. Perceived depth decreased as a function of depth of field beyond F3.5 on the stereoscope. For depth of field between F2 and F3.5, however, our results contradicted our hypothesis. The ratio of perceived depth/physical depth at the smallest depth of field F2 was less than that at F3.5. The most likely reason for this is that, at the smallest depth of field, the degree of blur introduced reduced the quality of disparity information provided (Costa et al., 2010). As such, a smaller depth of field may have resulted in less perceived depth under stereoscopic viewing conditions. Another possible reason is that the very small depth of field introduced so much blur that the visibility of details of the probes were firmly reduced. The third reason may be that the distance to the scene was more tightly constrained in the case of stereo viewing than in the case of non-stereo viewing (e.g., (Held et al., 2010)) or viewing through an aperture (e.g., (Vishwanath and Blaser, 2010)) which caused the divergence in results between stereo and non-stereo photographs. Changes in depth of field may thus be expected to have less of an effect on perceived egocentric distance under stereoscopic viewing.

On the 3D TV no significant difference in perceived depth was found with increasing depth of field in the range from F3.5 to F22 for photographs with a height-in-the-field cue. However, an additional control experiment showed that perceived depth decreased with increasing depth of field significantly if the height-in-the-field cue was removed from the stimuli. This suggests that the effect of depth of field on perceived depth on the 3D TV may be greatly weakened by other depth cues. But, the results of the control experiment also demonstrate that depth-of-field blur can indeed act as a depth cue even under binocular viewing conditions. When height-in-the-field depth cues are absent, the amount of depth indicated by our observers increased monotonically with increases in depth-of-field blur. As such, these results are consistent with the suggestion of Held et al. (2012) that blur provides a quantitative depth cue under binocular viewing. This study was criticized by Vishwanath (2012) in part because it did not directly demonstrate the effect of blur on the magnitude of perceived depth. For non-stereoscopic photographs, Nefs (2012) provided direct evidence for the effect of depth-of-field blur on perceived depth, consistent with the use of blur as a depth cue. Here, we show that these effects may also occur in stereoscopic viewing.

Physical depth was found to significantly affect perceived depth in stereo photographs in our experiments. We found a clear overestimation of depth between the focus object and the probe compared to the physical depth on both the stereoscope and the 3D TV, as shown in Figures 3.7(c) and 3.8(c). The reasons for this overestimation are different for the two display systems. For the stereoscope, the stereobase of the cameras recording the images was similar to the human adult interocular separation, and the viewing distance and projected size of the images matched those used for the cameras. Binocular depth cues thus indicated the correct depth in this case, so the biases observed represent a genuine overestimation of depth. It is known that perceived depth is overestimated in this way for objects viewed at close distances, both for stereoscopically viewed (Johnston, 1991; Scarfe and Hibbard, 2006, 2011) and natural objects (Loomis et al., 1998; Watt et al., 2005). The overestimation data from some previous studies are partly consistent with our results, indicating that perceived distance was overestimated compared to the physical distance (Elliott, 1987; Loomis et al., 1998). The overestimation of depth in the case of 3D TV may be related to the 4.3 up-scaling factor of the image size, viewing distance and rendered depth (as explained in section 2.3). As the ration of perceived depth/physical depth is close to or slightly smaller than 4.3 in figure 3.8, it seems that our participants perceived depth very close to the rendered depth on the 3D TV. The present study also showed that observers overestimated depth less with increasing physical depth. This latter finding is consistent with results reported by Sauer (2001) and confirms Gilinsky (1951) conclusion that perceived distance is a non-linear function of physical distance. As shown in Figure 3.6, we found that the perceived depth is farther away from the stereoscopically presented depth at larger depth on both the stereoscope and the 3D TV. This confirms the results of a previous study that the effective range of stereopsis as a depth cue attenuates with distance (Held et al., 2012). With increasing physical distance from the observer, the just-discriminable depth threshold also increases (Cutting and Vishton, 1995), implying that observers are less certain about differences in depth. As such, the observers' mapping of perceived depth to physical depth could vary more with increasing distance to the objects (Ernst and Banks, 2002).

Interestingly, there was a significant interaction between depth of field and physical depth on both the Wheatstone stereoscope and the 3D TV; the effect of depth of field on perceived depth was larger at a smaller physical depth. The reason could be that the human eye perceives depth of field cues most effectively within a certain depth range (Kaufman, 1974). Wang et al., however, reported a different result, namely that the effect of blur on perceived depth was smaller when the relative distance between a foreground object and the background was closer (Wang et al., 2011). The crucial difference between their work and ours was that the relative distance in their study ranged from 0cm to 30cm, whereas in our experiment it varied from 20cm to 80cm. This may be explained by the different importance of disparity and blur in different ranges (Held et al., 2012). Alternatively it could be that the differences resulted from the differences between the Gaussian blurring implemented by Wang et al. (2011) and the optical depth of field effects in the current study.

When comparing the results of the two experiments, we found not only many similarities but also an interesting difference. Perceived depth was found to decrease with increasing depth of field from F3.5 to F11 on the stereoscope, while no effect of depth of field from F3.5 to F22 was found on the 3D TV when a height-in-the-field cue was present in the images. In other words, it seemed that the effect of depth of field was smaller for the 3D TV than for the stereoscope. There are obviously a number of reasons why the results on the stereoscope and the 3D TV might have been different. Differences in viewing distance, stereobase of the camera and screen resolution might have directly or indirectly impacted the effect of depth of field on perceived depth. More interestingly, however, is the observation that the impact of depth-of-field blur on perceived depth increased on the 3D TV once the height-in-the-field cue was removed from the stimuli. So, especially on the 3D TV, more than on the stereoscope, the effect of depth of field on perceived depth was weakened by the height-in-the-field cue. The latter may be a direct consequence of the difference in viewing distance between both display systems. At larger (viewing) distance depth-of-field blur is less important and perceived depth may become dominated by other pictorial cues, such as the height-in-the-field cue.

Wang et al. (2011) made the point that depth-of-field blur could be used to supplement the subjective impression of depth in stereoscopically presented images. This means that a smaller disparity-defined depth cue would be needed to create the same perception of depth when the depth-of-field blur cue was used as well. As such, depth-of-field blur could indirectly (i.e., via smaller disparity values) reduce the level of accommodation-vergence conflict, and yet preserve the subjective impression of depth in the scene. However, given the influence of the height-in-the-field cue on depth-of-field effects in certain viewing conditions, it may be that also other combinations of non-stereoscopic or pictorial cues together with depth of field are useful to reduce the required disparity for a given perceived depth, and so, are useful to reduce visual discomfort. For example, it has been suggested that depth of field in combination with shading in a Bayesian framework allows optimal reconstruction of depth relationships in a scene (Li et al., 2013). An essential condition for the resulting reduction in visual discomfort, however, is that the additional cues and depth of field do not introduce discomfort on their own. For example, inappropriately simulated depth of field may be different from that introduced by human optics, and so lead to unstable perceived depth, and possibly discomfort. Vinnikov and Allison (2014) found that the viewing comfort was reduced when depth of field was added as a depth cue. However, it has been shown that depth of field itself does not cause discomfort, even when conflicts with binocular depth cues are large (O'Hare et al., 2013). So, all together our results indicate that depth-of-field blurring is a promising technique for reducing stereoscopically-induced discomfort whilst maintaining subjective impressions of depth. Nonetheless, future work is needed to determine exactly how disparity, depth-of-field blur and other pictorial cues need to be combined to provide the most comfortable depth impression in natural images displayed on stereoscopic screens.

3.6 Conclusion

In summary, the current study confirms the importance of depth of field on depth perception in stereo photographs. Stereo depth of field is different from non-stereo depth of field: in non-stereo images perceived depth monotonically decreases with depth of field, whereas in stereo images perceived depth may increase with depth of field, especially at small depth of field and short viewing distances. The latter may be a consequence of the high degree of perceived blur in these images. We may conclude that if depth of field is not very small (i.e., $>$ F3.5), it may be used as a depth cue in stereo photographs: a decrease in depth of field creates an increase in perceived depth.

Bibliography

Arndt, P., Mallot, H., and Blthoff, H. H. (1995). Human stereovision without localized image features. *Biological Cybernetics*, 72(4):279–293.

Banks, M. S., Read, J. C., Allison, R. S., and Watt, S. J. (2012). Stereoscopy and the human visual system. *SMPTE Motion Imaging Journal*, 121(4):24–43.

Baveye, Y., Urban, F., and Chamaret, C. (2012). *Image and video saliency models improvement by blur identification*, volume 7594 of *Lecture Notes in Computer Science*, chapter 34, pages 280–287. Springer Berlin Heidelberg.

Bradshaw, M. F., Glennerster, A., and Rogers, B. J. (1996). The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research*, 36(9):1255–1264.

Costa, M. F., Moreira, S. M. C. F., Hamer, R. D., and Ventura, D. F. (2010). Effects of age and optical blur on real depth stereoacuity. *Ophthalmic and Physiological Optics*, 30(5):660–666.

Cutting, J. E. and Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. *Academic Press*, 5:69–117.

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

Elliott, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation. *Journal of motor behavior*, 19(4):476–485.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433.

Gilinsky, A. (1951). Perceived size and distance in visual space. *Psychological Review*, 58(6)(6):460.

Gooding, L., Miller, M. E., Moore, J., and Kim, S.-H. (1991). Effect of viewing distance and disparity on perceived depth. volume 1457, pages 259–266.

Hamerly, J. R. and Dvorak, C. A. (1981). Detection and discrimination of blur in edges and lines. *Journal of the Optical Society of America*, 71(4):448–452.

He, Z. J. and Ooi, T. L. (2000). Perceiving binocular depth with reference to a common surface. *Perception*, 29(11):1313–1334.

Held, R., Cooper, E. A., and Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology*, 22(5):426–431.

Held, R., Cooper, E. A., O'Brien, J. F., and Banks, M. S. (2010). Using blur to affect perceived distance and size. *ACM Transactions on Graphics*, 29(2)(2):1–16.

Hillaire, S., Lecuyer, A., Cozot, R., and Casiez, G. (2008). Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE Computer Graphics and Applications*, 28(6):47–55.

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

Hubona, G. S., Wheeler, P. N., Shirah, G. W., and Brandt, M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. *ACM Transactions on Computer-Human Interaction*, 6(3):214–242.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, 31(78):1351–1360.

Kaufman, L. (1974). *Sight and mind: An introduction to visual perception.* Oxford U. Press.

Khan, R. A., Konik, H., and Dinet, E. (2010). Enhanced image saliency model based on blur identification. *Image and Vision Computing New Zealand (IVCNZ)*, pages 1–7.

Lambooij, M., Marten, F., Heynderickx, I., and Ijsselsteijn, W. (2009). Visual discomfort in stereoscopic displays : A review. *Journal of Imaging Science and Technology*, 53.

Li, C., Su, S., Matsushita, Y., Zhou, K., and Lin, S. (2013). Bayesian depth-from-defocus with shading constraints. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 217–224.

Ling, Y., Nefs, H. T., Brinkman, W. P., Qu, C., and Heynderickx, I. (2013). The relationship between individual characteristics and experienced presence. *Computers in Human Behavior*, 29(4):1519–1530.

Loomis, J. M., Klatzky, R. L., Philbeck, J. W., and Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, 60(6):966–980.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1367):169.

Mather, G. and Smith, D. R. R. (2000). Depth cue integration: stereopsis and image blur. *Vision research*, 40(25):3501–3506.

Nefs, H. T. (2012). Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing & Perceiving*, 25(6):577–595.

O'Hare, L., Zhang, T., Nefs, H. T., and Hibbard, P. B. (2013). Visual discomfort and depth-of-field. 4(3):156–169.

O'Kane, L. M. and Hibbard, P. B. (2007). Vertical disparity affects shape and size judgments across surfaces separated in depth. *Perception*, 36:696–702.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531.

Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. *Perception & Psychophysics*, 22(4):400–407.

Rogers, S. (1995). Perceiving pictorial space. *Perception of space and motion*, 5.

Sauer, C. (2001). Perceived depth of 3-d objects in 3-d scenes. *Perception*, 30(6):681–692.

Scarfe, P. and Hibbard, P. (2006). Disparity-defined objects moving in depth do not elicit three-dimensional shape constancy. *Vision Research*, 46(10):1599–1610.

Scarfe, P. and Hibbard, P. (2011). Statistically optimal integration of biased sensory estimates. *Journal of Vision*, 11(7).

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).

Vinnikov, M. and Allison, R. S. (2014). Gaze-contingent depth of field in realistic scenes: the user experience. *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 119–126.

Vishwanath, D. (2012). The utility of defocus blur in binocular depth perception. *i-Perception*, 3(8):541.

Vishwanath, D. and Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision*, 10(10).

Vishwanath, D. and Hibbard, P. B. (2013). Seeing in 3-d with just one eye stereopsis without binocular vision. *Psychological science*, 24(9):1673–1685.

Wang, J., Barkowsky, M., Ricordel, V., and Le Callet, P. (2011). Quantifying how the combination of blur and disparity affects the perceived depth. *Human Vision and Electronic Imaging XVI*, 7865:78650K –78650K–10.

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

Chapter 4

Change Detection

Change detection in pictorial and solid Scenes: the role of depth of field

This chapter has been submitted to a journal:

Change Detection in Pictorial and Solid Scenes: the Role of Depth of Field (2015) *Tingting Zhang, Harold T. Nefs, and Ingrid Heynderickx*

This paper investigated the influence of depth of field on change detection in pictorial and solid scenes. In order to assess the effects on change detection, we conducted a within-subjects experiment using a flicker paradigm, with which the hit rate and response time for change detection could be obtained. Three main factors were investigated in the current study, namely Depth of field (small, large, and no depth of field), Position (close to the focal plane, far away from the focal plane), and Viewing Condition (non-stereo and stereo). The results provided direct empirical evidence that depth of field has effects on change detection: the hit rate is smaller and response time is longer in the scene with small depth of field than in the scene with large depth of field or no depth of field (uniform blur). We concluded that when depth of field is small and binocular disparity is not zero in a picture, depth of field dominates the allocation of attention and hence leads to the result that the change in the sharp area is detected easier and faster than in the area that is closer to the observer.

4.1 Introduction

Change blindness is characterized by the fact that observers fail to notice substantial changes in a scene which are very obvious once they have been identified (Becker and Pashler, 2002; Di Lollo, 1980; Rensink et al., 1997). Change blindness suggests that limited information in the retinal image is stored as an enduring form across a saccade, which is against the traditional view that a complete representation of a real-world scene is constructed by our visual system (Henderson and Hollingworth, 1999b). The frequent reorientation of the eyes by saccadic eye movements leads to the impression that a complete representation can be constructed. Investigating change blindness and change detection can contribute to a better characterization of the nature of the visual representation, visual attention, and visual memory of the real world.

Research on change blindness started from proving the existence of change blindness (Di Lollo, 1980; Henderson and Hollingworth, 1999a), and then followed by discovering the experimental methods to study change blindness (Rensink et al., 1997). Later, researchers investigated the extent that humans are blind to large changes in contexts that simulate real-world perception (Henderson et al., 1999; Hollingworth, 1998; Hollingworth and Henderson, 2000). In these studies, it was found that change blindness can be introduced during a saccade (Grimes, 1996; Henderson and Hollingworth, 2003; McConkie, 1991), during a blink (Kevin O'Regan et al., 2000), or during a blank interval (Simons, 1996). Researchers also found that many factors could facilitate or inhabit change blindness or change detection. These factors included the fixation position (Henderson and Hollingworth, 1999b), scene context (Hollingworth, 1998), and location of the change (Mazza et al., 2005). One of the important findings in these studies was that changes in the foreground (closer to the observer) could be detected easier and faster than changes in the background (Mazza et al., 2005). The position of the change was mainly cued by monocular depth cues such as perspective, which leads to the question whether other depth cues such as depth of field may also facilitate change detection.

Manipulating depth of field is a popular photographic technique that is used widely by photographers and cinematographers to direct viewers' attention to interesting objects or important parts of a scene (Baveye et al., 2012; Colby and Scholl, 1991; Khan et al., 2010; Tsai and Wang, 1998). In these studies, researchers concluded that the sharp area in a scene with limited depth of field could attract viewers' attention. For example, Khan et al. (2010) used a head mounted eye-tracker to show that observers looked at objects in focus, and neglect the blurred background which suggested that sharp areas can attract observers' attention. Urban et al. (2011) investigated the importance of spatial scale to visual attention through an eye-tracking experiment and concluded that intermediate spatial frequencies (0.7-1.3 cycles per degree) attract attention, with some variability in the frequency range between different content categories (e.g., whether the photograph depicts a street or open countryside).

Thus, previous literature indicates the possibility that the fixation positions in a scene with a small depth of field are mainly in the sharp area under free-viewing conditions. The reason could be that the sharp area is more informative compared to the blurred areas. This observation may lead to the hypothesis that changes in sharp areas can be detected easier and faster than in blurred areas and that change blindness is experienced more in blurred areas because sharp areas hold attention (Friedman, 1979; Loftus and Mackworth, 1978). However, there could be conflicting situations. In most pictures, the focal plane is chosen such that with a limited depth of field the foreground is sharp and the background is blurred; however, this is not necessarily always the case. And so, for pictures where the focal plane is not in the foreground, the foreground - closer to the observer - is blurred, while the background - farther away from the observer - is sharp. The question then is which of the two findings related to change detection prevails. Are changes more easily detected in the background since that is sharp (according to the findings of Baveye et al. (2012)) or are changes more easily detected in the foreground since it is closer to the viewer (according to the observations of Jansen et al. (2009), or Mazza et al. (2005))?

In a scene with a large depth of field, the sharp area is large and the effect of depth of field on depth perception is weak. However, other monocular depth cues such as relative size, texture, or perspective may distinguish the foreground from the background, and so, result in changes closer to the observer to be detected faster (Jansen et al., 2009; Mazza et al., 2005). In a scene that is uniformly blurred, the objects are less informative than in a scene with a large depth of field. Hence the hypothesis could be that changes in a uniformly blurred scene are more difficult and slower to be detected than in a scene with a large depth of field.

Apart from the monocular depth cues, binocular depth cues may make the allocation of attention more efficient and then facilitate the detection of change (Chau and Yeh, 1995; Chen et al., 2012). The impact of binocular disparity on visual attention has already been extensively demonstrated (Atchley et al., 1997; He and Nakayama, 1995; Itti and Koch, 2001; Nakayama and Silverman, 1986). For example, Itti and Koch (2001) suggested that stereo disparity could be a relevant attentional cue. Other studies are even more specific and suggest that the visual system may focus attention to a particular depth cued by binocular disparity (Atchley et al., 1997; Bauer et al., 2012). However, these studies still insufficiently answer the question whether the combination of binocular disparity with small depth of field affects change blindness similarly to only binocular disparity or only a limited depth of field.

Therefore, to address the above mentioned questions, we conducted a within-subjects experiment using a flicker paradigm. This paradigm was developed by Rensink et al. (1997) to simulate visual events caused by eye movements. Using this paradigm, we were able to measure the change detection rate and also to record the response time. We created pictures with a small and large depth of field, and compared them to pictures that were uniformly blurred. Since depth of field directs a viewer's attention to sharper areas in the scene (Khan et al., 2010) and change detection requires attention, we may hypothesize that for a small depth of field change detection is facilitated in sharp areas, while suppressed in blurred areas. As such, it is also reasonable to infer that the chance to experience change blindness is higher in pictures with a small depth of field than in the pictures with a large depth of field, since in the latter case a bigger part of the picture asks for visual attention. But apart from having a difference in attention, also the difference in sharpness, and so in detail visibility, between in-focus and out-of-focus areas in the picture may have an impact on change detection. Therefore, it is also interesting to compare the effect of a limited depth of field to the effect of overall blur in the picture in terms of change detection. The position of the change was either in the area close to the focal plane or in the area far away from the focal plane. The focal plane was always in the middle of the scene (in terms of depth), and so, the area far away from the focal plane could be either in the foreground or background, as such also manipulating the distance in depth to the viewer. Finally, binocular disparity may affect the impact of depth of field on change detection, and therefore we perform our investigation with pictorial scenes (used in the remainder of the paper to refer to non-stereo scenes) and solid scenes (used in the remainder of the paper to refer to stereo scenes).

The main contribution of the current work is to reveal the processing priority of various depth regions in complex scenes that contain multiple pictorial cues such as depth of field and binocular disparity. Our study explores how to use an image feature as depth of field to control the selection of information in pictures. In addition, the current work can also help to understand the mechanisms of visual attention and visual representation without using eye-tracking methods to determine so-called saliency maps (Khan et al., 2010; Shepherd et al., 1986).

4.2 Method

4.2.1 Participants

Seventeen male and eighteen female participants took part in this experiment. Their age ranged between 21 and 35 years with a mean of 26.8 years and a standard deviation of 3.3 years. All participants had normal or corrected-to-normal vision, and none of them had color vision deficiencies. All participants gave written informed consent prior to participation. This study was in line with ethical regulations of Delft University of Technology, Dutch Law and the Declaration of Helsinki.

4.2.2 Apparatus

Pictures were presented on a 39" 3D LG television screen. Stereo and non-stereo pictures were viewed in the same experimental set-up. For the stereo pictures, the 3D TV was set to its "Side by Side" mode, and the participants wore passive, circularly polarized glasses.

4.2.3 Stimuli

We took photographs of a scene consisting of a group of 28/29 scattered figurines as shown as an example in Figure 4.1. The photographs were taken with a full-frame Nikon D7000 camera in combination with a Nikon AF - D 50mm lens. We selected an aperture setting of F3.6 to create a small depth of field (i.e., more blur) and of F16 to create a large depth of field (i.e., having the whole photograph sharp). Examples of a photograph with a small and a large depth of field are shown in Figures 4.1a and b, respectively. Figure 4.1c shows an example of a photograph without depth of field effect. In the remainder of this paper, the phrase "uniform blur" is used to represent no depth of field effect. This image was created by convolving the photograph having an aperture of F16 with a Gaussian blur kernel. The width of the Gaussian kernel was determined by the maximum blur in Figure 4.1a (i.e., by calculating the blur circle on the object that was farthest away from the focal plane). As a result the Gaussian blur kernel had a radius (i.e., at 1 standard deviation) of 4.68 pixels. This kernel was subsequently used in all the scenes for creating uniform blur over the whole photograph. When taking the photographs, the "Snow White" figurine positioned at the center of the scene was at a focal distance of 110cm from the center of the camera lens. In the scene, the frontal depth (that is, the depth from the first object to Snow White) was 45cm, and the posterior depth (that is, the depth from Snow White to the farthest object) was 100cm. The setup used to make the photographs is shown in more detail in Figure 4.2. Non-stereo and stereo photographs were created separately. For the latter, the focal plane was again set at the position of the Snow White figurine. Left and right stereo half-images were taken with a stereo base of 2.1 cm, in line with the viewing distance and magnification on the display screen in the experiment. That is, images were shown with the same visual angle to the participants, as the original scene would have if the observer were at

Figure 4.1: An example of the stimuli created with the aperture size of (a) F3.5; (b) F16; (c) uniform blur.

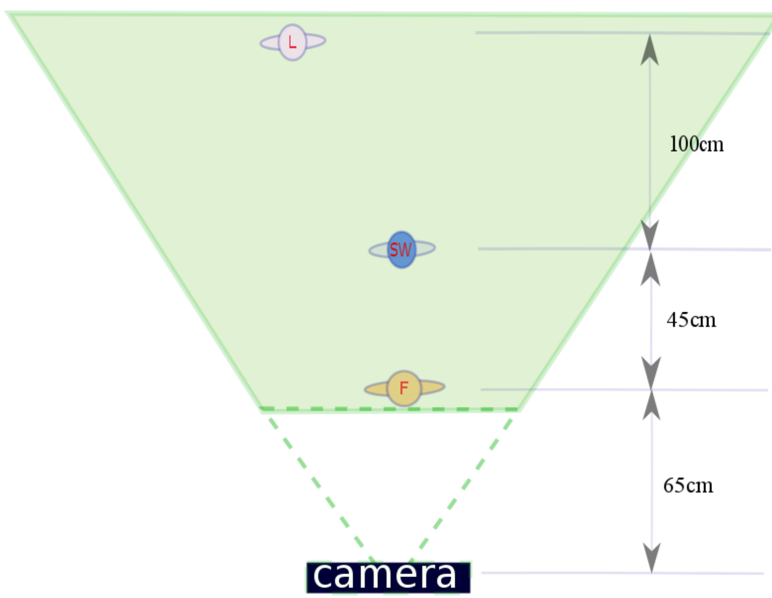

Figure 4.2: A diagram of the setup of the scene. The object named SW shows the position where the camera focused on, the object named F is closest to the camera and the object named L is farthest away from the camera.

the location of the camera. The cameras were set to converge on the object closest to the camera. The screen disparity for the object closest to the camera was hence zero degrees. As a consequence, the closest object was simulated at the distance of the display screen and all other objects in the scene were thus simulated to be behind the display screen. Although the stereo images were toed-in, the actual distortions were small. Earlier tests revealed that images such as these did not lead to visual discomfort (O'Hare et al., 2013).

The actual trials in the experiment consisted of a sequence of photographs, based on the same scene, but either with or without a change in the alternating photographs. In other words, the stimuli were generated from two slightly different photographs, as shown in Figure 4.3. The first photograph contained the original scene and was presented for 320 milliseconds, followed by a grey mask for 160 msec. Then the second photograph was shown for 320 milliseconds, followed again by the grey mask for 160 msec. In the change trials, the second photograph showed a similar scene as the first photograph, but with a new figurine added to it (see Figure 4.3). We call this figurine "change object" in the remainder of this paper. In the no-change trials, the second photograph was exactly the same as the first photograph. The sequence of photographs and gray masks was repeated until observers responded. The whole procedure is shown in Figure 4.4. In the experiment, the change object and its position

Figure 4.3: An example of an original scene and changed scene. The object in the right image in the yellow circle is new and regarded as the change compared to the original scene.

in the scene were not the same for each condition. We only made sure that for each condition six positions close to the focal plane (i.e., with a distance in depth within 10 cm from the plane of the Snow White figurine in the real scene, as shown in Figure 4.5(a)) were selected and six positions far away from the focal plane (i.e., with a distance in depth beyond 10 cm from the plane with the Snow White figurine) were selected. In both cases, the actual position could be in front of or behind Snow White, as such varying the distance in depth between the change object and the observer. This latter variation is illustrated in Figure 4.5(b) for further clarification.

4.2.4 Procedure

The participant was seated at a distance 195cm in front of the television screen in a dimly lit room with an illuminance of 16.6 lux, measured horizontally at the

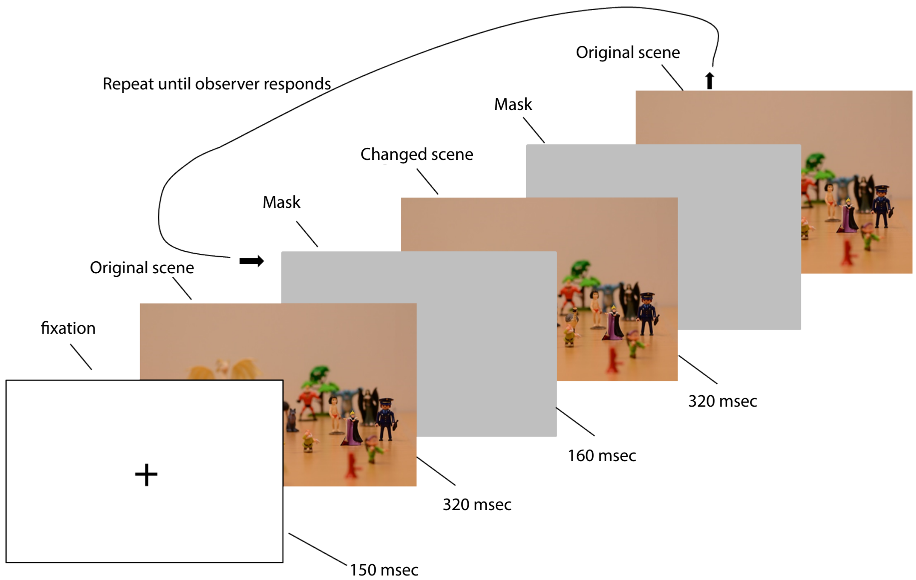

Figure 4.4: A sample trial for this flicker experiment. After an initial fixation, a sequence of photographs is shown, alternating between an original scene and a changed scene (displayed for 320 msec each), separated by a grey mask (displayed for 160 msec).

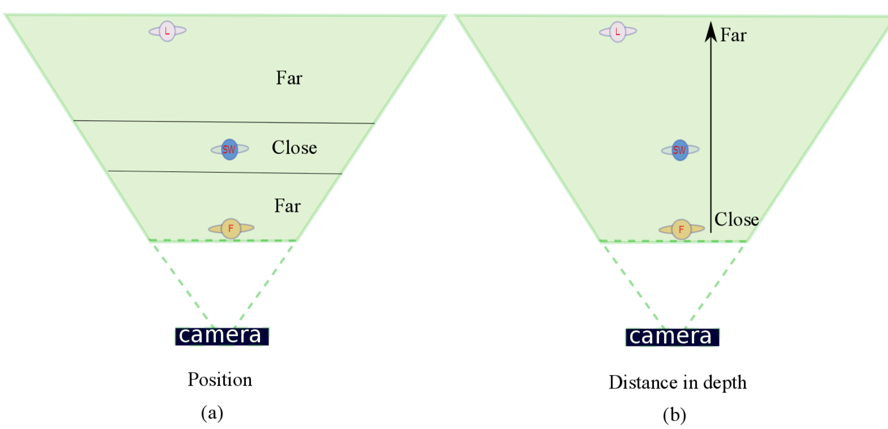

Figure 4.5: Clarification of terminology used in the experiment: the position of the change can be close to the Snow White (focal plane) or far away (a), while the actual distance in depth of the change may be close to or far away from the observer (b).

desk in front of the observer (with VOLTCRAFT MS-1300 Luxmeter having a basic accuracy ±5%). No chin rest was used, hence participants were free to move and to choose a position comfortable to view the screen and operate the keyboard. The experiment was split in two sessions: one for the trials using the stereo photographs and one for the trials using the non-stereo photographs.

The second session started after a break of about fifteen minutes when the first session finished. Seventeen participants first assessed the trial on the non-stereo photographs and eighteen participants started with trials on the stereo photographs.

The experiment consisted of 84 trials in total, of which 42 were based on stereo photographs and 42 on non-stereo photographs. Each set of 42 trials included 36 trials with a change and 6 trials without change. The latter trials were included to provide a measure of false alarms, and also to minimize the chance that participants just pressed the button to speed up their responses.

Before the real experiment started, we asked participants to read the instructions carefully and to fill in the consent form. After that, there was a short training containing four trials so that participants could get familiar with the interface and the task. Two of these four trials contained a change, whereas the other two did not. None of the training trials appeared in the actual experiment.

During the actual experiment, each trial started with requesting the participants to look at a fixation cross (shown in Figure 4.4) until the first photograph appeared. The fixation cross was at the location of the Snow White figurine, and served as a reference to make the initial start location identical across participants. Participants were asked to press the space bar as soon as they identified the change in the trial. It was emphasized that their only concern was to detect the change as soon as possible. They did not need to spend extra time in remembering which object it was nor its location. If a participant could not detect a change and was confident that there was no change, he/she was asked to press the space bar as well. After pressing the space bar, participants needed to answer the question "Did you see any change in the flickering images?". If they pressed the "No" key, they would be directed to the next trial. If they pressed the "Yes" key on the keyboard, they had to first answer the question: "which object appeared in the scene as the change?". Four different figurines, of which only one was correct, were displayed below the question. After this question, they had to answer the question: "In which area did the change happen?". To answer this question, the area of the scene was separated into four subareas, and participants were asked to pick the correct one. Then the instruction "Please press the ENTER button on the keyboard to start next trial" appeared on the screen. Participants could take a break between trials as they wanted, and time to answer any of the questions was not restricted.

4.3 Results

4.3.1 Data Preparation

To prepare our data for further analysis, we first checked the correctness of the responses. According to previous studies (Schoenmakers et al., 2007; Townshend and Duka, 2001), no change can be detected below 200ms from the initial presentation of a change. Therefore, any response within 680ms (i.e., 320ms of the original scene plus 160ms of the blank screen plus 200ms) was considered incorrect. Based on this criterion, we excluded two data points. Additionally, we excluded the data of one of our participants because 67% of his/her data were missing. The reason for this low performance was unclear, but we considered it likely that this participant did not understand the task well enough.

We did not find any false alarm in the experiment, indicating that none of the participants detected a change when there was no change in the trial. Hence, the rate of false alarms is zero and the rate of correct rejections is 1, and so these two types of responses are not further considered in the paper. We were mainly interested in the rate of correct hits. As mentioned before, participants had to answer two questions after they pressed the "Yes" key. If both answers were incorrect, the response was not considered as a correct hit even if they pressed the "Yes" key. This situation only happened once in our data, and this single data point was considered thus as a miss and was not included when calculating the average response times. The response time for those correct hits was measured starting from the first appearance of the original scene to the time participants responded by pressing the Space key.

4.3.2 Hit Rate

We calculated the hit rate (H) as the proportion of correct responses over all responses per participant and per condition. Figures 4.6 (a), (b) and (c) show the hit rate as a function of Depth of field, Position, and Viewing condition, respectively. The first graph shows that the hit rate is smaller in photographs with small depth of field than in photographs with large depth of field or uniformly blurred. Figure 4.6(b) shows the trend that the hit rate is higher in the area close to the focal plane than in the area farther away from it. Figure 4.6(c) only shows a small difference in hit rate between non-stereo and stereo photographs, and in the direction opposite to our expectation: the hit rate is slightly higher in the absence of binocular disparity. Finally, Figure 4.6(d) illustrates the interaction between Depth of field and Position, indicating that the difference in hit rate for objects close by or far away from the focal plane is only found in photographs with a small depth of field, while photographs with a large depth of field or uniformly blurred do not exhibit a difference in hit rate depending on the position of the change object. The hit rates were

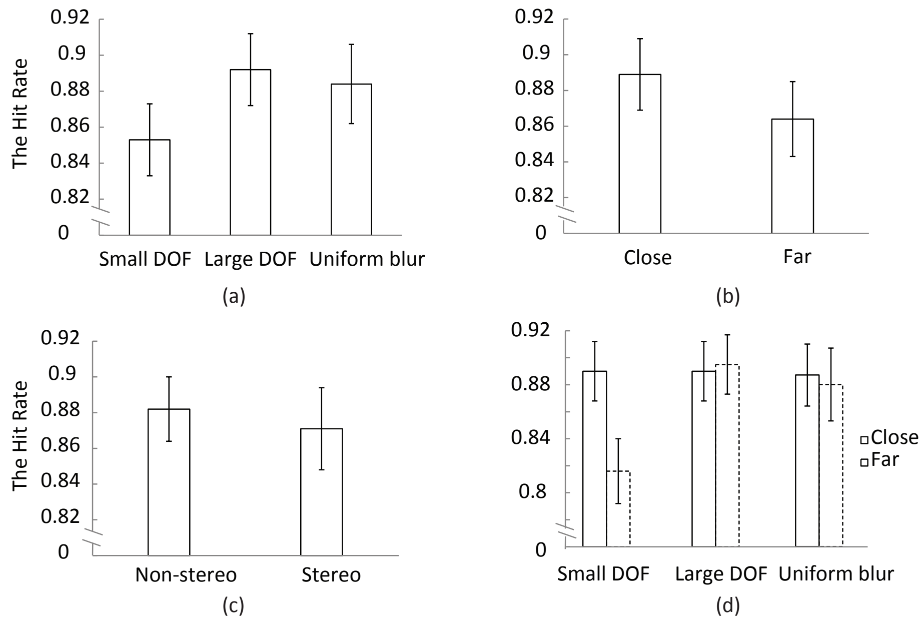

Figure 4.6: Hit rate for different levels of Depth of field (a), across Position (b), and under the two viewing conditions (c); (d) illustrates the interaction between Depth of field and Position. All graphs show the mean ±1 standard error.

entered into a three-way repeated-measures ANOVA, the factors being Viewing condition (two levels: non-stereo and stereo), Depth of field (three levels: small, large and uniform blur), and Position (two levels: close and far). The main effect of Depth of field was significant ($F(2, 66) = 4.04$, $p = .022$), as was the interaction between Depth of field and Position ($F(2, 66) = 4.78$, $p = .012$). The main effect of Position was just not significant ($F(1, 33) = 3.81$, $p = .059$). Viewing condition was not found to have a significant main effect on change detection. Pairwise comparisons between the different levels of Depth of Field showed that the hit rate was significantly smaller for photographs with small depth of field than for photographs with large depth of field or uniformly blurred. There was no significant difference in change detection between the latter two sets of photographs.

4.3.3 Response time

The effect of the independent variables on the response time is summarized in Figure 4.7. Figure 4.7(a) shows that the response time is longer in photographs with small depth of field than in photographs with large depth of field or uniformly blurred. Figure 4.7(b) shows that a change happening in an area close to the focal plane needs less time to be detected compared to a change happening in an area far away from the focal plane. Figure 4.7(c) illustrates that there is a trend that the response time under stereo viewing is shorter than under non-stereo viewing. Figure 4.7(d) shows the interaction between Viewing condition and Position, illustrating that the difference between the two viewing conditions is larger in the area close to the focal plane than in the area farther away from the focal plane.

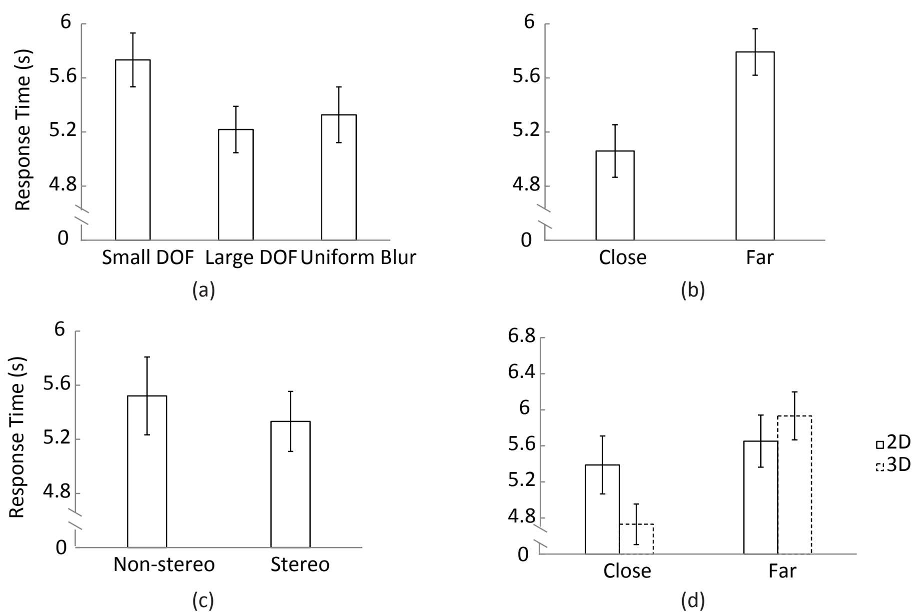

Figure 4.7: Response time as a function of Depth of field (a), as a function of Position (b), and for the two Viewing conditions (c); (d) illustrates the interaction between Viewing condition and Position. All graphs show the mean ±1 standard error.

We performed a repeated-measures ANOVA with three factors: Viewing condition (non-stereo and stereo), Depth of field (small, large and uniformly blurred), and Position (close and far). We found that Depth of field ($F(2, 66) = 5.46$, $p = 0.006$) and Position ($F(1, 33) = 24.1$, $p < .001$) significantly influenced the response time for change detection. Pairwise comparisons between the different levels of Depth of Field showed that the time to detect a change in the photographs with small depth of field was significantly larger than in the photographs with large depth of field or uniformly blurred. There was no significant difference between non-stereo and stereo viewing. The interaction between Viewing Condition and Position was significant ($F(1, 33) = 10.51$, $p = 0.003$).

The results shown above did not exclude any outliers in the data. But we had some extreme outliers, and so we repeated the analysis after excluding data points which deviated more than three standard deviations from their mean response time (calculated per condition). The effects of the main factors were constant with the results including those outliers. The difference was that a significant interaction between Depth of field and Viewing condition ($F(2, 66) = 3.7$, $p = 0.03$), and a significant interaction between all three factors ($F(2, 66) = 4.04$, $p = 0.022$) were found.

In our data, we noticed that the extreme outliers appeared more often in the beginning of the experiment, and also the number of misses was larger for trials in the beginning than in the end of the experiment. Hence, we wanted to check whether there was a learning effect in our experiment. We had two sessions (non-stereo and stereo) including 84 trials and the order of the trials was random across participants. We separated these trials into 14 periods following the time, and each period contained six trials: the first to sixth trials were in the first period, seventh to twelfth trials were in the second period, and so on. The statistical analysis showed that for the participants who first did the non-stereo session followed by the stereo session, the response time decreased with the increasing trials in both the non-stereo session ($F(6, 96) = 5.84$, $p < .001$) and stereo session ($F(6, 96) = 3.13$, $p = .007$), as shown in Figure 4.8(a). In the opposite case when the participants first did the stereo session and then the non-stereo session, there was a decreasing trend for response time with increasing trial for the participants in the stereo session. However, the result was not significant. For the non-stereo session, the response time was not influenced by the order of trials significantly neither, as shown in Figure 4.8(b).

4.3.4 Response time as a function of depth from the first object

We already analyzed the effect of Position, being the distance of the change object to the focal plane (see Figure 4.5(a)) on the response time for change detection. Here we would like to analyze how the actual depth of the change - or the distance of the change with respect to the observer (see Figure 4.5(b)) -

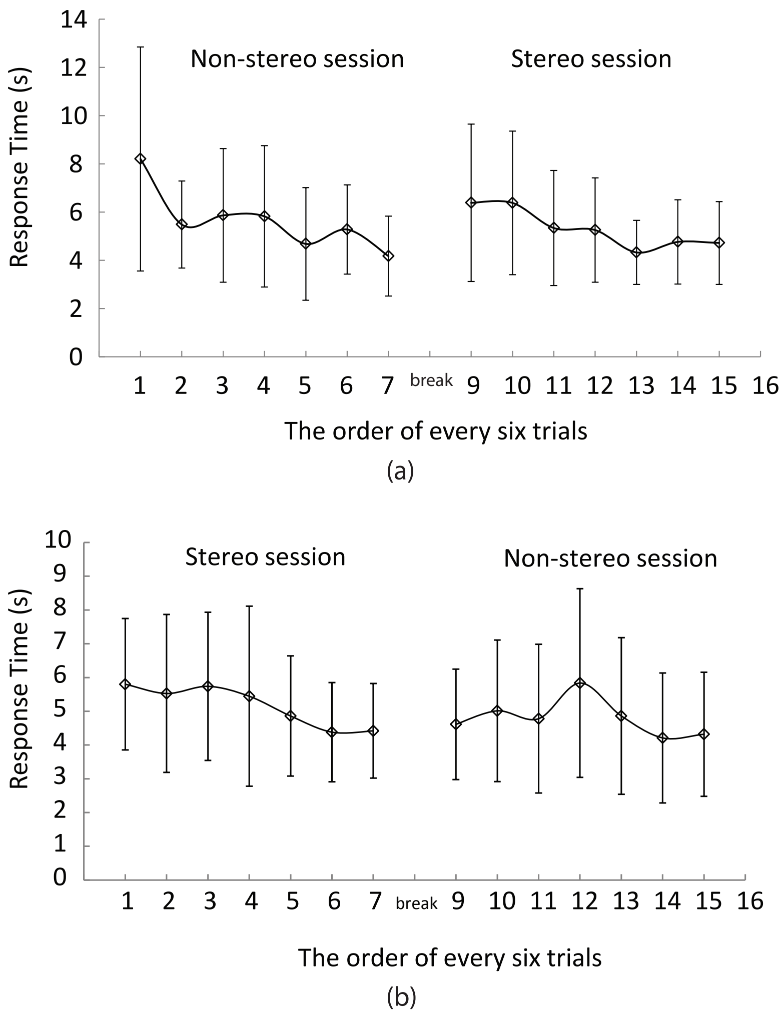

Figure 4.8: The average response time (a) a function of the order of the trials from the participants who first did the non-stereo session and then did the stereo session; (b) as a function of the order of the trials from the participants who first did the stereo session and then did the non-stereo session. The two graphs show the mean 1 standard deviation.

influenced the change detection. Figure 4.9 shows a scatter plot of the response time as a function of depth from the change to the first object (or equivalently, to the observer if we add a constant distance of 65cm to each data point, as can be deduced from Figure 4.2). Each point represents the mean response time across participants for a change at a specific depth distance. Obviously, the response time increases with depth in the stereo viewing condition, but there is no effect of depth in the non-stereo viewing condition. A correlation analysis showed that there is indeed a positive significant relationship between depth and response time under stereo viewing with a Pearson Correlation coefficient of 0.552. So, generally speaking, we found that change detection was significantly influenced by solid depth (in stereo viewing), but not by pictorial depth (in non-stereo viewing).

With the previous analyses, it is still not clear whether depth of field or binocular disparity is more important in influencing change detection. To answer this question, we analyzed the relationship between response time and the distance in depth of the change object across depth of field under stereo viewing conditions. The results are shown in Figure 4.10. The correlation analysis shows that there is no significant correlation when depth of field is small. However, response time and the distance in depth of the change object is significantly correlated when the depth of field is large (Pearson Correlation coefficient is 0.677) and also in uniformly blurred pictures (Pearson Correlation coefficient is 0.583). These results suggest that when a picture with small depth of field is observed under stereo viewing conditions, the allocation of attention is mainly influenced by the effect of depth of field and not by the binocular disparity.

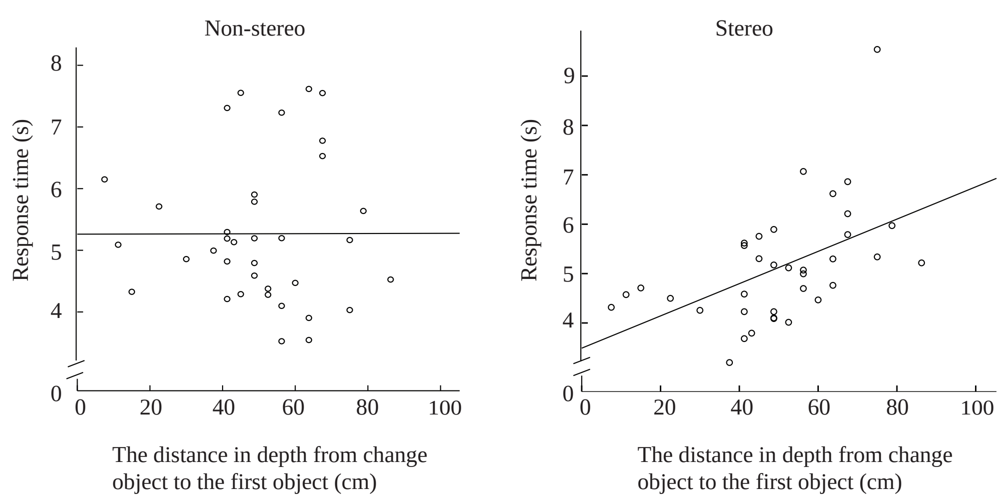

Figure 4.9: Response time as a function of the absolute depth from the change object to the first object (the object closest to the camera) under stereo and non-stereo viewing conditions.

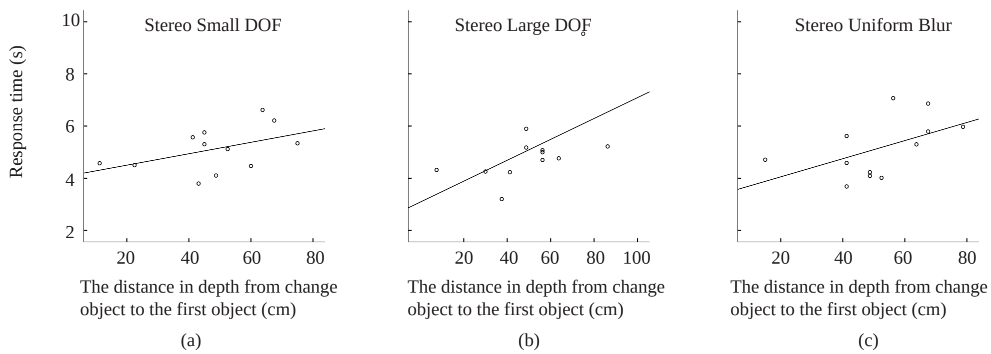

Figure 4.10: Response time as a function of the absolute depth from the change object to the first object (the object closest to the camera) across depth of field under stereo viewing conditions.

4.4 Discussion

The current study provides empirical evidence that depth of field has a significant influence on change detection. The hit rate is lower and the response time is longer for photographs with small depth of field than for photographs with large depth of field or uniformly blurred. The lower hit rate for photographs with small depth of field is mainly due to changes occurring farther away from the focal plane. The hit rate for changes close to the focal plane is very comparable to that found in photographs with large depth of field or uniformly blurred, with the hit rate being independent of the position of the change in both latter cases. In other words, there is a higher chance that observers would ignore the change in the area far away from the focal plane when the small depth of field generates a considerable blur gradient across the photograph. Since Rensink et al. (1997) suggested that change detection occurs only when focused attention is given to the change part, we may conclude that observers' attention is predominantly given and held in the sharp area of the scene with small depth of field, even if the observers have to search the whole photograph to finish a certain task. Therefore, our results are in line with earlier findings obtained with saliency maps based on eye movements, showing that depth of field may direct viewers attention (Colby and Scholl, 1991; Khan et al., 2010). This observation also suggests that our study supports previous studies using saliency maps, and so, that both the flicker paradigm and saliency maps are effective ways to investigate visual attention.

The previous reasoning suggests that the very low hit rate in the area far away from the focal plane in a photograph with small depth of field is only caused by directing visual attention. However, it may also be caused by the limited visibility of details in the blurred foreground or background. If this were the case, we would then expect the hit rate to be also lower in photographs that are uniformly blurred, especially compared to photographs with large depth of field. Our results showed no significant difference in hit rate or response time between photographs with large depth of field and photographs that were uniformly blurred. The results suggest that blur itself is not sufficient to facilitate the experience of change blindness.

The response time is found to be longer for photographs with small depth of field no matter whether the change occurs in the area closer by or farther away from the focal plane. We do not have a clear explanation for this finding. It could be that participants move their fixation from time to time across the scene searching for a change in the photographs with large depth of field or being uniformly blurred. In the photographs with small depth of field, large blur gradients may inhibit such visual processing, leading to longer times to detect a change.

We also find that the position of the change influences the time for change detection significantly and the hit rate at a marginal significance level. Response times to detect changes are significantly shorter in the area closer to the focal plane than in the area farther away from the focal plane. The latter finding does not only hold for photographs with a small depth of field, but also for photographs with a large depth of field or that are uniformly blurred. This may be caused by the initial fixation, positioned always at the center of the focal plane. As such, our results confirm previous studies showing that a change is more easily detected if it appears close to the fixation location (Hollingworth et al., 2001; Posner et al., 1980).

Viewing condition is not a main factor influencing change detection; we only find a slight, but not significant trend that the response time is shorter for stereo viewing than for non-stereo viewing. As such, our results confirm the conclusion drawn earlier by Steinicke et al. (2010) that observers require the same time to detect a change in non-stereo scenes as in stereo scenes when using the flicker paradigm. Although we do not find a significant main effect of viewing condition, we find a significant interaction on response time between viewing condition and position. The response time for change detection is more than one second shorter in stereo photographs than in non-stereo photographs when the change is close to the focal plane. This may be explained by previous studies showing that the visual system is sensitive to binocular disparity in the early visual processing (Han et al., 2005; Hubel and Livingstone, 1987). Furthermore, earlier research also showed that the response time may be speeded when there is a valid cue to the location where a change may happen (Posner et al., 1980; Theeuwes et al., 1998). Hence, cuing the viewer more strongly to the focal plane through a combination of binocular disparity and blur - both known to be valid cues for areas close to the focal plane (Held et al., 2012) - may explain the shorter response time. Contrary to changes close to the focal plane, the response time is slightly longer for stereo viewing than for non-stereo viewing when the change is farther away from the focal plane. This may suggest that binocular disparity does not facilitate change detection in areas far away from the focal plane, and instead seems to inhibit detection of change.

As explained in the introduction, we were interested in whether distance to the observer or blur due to a limited depth of field would dominate the chance of detecting change. To this end, we first plotted the response time for change detection as a function of actual distance in depth to the observer (or to the first figurine in the scene, as given in Figure 4.9) for stereo and non-stereo viewing separately and then plotted the relationship across depth of field under stereo viewing conditions (as shown in Figure 4.10). The depth of the change does not affect response time for non-stereo viewing, but it does for stereo viewing; in the latter case the conclusion that the response time is shorter when the change occurs closer to the viewer holds when depth of field is large or for uniformly blurred pictures. The reason for the shorter response time at closer distance may be related to the larger disparity at increased depth, where the latter requires more effort from eye movements, and more discomfort as a consequence of the conflict between vergence and accommodation (Hoffman et al., 2008); both aspects may lengthen the time to detect a change at larger depth. However, when depth of field is small under stereo viewing conditions, response time does not decrease with decreasing distance anymore. It suggests that a small depth of field dominates the change detection when combined with binocular disparity. This finding indicates that the conclusion drawn by previous researchers that observers detect changes happening closer to the viewer faster than changes happening farther away (Mazza et al., 2005; Jansen et al., 2009) only holds when depth of field is large or when the picture is uniformly blurred.

4.5 Conclusion

In the beginning of the paper, we hypothesized that depth of field may influence change detection because it influences visual attention. The current study directly shows that depth of field indeed has effects on change detection: the hit rate is smaller and response time is longer in scenes with a small depth of field compared to scenes with a large depth of field or uniformly blurred, basically because changes in the blurred background or foreground are not as easily detected in photographs with a small depth of field. Furthermore, we confirm that a small depth of field attracts viewers' attention to the sharp region in a scene, since changes in the area close to the focal plane or close to the fixation location are detected faster than changes in the area farther away from the focal plane. Additionally, although we did not find a main significant effect of viewing condition, we still found that binocular disparity strengthens the effect of depth of field as a depth cue, leading to the conclusion that the response time increases with increasing depth of the change only under stereo viewing conditions.

In the end, we conclude that when depth of field and binocular disparity are combined in complex scenes, the sharp area resulting from a small depth of field has a higher processing priority than the closer area cued by binocular disparity. Depth of field, therefore, is an important image feature that can control the selection of information in pictures under both non-stereo and stereo viewing conditions.

Bibliography

Atchley, P., Kramer, A. F., Andersen, G. J., and Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a depth-aware attentional focus. *Psychonomic Bulletin & Review*, 4(4):524–529.

Bauer, D., Plinge, A., Ehrenstein, W. H., Rinkenauer, G., and Grosjean, M. (2012). Spatial orienting of attention in stereo depth. *Psychological research*, 76(6):730–735.

Baveye, Y., Urban, F., and Chamaret, C. (2012). Image and video saliency models improvement by blur identification. *Computer Vision and Graphics*, pages 280–287.

Becker, M. W. and Pashler, H. (2002). Volatile visual representations: Failing to detect changes in recently processed information. *Psychonomic Bulletin & Review*, 9(4):744–750.

Chau, A. W. and Yeh, Y.-Y. (1995). Segregation by color and stereoscopic depth in three-dimensional visual space. *Perception & Psychophysics*, 57(7):1032–1044.

Chen, Q., Weidner, R., Vossel, S., Weiss, P. H., and Fink, G. R. (2012). Neural mechanisms of attentional reorienting in three-dimensional space. *The Journal of Neuroscience*, 32(39):13352–13362.

Colby, G. and Scholl, L. (1991). Transparency and blur as selective cues for complex visual information. *Proceedings of SPIE 1460 in Electronic Imaging, Image Handling and Reproduction Systems Integration*, pages 114–125.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, 109(1):75–97.

Friedman, A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of experimental psychology: General*, 108(3):316–355.

Grimes, J. (1996). *On the failure to detect changes in scenes across saccades*, volume 5. Oxford University Press.

Han, S., Wan, X., and Humphreys, G. W. (2005). Shifts of spatial attention in perceived 3-d space. *The Quarterly Journal of Experimental Psychology*, 58(4):753–764.

He, Z. J. and Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences*, 92(24):11155–11159.

Held, R. T., Cooper, E. A., and Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology*, 22(5):426–431.

Henderson, J. M. and Hollingworth, A. (1999a). High-level scene perception. *Annual review of psychology*, 50(1):243–271.

Henderson, J. M. and Hollingworth, A. (1999b). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, 10(5):438–443.

Henderson, J. M. and Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, 65(1):58–71.

Henderson, J. M., Weeks Jr, P. A., and Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of experimental psychology: Human perception and performance*, 25(1):210–228.

Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S. (2008). Vergenceaccommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3).

Hollingworth, A. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General*, 127(4):398–415.

Hollingworth, A. and Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, 7(1-3):213–235.

Hollingworth, A., Schrock, G., and Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*, 29(2):296–304.

Hubel, D. H. and Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *The Journal of Neuroscience*, 7(11):3378–3415.

Itti, L. and Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203.

Jansen, L., Onat, S., and Knig, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1).

Kevin O'Regan, J., Deubel, H., Clark, J. J., and Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, 7(1-3):191–211.

Khan, R. A., Konik, H., and Dinet, E. (2010). Enhanced image saliency model based on blur identification. *25th International Conference of Image and Vision Computing*, pages 1–7.

Loftus, G. R. and Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human perception and performance*, 4(4):565–572.

Mazza, V., Turatto, M., and Umilt, C. (2005). Foregroundbackground segmentation and attention: A change blindness study. *Psychological Research*, 69(3):201–210.

McConkie, G. (1991). Perceiving a stable visual world. In *Proceedings of the 6th European Conference on Eye Movements*, pages 5–7.

Nakayama, K. and Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059):264–265.

O'Hare, L., Zhang, T., Nefs, H. T., and Hibbard, P. B. (2013). Visual discomfort and depth-of-field. *i-Perception*, 4(3):156–169.

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, 109(2):160–174.

Rensink, R. A., O'Regan, J. K., and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological science*, 8(5):368–373.

Schoenmakers, T., Wiers, R. W., Jones, B. T., Bruce, G., and Jansen, A. (2007). Attentional retraining decreases attentional bias in heavy drinkers without generalization. *Addiction*, 102(3):399–405.

Shepherd, M., Findlay, J. M., and Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, 38(3):475–491.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, 7(5):301–305.

Steinicke, F., Bruder, G., Hinrichs, K., and Willemsen, P. (2010). Change blindness phenomena for stereoscopic projection systems. *IEEE Virtual Reality Conference (VR)*, pages 187–194.

Theeuwes, J., Atchley, P., and Kramer, A. F. (1998). Attentional control within 3-d space. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5):1476.

Townshend, J. and Duka, T. (2001). Attentional bias associated with alcohol cues: differences between heavy and occasional social drinkers. *Psychopharmacology*, 157(1):67–74.

Tsai, D.-M. and Wang, H.-J. (1998). Segmenting focused objects in complex visual images. *Pattern Recognition Letters*, 19(10):929–940.

Urban, F., Follet, B., Chamaret, C., Le Meur, O., and Baccino, T. (2011). Medium spatial frequencies, a strong predictor of salience. *Cognitive Computation*, 3(1):37–47.

Chapter 5

Aesthetics and Overall Quality

The role of depth of field in the subjective evaluation of photographs' aesthetic appeal and overall quality

This chapter has been submitted to a journal. Part of this chapter has been published in:

The aesthetic appeal of depth of field in photographs (2014) *Tingting Zhang, Harold T. Nefs, Judith Redi, and Ingrid Heynderickx,* Sixth International Workshop on Quality of Multimedia Experience (QoMEX), page *81-86.*

Image blur is commonly assumed to degrade image quality, but this may not be necessarily the case. Defocus blur introduced by depth of field in particular is considered to be able to enhance the aesthetic appeal of image material, though empirical evidence for that is lacking. Therefore, we here investigate the effect of depth of field on aesthetic appeal and overall quality of photographs. Three hundred and thirty nine photographs covering eight different content categories were selected from internet or personal collections. In a preliminary experiment, sixteen participants classified the size of the depth of field of these photographs into three levels, namely a small, medium, or large depth of field. In the main experiment, thirty-two new participants scored each photograph separately in terms of aesthetic appeal and overall quality. We found that content category affected aesthetic appeal and overall quality significantly and the effect of depth of field on aesthetic appeal and overall quality was different across content categories. The impact of content category may be caused by differences in photographic characteristics across these categories, such as semantics or clutter. We conclude that depth of field may indeed enhance the aesthetical appeal of photographs, but this conclusion is highly dependent on the content of the photograph.

5.1 Introduction

Blur is a familiar photographic property, but its perception and appreciation are not yet well understood. In general, blur is characterized by a loss of power in the higher spatial frequencies of image content. It may originate at several levels: it can be present in photographs as a consequence of camera movement, camera defocus or a large aperture of the camera, but it can also be created in the human visual system by aberrations in the eye lens (Ogle and Schwartz, 1959) or by an incorrect focal distance. Nowadays, image blur can also be caused by digitization or format conversion. Blur is often regarded as image degradation, and so, as decreasing visual quality, which has already been extensively investigated (Crete et al., 2007; de Ridder, 2001).

However, contrary to the common first impression, blur does not always degrade image material. It is sometimes intentionally introduced photographically for aesthetic or artistic reasons, most frequently in the form of depth of field. Depth of field is defined as the depth range around the focal plane that is perceived as sharp. When taking photographs, the reproduction of the focal object is always sharp, whereas the reproduction of other objects that are not in the focal plane becomes blurred. The amount of blur depends on the distance of the object from the focal plane. As such, photographs with depth of field blur usually have both sharp areas and blurred areas (depending on the placement of objects in the scene relative to the focal plane).

Depth of field was first used in a movie by cinematographer Gregg Toland and director Orson Welles in 1941. They used a large depth of field to make details sharp in both foreground and background, probably to improve overall visual quality through overall sharpness. Compared to the use of a large depth of field in microscopy and photography, the use and understanding of a small depth of field in photography is still not well documented. Researchers and photographers alike claimed that a small depth of field might have an important contribution to increased aesthetic appeal of photographs (Datta et al., 2006; Dhar et al., 2011). Although researchers reported statements relating depth of field blur to visual quality, these statements are only marginally substantiated by empirical data. Therefore, the main goal of this paper is to quantify to what extent depth of field affects aesthetic appeal and overall quality in photographs.

Literature reports several reasons for the impact of depth of field on aesthetic appeal of photographs. First, photographers always use a small depth of field to bring the main topic of their photograph in focus, while blurring other objects that are less important in the foreground or the background (Cerosaletti and Loui, 2009). As such, depth-of-field blur contributes to reduce clutter in a photograph, a property known to be inversely proportional to aesthetic appeal. Second, Luo and Tang (2008) argued that depth-of-field blur influenced clarity contrast and simplicity in a photograph, and thus, might impact aesthetic appeal. Third, depth of field is known to enhance the impression of depth in photographs (Nefs, 2012; Pentland, 1987; Vishwanath and Blaser, 2010; Zhang et al., 2014b), and so, may make objects appear more vividly in depth, more realistic, and more stereoscopic (Hillaire et al., 2007), which are all properties that may contribute to aesthetic appeal. But depth of field can also have a tilt-shift miniaturization effect in photographs (Held et al., 2010; Vishwanath, 2007). Such a miniaturization effect changes the familiar size of objects in the scene, making them appear toy-like. As such, depth of field may have a negative effect on the aesthetic appeal of photographs. Hence, based on the above reasons, no general conclusion on the relationship between aesthetic appeal and depth of field can be inferred, and so, the first goal of this paper is to evaluate this relationship directly.

The second goal of our research is to quantify to what extent depth of field blur affects the overall quality of photographs. Assuming a positive effect of depth of field blur on aesthetic appeal, and relying on the assumption that aesthetics contribute to overall quality (Fedorovskaya, 2002; Loui et al., 2008), one would expect a positive effect of depth of field blur on overall quality as well. In addition, Vishwanath (2014)reported that depth of field could introduce the impression of stereopsis, while IJsselsteijn et al. (2000) found that stereoscopic depth improved image quality because of enhanced naturalness. Hence, it seems reasonable to hypothesize that depth of field may improve overall quality. On the other hand, the negative impact of 'distortion' blur on overall quality (e.g. introduced by compression, re-scaling or transcoding) is well documented (Crete et al., 2007; de Ridder, 1998; Roufs, 1989). A large body of research has been dedicated to blur detection and annoyance quantification (Ferzli and Karam, 2009; Marziliano et al., 2002, 2004) as well as to sharpening algorithms (Biemond et al., 1990; Yuan et al., 2007). In addition, Vishwanath's research is inconclusive on how variations in depth of field blur affect the impression of stereopsis.

Moreover, the findings of IJsselsteijn et al. (2000) on the relation between stereoscopy and image quality were not always confirmed by later studies (Kaptein et al., 2008; Kuijsters et al., 2009; Seuntiens et al., 2006). Therefore, the effect of depth of field on overall quality is far from clear from the existing literature. Only one contribution in literature makes an explicit distinction between depth of field blur and distortion blur. Liu et al. (2011) found that blur quality metrics and sharpening algorithms should be able to discriminate between depth of field blur and unwanted 'distortion' blur, towards preserving the former and, as a consequence, the overall quality of an image. But as research on the distinction between depth of field blur and distortion blur in terms of image quality is limited, we considered the investigation on the effects of depth of field on overall quality necessary.

In practice, depth of field is not equally used in photographs of all content categories. Small depth of field is preferred in macro photography, resulting in a sharp area of interest against a very blurred background. Large depth of field, on the other hand, is commonly used in landscape photography (Joshi et al., 2011). These observations suggest that people might expect, and hence, appreciate a different depth of field for different content categories. Content category in itself is known to influence aesthetic appeal (Joshi et al., 2011; Ke et al., 2006; Obrador et al., 2012). For example, Ke et al. (2006) designed a model to discriminate high from low quality photographs, and suggested that the importance of the perceptual features such as color and blur used for discrimination may change over content categories. Additionally, Obrador et al. (2012) showed that category-based aesthetic models of photographs performed better than a generic aesthetic model. The only evidence that the effect of depth of field on aesthetic appeal is also content dependent is given by Joshi et al. (2011); they suggested that a small depth of field has a positive effect on aesthetic appeal, but only in an appropriate semantic context. Because of the limited knowledge about how content category affects the relation between depth of field and aesthetic appeal, we included different content categories in the set of evaluated photographs.

To some extent, the impact of depth of field on aesthetics and overall quality also depends on a person's background. Previous research suggests that experts are preferred for assessing image aesthetics because aesthetic perception may be a learned process (Axelsson, 2007; Szechter and Liben, 2007). In our study, photographic experts may understand the meaning of depth of field blur to the aesthetics of the photograph, and as such may accept the resulting blur when assessing overall quality. Naive image consumers may instead be less impressed by the effect of depth of field, and be more annoyed by the presence of blur in (parts of) the photograph. Therefore, we added another research question to our study, in which we evaluated whether the impact of depth of field on aesthetics and overall quality depended on the participants' level of expertise in photography.

To achieve our goals, we conducted a subjective study to obtain observers' evaluations of aesthetic appeal and overall quality for photographs selected from the Internet or personal collections, and covering eight content categories. As the amount of depth of field in these photographs was not clearly defined, we started with a preliminary experiment, in which participants subjectively quantified the size of depth of field in our photographs. In the main experiment, we asked new participants to rate these photographs on aesthetic appeal and overall quality. The latter group of participants was divided in experienced or naive, so that we could evaluate whether their level of expertise had an effect on the impact of depth of field on aesthetic appeal and overall quality.

5.2 Preliminary experiment: Subjective perception of depth of field

To measure the effect of depth of field on aesthetic appeal and overall quality, the first essential step was to quantify the size of the depth of field for the photographs included in our testbed. When selecting photographs from online repositories, calculating the depth of field is not always possible, due to the lack of information on essential parameters, such as the distance between the focus object and the camera. In addition, for photographs taken from a real-world scene, the physical scale of the scene varies greatly over different content categories. For example, the scale of the scene for photographs of landscapes may vary from meters to hundreds of meters, whereas for photographs of animals or flowers it may vary within centimeters or meters. The scale of the scene may influence the way observers perceive depth of field in the photographs, suggesting that content category may have an effect on the perception of depth of field. As a result, we opted to measure the perceived size of depth of field for the photographs that we planned to use in our study subjectively.

5.2.1 Method

Participants. Sixteen participants took part in this preliminary experiment and seven of them were male. Most of these participants were from Delft University of Technology. Their ages ranged from 24 to 40 years with a mean age of 29 years (SD=3.5). The experimental protocol was approved by the ethics committee of Delft University of Technology.

Stimuli. The experiment used 339 photographs, which were selected from the Internet or created by the authors. Most of the photographs (N=315) were selected from Internet sources, such as photo.net [1], Flicker [2] , Google+ [3], or personal collections. These photographs were classified in seven content categories: 50 photographs of animals, 25 of architecture, 50 of flowers, 48 of food, 53 of landscapes, 39 of streets, 50 of sport. When available, the content labels were taken over from Flickr and photo.net. When not available, as for the photographs from personal collections, the content labels were assigned by the authors. When selecting the photographs, we noticed that the amount of blur in photographs of streets and architecture did not vary that much. Hence, we decided to include fewer photographs of these two content categories to reduce experimental time. The remaining 24 photographs not selected from the Internet were created by the authors (Nefs, 2012; Zhang et al., 2014b).

[1] www.photo.net
[2] www.flickr.com
[3] https://plus.google.com

They portrayed artificial scenes, in which depth of field was fully controlled. The details about how the photographs were created can be found in previous work (Nefs, 2012; Zhang et al., 2014b). In the current work, camera aperture settings of F2, F6.3 and F22 were chosen to create three levels of depth of field. These three levels of depth of field could be discriminated by humans (Zhang et al., 2014a). The size of the depth of field for F2, F6.3, F22 was calculated to be 5.7mm, 18.4mm, and 68.9mm respectively in the original physical scene. In the remainder of this paper, we refer to these photographs as the artificial photographs.

Procedure. The experiment was performed in a standard office room, in which photographs were shown to the participants on a PLE48 19 inches Iiyama monitor screen with a resolution of 1280 pixels wide by 1024 pixels high. The screen was calibrated with a ColorMunki spectrophotometer towards sRGB. The room was illuminated only by the light from the monitor screen. The test leader explained the definition of depth of field to the participants, using the words: "the distance within which objects are perceived as sharp". Subsequently, the 339 photographs were displayed in the middle of the screen in a random order for each participant. The position of the participants was fixed, though we did not use a chinrest to fix their head. For each photograph, the participant was asked to indicate whether the size of the depth of field was small, medium, or large by clicking the corresponding button. The participants could look at each photograph as long as they wanted, and had to click the "Next" button to proceed to the next photograph. As the task was relatively easy, it took most participants about 30 minutes to complete the experiment.

5.2.2 Results

We calculated the mode of the scores for depth of field in terms of the three labels (i.e., small, medium and large) over all participants. For those photographs where the mode corresponded to the score given by more than eight participants, we took the mode as the size of the depth of field for that photograph. When the mode of the scores was obtained with fewer participants, which meant that the number of participants choosing a given value of depth of field was smaller than nine for each of the three labels (i.e. small, medium or large), we concluded that the size of the depth of field for that particular photograph was ambiguous.

Out of 339 photographs, we found that 67 were classified as having a small depth of field, 96 as having a medium depth of field, and 156 as having a large depth of field. The size of depth of field was ambiguous for 20 photographs, in the sense that there was no clear agreement among the participants on whether the depth of field was small, medium or large. Figure 5.1 shows the distribution of depth of field of the photographs over the different content categories. The size of depth of field was perceived as large for most photographs of architecture and landscape scenes, whereas none of the artificial photographs was considered as having a large depth of field. Five content categories (i.e., animal, flower, food, people, and sport) contained photographs with the depth of field spread over the three labels, implying that for these contents, photographs with small and large depth of field were available in our selection. As such, these content categories were most valuable to investigate how depth of field influences the perception of aesthetic appeal and overall quality.

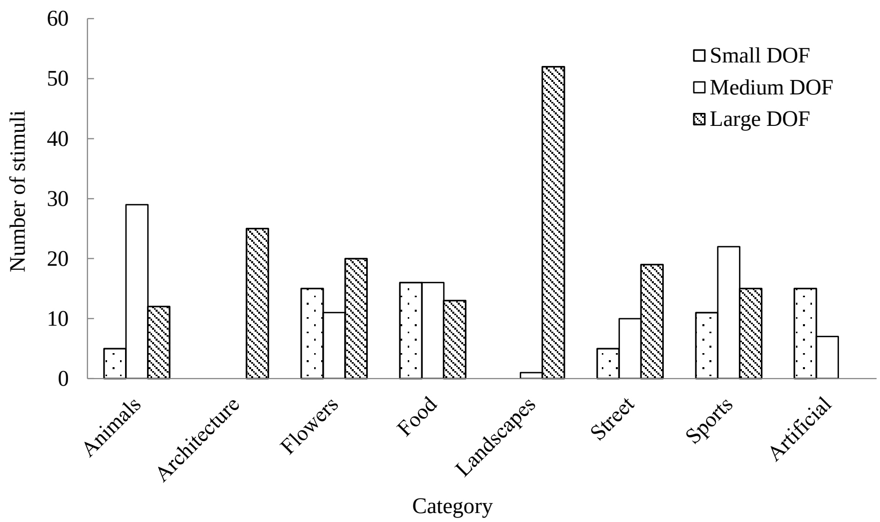

Figure 5.1: The distribution of the size of depth of field for photographs in eight content categories.

5.3 Main experiment: Subjective assessment of aesthetic appeal and overall quality

5.3.1 Method

Participants. The main experiment was performed by 32 participants, of which 16 were male. These participants were spread across eleven different nationalities (i.e., 18 Chinese, 4 Dutch, 2 Romanian, and 1 Brazilian, 1 British, 1 Indonesian, 1 Portuguese, 1 Mexican, 1 Russian, 1 Indian, and 1 Italian). The age of the participants varied between 17 and 39 years with a mean age of 27 years (SD=3.7). None of the participants participated in the preliminary experiment. All participants were naive with respect to our hypotheses. This experimental protocol was approved by the University Ethics Committees of Delft University of Technology.

Stimuli. The same 339 photographs as used in the preliminary experiment (see section 5.2.1) were used again in this experiment. Although 20 of them were classified as having an ambiguous depth of field in the preliminary experiment, they were still scored on aesthetic appeal and overall quality because it could provide more data for further research, but they were excluded from further analyses in this paper.

Procedure. Because of the large amount of photographs that had to be assessed, the experiment was separated into six sessions. Three sessions were devoted to assessing aesthetic appeal and three to assessing overall quality. Sessions on aesthetic appeal and on overall quality were alternated for each participant. At the beginning of each session, participants received instructions in order to remind them of their task. Each session included 113 photographs, shown in a random order to each participant. Participants were asked to rate aesthetic appeal or overall quality (depending on the session) on a continuous rating scale. To this end, they used a slider bar that was displayed just below the photograph on the same monitor screen. Both rating scales contained marks ranging from "1" to "7". When assessing aesthetic appeal, "1" was labeled as "ugly", and "7" was labeled as "beautiful". When assessing overall quality, "1" was labeled as "bad", and "7" was labeled as "excellent". Intermediate marks on the rating scale were not labeled. Participants could take as much time as they wanted to express their score, and had to click the "Next" button to proceed to the next trial.

Before the first session started, participants were given the definition of the quantity they had to score. Aesthetic appeal was defined as "concerned with beauty and art, the understanding of beautiful things, and made in an artistic way and beautiful to look at" (Datta et al., 2006). To ensure that participants understood what we meant, we showed some example reasons for a rating close to 7 on aesthetic appeal. These reasons included "looks good, pleasing to your eyes, attracts attention, interesting composition, great use of color and contrast". We copied these descriptions from Photo.net, where they were used as the standard to rate aesthetic appeal. Overall quality was described as: "if no degradation or artifact is perceived in an image, it is considered to be of high quality". In this case, we provided some reasons for a low rating in overall quality, including "main subject out of focus, obvious visible distortions". At the end of the experiment, observers completed a short questionnaire, the questions of which are shown in Table 5.1. With this questionnaire we wanted to collect their personal background, interest, and knowledge related to our study.

Table 5.1: Questionnaire in this study.

Questions
1. What is your standard for good aesthetic quality?
2. What is your standard for good overall quality?
3. Do you like taking photographs? (1 Strong disagree, 5 Strong agree)
4. Do you like looking at photographs? (1 Strong disagree, 5 Strong agree)
5. Do you have a lot of experience in taking photographs?
(1 Strong disagree, 5 Strong agree)
6. Do you rate photographs online? (1 Strong disagree, 5 Strong agree)
7. Have you ever done courses about photography?
(1 Strong disagree, 5 Strong agree)
8. Do you know what Depth of Field is? (1 Strong disagree, 5 Strong agree)
9. If you know Depth of Field, would you please describe it in your own words?
10. Which category of the photographs do you like best? (multiple choices)
11. Which category of the photographs do you like worst? (multiple choices)
12. Would you expect that Depth of Field affects aesthetic appeal of photographs?
13. Would you expect that Depth of Field affects overall quality of photographs?

5.3.2 Results

Overview of the subjective aesthetic appeal and overall quality

Before analyzing the effects of depth of field on aesthetic appeal and overall quality, we first examined the between-subjects' agreement on the ratings, using the so-called "SOS Hypothesis". SOS refers to the standard deviation over the opinion scores. Hofeld et al. proposed that the relationship between the squared standard deviations over the opinion scores (SOS^2) and the mean opinion scores (MOS) is quadratic (Hobfeld et al., 2011). Indeed, the SOS is expected to be smaller near the end of the scoring scale, because the end of the scoring scale truncates the possibility of variations in single opinion scores. As such, the SOS hypothesis can be described as:

$$SOS(X)^2 = a(-x^2 + (\nu_1 + \nu_k)x - \nu_1\nu_k) \quad (5.1)$$

with x representing MOS, ν_1 and ν_k the beginning and ending points of the scale, and a the SOS parameter. If a is small, the SOS is relatively small. If a from one study is in the same range as a from other studies, then those studies are comparable in terms of scoring consistency among subjects. We applied Equation 5.1 to our data set using ν_1 =1 and ν_k =7 according to our scoring scale. The result is visualized in Figure 5.2, and resulted in a value of a = 0.2 for aesthetic appeal and a value of a = 0.24 for overall quality. In Figure 5.2, each point represents the standard deviation of each mean opinion score and the solid line represents the fitted square functions according to Equation 5.1, whereby the SOS parameter is obtained by minimizing the least squared errors between the measurement data and the fitting function. The SOS parameters for both aesthetic appeal and overall quality are comparable to that found for overall quality assessment as well as for VOIP (Voice over IP) QoE (Quality of Experience) assessment (Redi et al., 2010), and thus, the between-subjects' agreement in scoring overall quality and aesthetic appeal is what we could expect from other research. In addition, the SOS parameter of aesthetic appeal is slightly smaller than that of overall quality, suggesting that the subjective agreement on aesthetic appeal is equal or slightly higher than for overall quality.

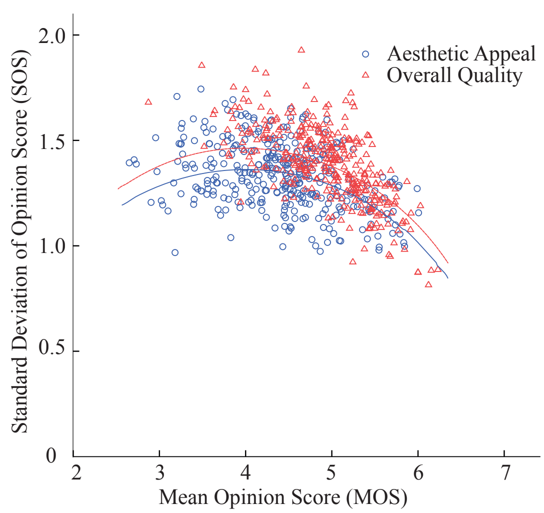

Figure 5.2: Standard deviation as a function of mean rating of aesthetic appeal and overall quality separately for each of the 339 photographs.

Subjective evaluation of aesthetic appeal

To calculate the Mean Aesthetic Appeal Scores (hereafter referred to as MAAS), we first averaged the aesthetic appeal scores given per photograph over the 32 participants. Subsequently, we averaged the MAAS scores over all photographs of a given content category in order to calculate the MAAS per content. These

MAAS (1 standard error) per content category are shown in Figure 5.3. In general, photographs of streets and flowers were rated as less beautiful than photographs of other categories, whereas architecture, landscape and artificial photographs were rated high on aesthetic appeal. To evaluate whether these differences in aesthetic appeal between the eight different content categories were significant, we performed a repeated-measures ANOVA, and found that indeed aesthetic appeal was significantly affected by Content category ($F(7,217)=14.47$, $p< .001$). An LSD post-hoc test, the results of which are summarized in Table 5.2, revealed that not all content categories were necessarily significantly different from each other in aesthetic appeal.

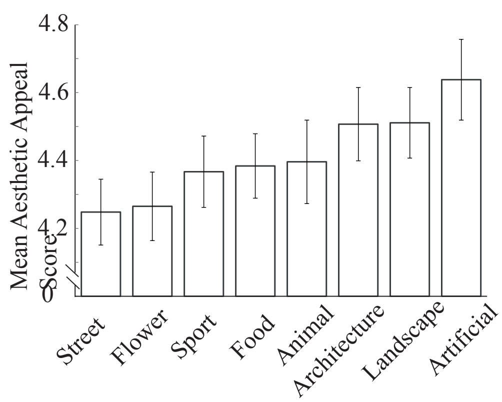

Figure 5.3: Mean Aesthetic Appeal score with ±1 standard error across content categories.

Table 5.2: Statistical significance of the difference in aesthetic appeal among content categories; only p-values smaller than .05 are shown in the table.

	Animal	Architecture	Flower	Food	Landscape	Street	Sport	Artificial
Animal	-	.03	.012	-	.011	.033	-	.001
Architecture	.03	-	.001	.012	-	.001	.003	.007
Flower	.012	.001	-	.012	.001	-	.009	.001
Food	-	.012	.012	-	.004	.01	-	.001
Landscape	.011	-	.001	.004	-	.001	.003	.004
Street	.033	.001	-	.01	.001	-	.002	.001
Sport	-	.003	.009	-	.003	.002	-	.001
Artificial	.001	.007	.001	.001	.004	.001	.001	-

The results of the preliminary experiment, summarized in Figure 5.1 (see Section 5.2), revealed that the photographs of architecture and landscape all had a large depth of field, whereas the artificial photographs mainly had a small to medium depth of field. Hence, the photographs in these content categories could mask an overall effect of depth of field on aesthetic appeal. As such, we excluded these three content categories from further analysis. As a result, a 3 (Depth of field) $\times$ 5 (Content category) repeated-measures ANOVA was performed on the aesthetic appeal ratings. Depth of field was not found to have a significant main effect on aesthetic appeal ($F(2, 62)=0.376$, $p = 0.69$); there was only a slightly increasing trend of aesthetic appeal with depth of field, as shown in Figure 5.4. The analysis, however, revealed a significant interaction between Depth of field and Content category on aesthetic appeal ($F(8,248)=10.18$, $p< .001$), implying that the effect of depth of field on aesthetic appeal was not the same across the five content categories, considered in this analysis. Figure 5.5 illustrates the effects of depth of field on aesthetic appeal among different content categories. Depth of field influenced aesthetic appeal significantly for photographs of animals ($F(2, 62) = 6.76$, $p = .002$), and sports ($F(2,62) = 45.02$, $p< .001$). Photographs of animals with a medium depth of field were found to be less beautiful than photographs with a large or small depth of field. For photographs of sport, we found an opposite effect: photographs with a medium depth of field were scored as more beautiful than photographs with a large depth of field or a small depth of field. No significant main effect of depth of field on aesthetic appeal was found for the photographs of flowers, food, or street.

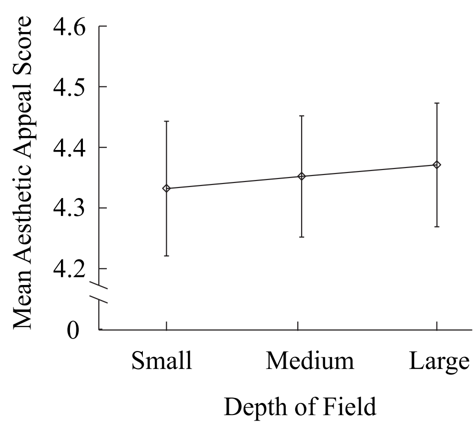

Figure 5.4: Main effect of depth of field on aesthetic appeal (mean ±1 standard error).

Since it was not clear whether the effect of depth of field on aesthetic appeal was systematic within a content category or caused by the specific selection of photographs used in this experiment, we used bootstrapping to resample

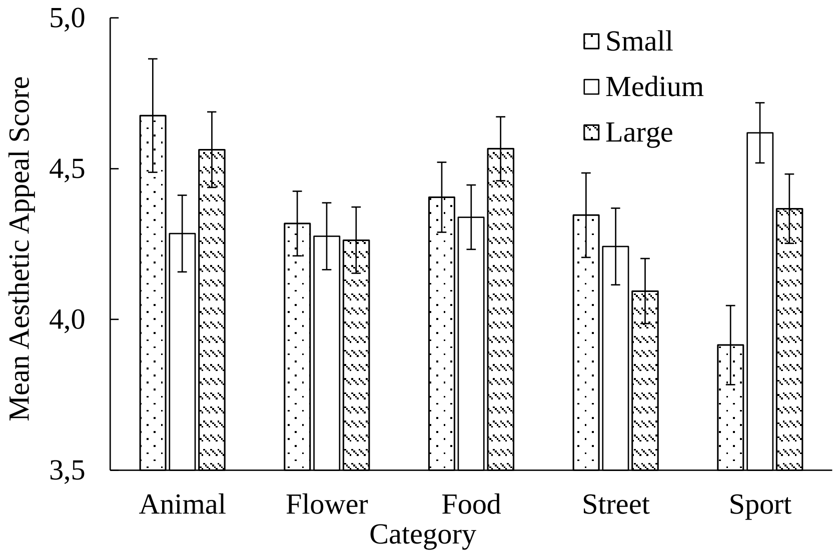

Figure 5.5: Influence of depth of field on aesthetic appeal for the different content categories (mean 1 standard error).

the photographs of animals and sport. In these two categories, the minimum number of the photographs was five for all the levels of depth of field. Hence, we picked four photographs from each level of depth of field for those two content categories eight times, resulting in new selections of photographs labeled as S1, S2, S3, S4, S5, S6, S7, and S8 respectively. Figure 5.6(a) and (b) show the effect of depth of field on aesthetic appeal for the eight sets of photographs. We can see in Figure 5.6(a) that photographs with a medium depth of field were rated as less beautiful than photographs with a large depth of field or a small depth of field in seven out of the eight samples. Figure 5.6(b) also shows that the effect of depth of field on aesthetic appeal in photographs of sport is consistent again in seven samples out of eight. From these results, we can conclude that a medium depth of field makes photographs of animals less beautiful, whereas oppositely, a medium depth of field makes photographs of sports more beautiful.

Subjective evaluation of overall quality

For overall quality, we first averaged the scores over the 32 participants in order to obtain a mean overall quality score (MOQS) per photograph. We plot

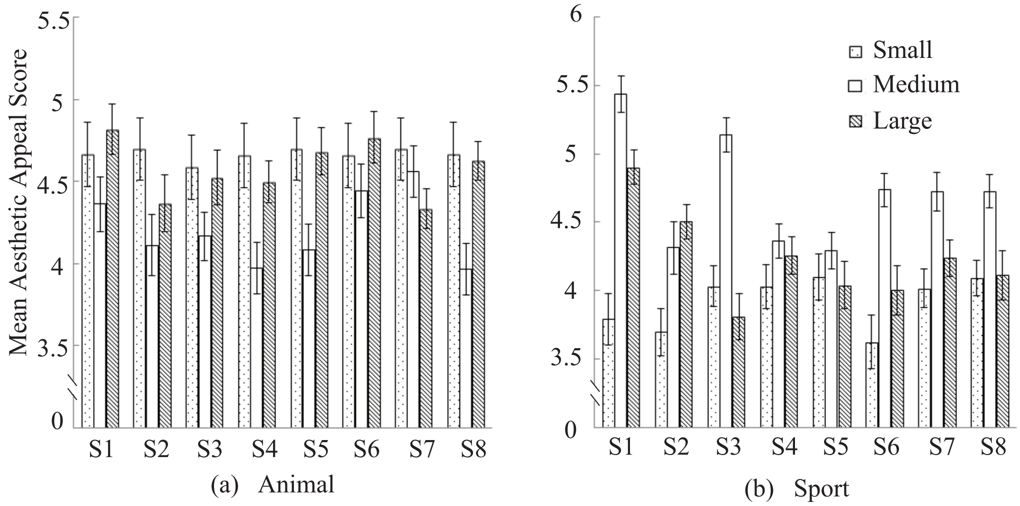

Figure 5.6: Effect of depth of field on mean aesthetic appeal for different samples of (a) photographs of animals and (b) photographs of sport, with ±1 standard error. In the figures, the label "S1 to S8" under the horizontal ordinate represent a randomly selected sample from the stimuli.

in Figure 5.7 the MOQS with ±1 standard error averaged over all photographs per content category. A repeated-measures ANOVA showed a significant impact of Content category on the subjective evaluation of overall quality ($F(7, 217) = 7.09$, $p < .001$). Figure 5.7 shows that the overall quality of photographs of architecture, landscape, or artificial is higher than that of photographs of street, flower, or sport. The results of the LSD post-hoc test (shown in Table 5.3) reveals that not all content categories were significantly different from each other in overall quality.

Table 5.3: Statistical significance of the difference in overall quality among content categories; only p-values smaller than .05 are shown in the table.

	Animal	Architecture	Flower	Food	Landscape	Street	Sport	Artificial
Animal	-	-	-	-	.009	.02	-	.003
Architecture	-	-	.048	-	-	.001	-	.042
Flower	-	.048	-	.001	.001	-	-	.007
Food	-	-	.001	-	-	.001	.019	-
Landscape	.009	-	.001	-	-	.001	.003	-
Street	.02	.001	-	.001	.001	-	-	.001
Sport	-	-	-	.019	.003	-	-	.006
Artificial	.003	.042	.007	-	-	.001	.006	-

As done before for aesthetic appeal, we also here omitted the content categories of architecture, landscape and artificial photographs in the investigation of the

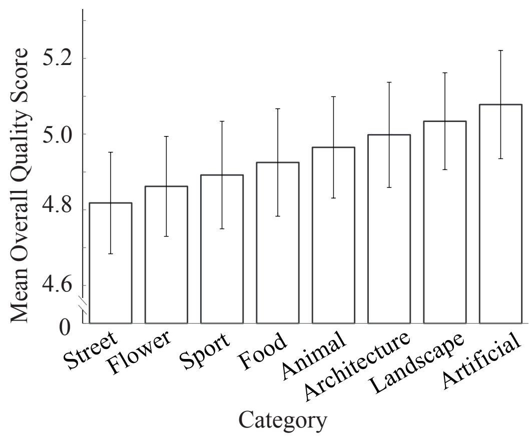

Figure 5.7: Mean overall quality score (±1 standard error) across content categories.

impact of depth of field on overall quality. Hence, we again performed a 3 (Depth of field) × 5 (Content category) repeated measure ANOVA, but now with the overall quality scores as the dependent variable. It revealed that Depth of field had a significant effect on overall quality ($F(2,62) = 5.27$, $p = .008$), with overall quality being higher for larger depth of field, as shown in Figure 5.8. The LSD post-hoc test for the effect of depth of field on overall quality showed that photographs with a large depth of field were rated significantly higher than photographs with a small depth of field ($p = .008$), or a medium depth of field ($p = .008$). There was no significant difference in overall quality between photographs with a small or medium depth of field.

We also found a significant interaction between Depth of field and Content category ($F(8, 248) = 7.99$, $p < .001$). Figure 5.9(a) shows how the influence of depth of field on overall quality differed across content categories. Depth of field significantly affected overall quality for the photographs of animals ($F(2,62)=9.86$, $p< .001$), flower ($F(2,62)=13.16$, $p< .001$), and sport ($F(2,62)=18.26$, $p< .001$), whereas it had no significant effect on overall quality for the photographs of food or street. To investigate whether the effect of depth of field on overall quality for these three content categories was systematic or caused by the specific selection of photographs used in this experiment, we repeated the analysis selecting randomly five photographs in each level of depth of field for the content categories animal, flower and sport. Figure 5.9 (b), (c), and (d) show the effect of depth of field across eight different samples (i.e., different subsets of the original photographs). It can be seen in Figure 5.9 (c) and (d) that the effect of depth of field was consistent in most samples for photographs of flowers and sport; for photographs of flowers the medium

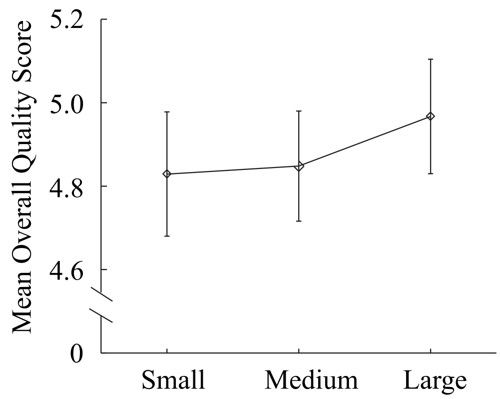

Figure 5.8: Main effect of depth of field on overall quality with mean 1 standard error.

depth of field reduced the overall quality compared to a small or large depth of field, whereas for photographs of sport a medium depth of field improved the overall quality compared to a small depth of field, and was comparable in overall quality to a large depth of field. For photographs of animals, however, we found that the effect of depth of field on overall quality was scattered, which suggests that the specific selection of photographs used in this experiment may have had an effect on overall quality.

Correlation between aesthetic appeal and overall quality

Figure 5.10 shows a scatter plot of the mean aesthetic appeal score against the mean overall quality score, with each point representing a photograph. A correlation analysis showed that there was a positive significant relationship between aesthetic appeal and overall quality with a Pearson correlation coefficient of $0.775(\sqrt{0.601})$. Table 5.4 shows the correlation coefficient per content category, and suggests that aesthetic appeal and overall quality are positively correlated for every content category. Figure 5.11 shows that the significant correlation is not influenced by the depth of field.

Table 5.4: The correlation between aesthetic appeal and overall quality for the eight different content categories.

Category	Animal	Architecture	Flower	Food	Landscape	Street	Sport	Artificial
Correlation coefficients	0.586	0.615	0.598	0.509	0.539	0.643	0.683	.64

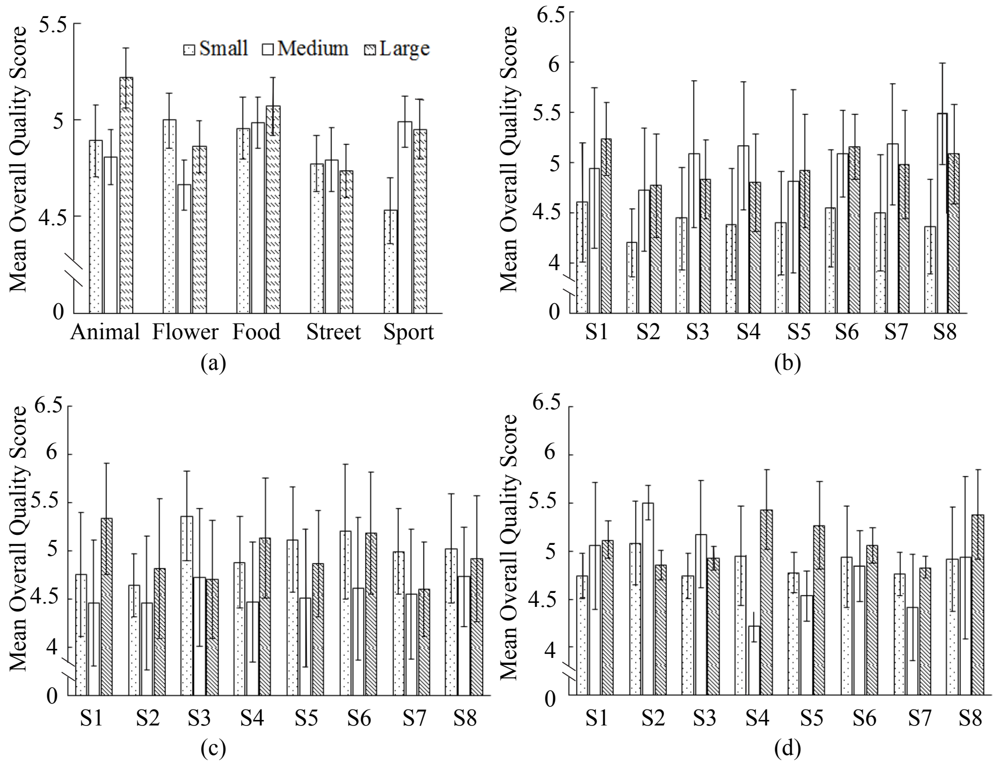

Figure 5.9: (a) Influence of depth of field on overall quality across categories. This effect is evaluated for different sets of selected photographs in (b) for photographs of animals , (c) photographs of flower, and (d) photographs of sport. All graphs show the mean ±1 standard error.

Effect of controlled depth of field on aesthetic appeal and overall quality

We considered that our findings on the effect of depth of field on aesthetic appeal and overall quality might to some extent be confounded with the particular set of photographs that we selected. Hence, we also investigated the effect of depth of field on both attributes for a particular set of photographs, for which we had full control on the scene content and depth of field. This set of photographs consisted of the content category artificial, for which the depth of field was changed at three levels. All other relevant parameters, such as the image content, lighting, and camera settings were controlled.

The example stimuli used as the artificial photographs, with three controlled

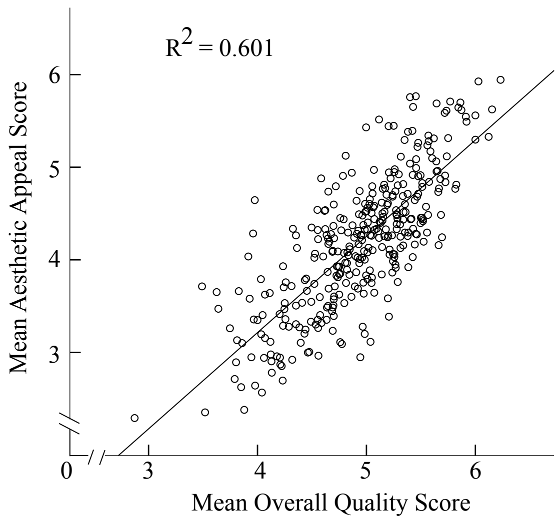

Figure 5.10: The correlation between aesthetic appeal and overall quality in general.

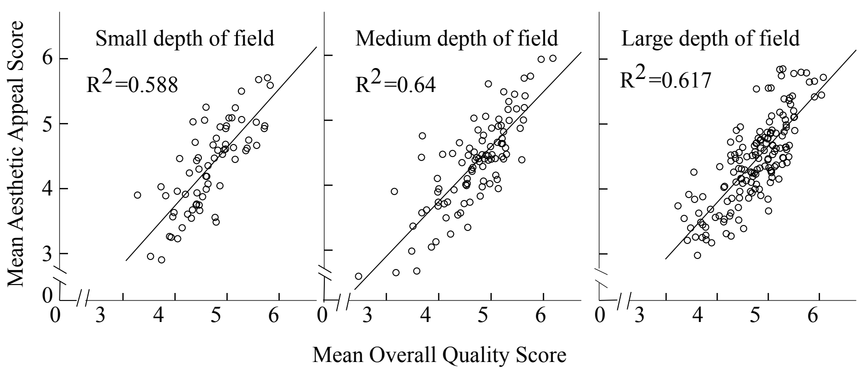

Figure 5.11: The correlation between aesthetic appeal and overall quality for the three depth of field levels separately.

values of depth of field (i.e., 5.7mm, 18.4mm, and 68.9mm) are shown in Figure 5.12. The aesthetic appeal and overall quality scores averaged across participants and scenes are shown as a function of depth of field in Figure 5.13. We performed a 4 (Scene)×3 (Depth of field) repeated measure ANOVA on the ratings of aesthetic appeal and overall quality, and found that Scene did not influence aesthetic appeal significantly, but influenced overall quality significantly ($F(3, 93) = 5.31$, $p = .002$). The third scene was scored the lowest among the four scenes. Depth of field was found to significantly influence the aesthetic appeal ratings ($F(2,62) = 9.76$, $p = .004$), whereas it was not a main significant factor for the overall quality ratings ($F(2, 62) = 2.77$, $p = .07$). The interaction between Scene and Depth of field was found to be significant for both aesthetic appeal and overall quality ($F(6, 186) = 12.07$, $p < .001$ and $F(6, 186) = 3.63$, $p = .002$, respectively).

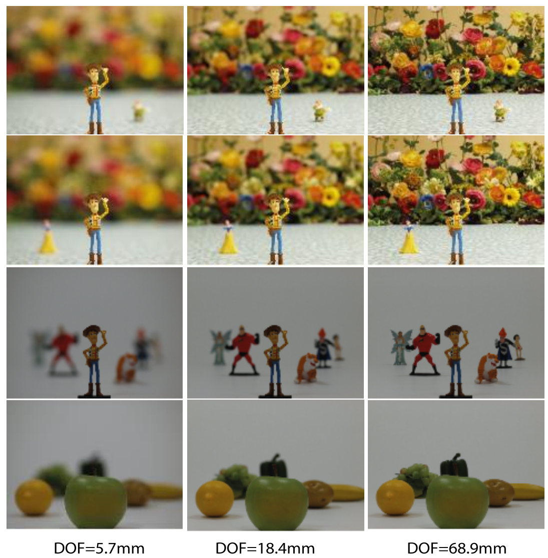

Figure 5.12: Stimuli with controlled depth of field.

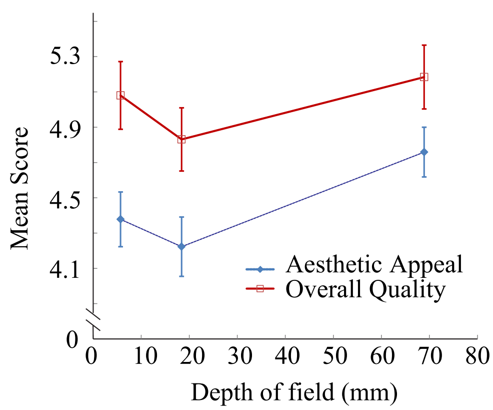

Figure 5.13: Mean aesthetic appeal (±1 standard error) as a function of depth of field for artificial photographs.

Results of the questionnaire

Although we were mainly interested in how depth of field affected aesthetic appeal and overall quality, we asked the participants to fill in a questionnaire to better understand which other image characteristics could potentially influence aesthetic appeal and overall quality. When answering the first and second question, regarding their standards for aesthetic appeal and overall quality (see Table 5.1), the participants could write down as many terms as they wanted to express their standard for good aesthetic appeal or overall quality. We collected these terms, and ranked them based on how many times they were mentioned by the participants. 31 out of 32 participants answered these questions, and the results are shown in Table 5.5 as percentages, calculated by the total number of the participants who mentioned the term divided by the total number of the participants who answered the question (times 100). We only picked the terms that were mentioned by at least five participants, and ignored those terms that were mentioned fewer times. The responses showed that color, composition, image content, and the main subject (i.e., the salient subject) were the major factors that influenced aesthetic appeal whereas sharpness, color, artifacts, and focus were important for overall quality.

Table 5.5: Summary of characteristics influencing aesthetic appeal and overall quality according to the participants.

Aesthetic appeal characteristics	Percentage (n=31) %	Overall quality characteristics	Percentage (n=31) %
Color	61.3	Sharpness	51.6
Composition	41.9	Color	32.3
Image content	29	Artifacts	25.8
Main subject	22.6	Focus	22.6
Focus	16.1	Luminance	19.4
Sharpness	16.1	Main subject	19.4
Contrast	16.1	Contrast	16.1

To better understand whether participants' preference on content category has any influence on aesthetic appeal and overall quality, we also asked participants to list the content categories they like best and worst separately (questions 10 and 11 in Table 5.1). They could answer with multiple categories. 21 out of 32 participants answered this question and we show these results in Table 5.6, calculating the percentage in the same way as we did in Table 5.5. Whether participants expected depth of field to have any effect on aesthetic appeal or overall quality was asked once they finished the experiment (questions 12 and 13). We found that 64.5% of the participants expected depth of field to have

Table 5.6: Participants preference on content category.

Image category	Like best (n=21) %	Like worst (n=21) %
Landscape	76.2	0
Animal	33.3	4.8
Flower	28.6	4.8
Food	23.8	19
Architecture	14.3	4.8
Street	14.3	38.1
Sport	14.3	28.6

an effect on the aesthetic appeal of the photographs, whereas 16.1% of the participants expected no influence of depth of field on aesthetic appeal. The remaining 19.3% of the participants did not know whether depth of field had any effect or not. For overall quality, the corresponding percentages were: 41.9% expected an effect, 35.5% expected no effect, and 22.5% were not sure.

Comparison between experienced participants and naive participants

According to the third and fifth question in the questionnaire (i.e., regarding the degree of expertise in taking photographs and the extent to which doing so was a habit, again see Table 5.1), we separated the participants into two groups, namely: experienced participants and naive participants. If the scores for the third and fifth question were both higher than 4, the participant was regarded as experienced in the sense of liking to take photographs and having a lot of experience in doing so. Otherwise, the participant was considered as naive. In total, we had 15 experienced participants and 17 naive participants.

Figure 5.14(a) and (c) show that the experienced participants' MAAS and MOQS are smaller than the scores from naive participants. Additionally, the effect of content category on aesthetic appeal and overall quality is somewhat different between experienced and naive participants. Figure 5.14 (b) shows that there is an increasing trend of MAAS with increasing depth of field for experienced participants, whereas the trend from naive participants is the opposite. In terms of overall quality, we retrieve instead an increasing trend of MOQS with depth of field for both experienced and naive participants, as illustrated in Figure 5.14(d). To check the above mentioned effects, we performed a 3 (Depth of field) $\times$ 5 (Content category) $\times$ 2 (Expertise) repeated-measures ANOVA on overall quality and aesthetic appeal, in which Depth of field and Content category were within-subject factors and Expertise was the between-subject factor. The statistical analysis showed that the difference in aesthetic appeal between experienced and naive participants was not significant (F(1,

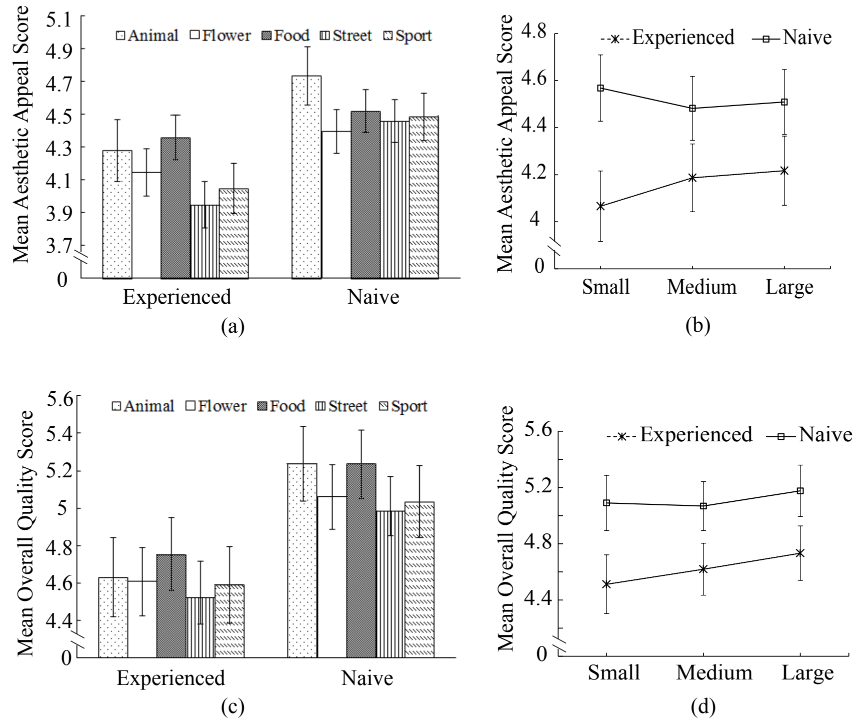

Figure 5.14: The effect of (a) content category on aesthetic appeal (b) depth of field on aesthetic appeal; (c) content category on overall quality; (d) depth of field on overall quality assessed by experienced and naive participants. All graphs show the mean ± 1 standard error.

30) = 3.44, p = 0.073). We did find an interaction for both Content category (F(4, 120) = 2.79, p=0.029) and Depth of field (F(2, 60) = 4.07, p = 0.022) with expertise level; so, the effect of content category and depth of field on aesthetic appeal was different between experienced and naive participants. Also the difference in overall quality between experienced and naive participants was not significant (F(1, 30) = 3.45, p = 0.073). In this case, we did not find a significant interaction of Expertise with Content category or Depth of field. We only found that the interaction between Category and Depth of field was significantly different between experienced and naive participants (F(8, 240) = 2.08, p = .038).

5.4 Discussion

5.4.1 Subjective perception of depth of field

In the preliminary experiment, we found that all photographs in the content categories of architecture and landscape were perceived as having a large depth of field, whereas none of the photographs in the artificial content category were perceived as having a large depth of field. Before discussing possible reasons, we would like to emphasize that all but the artificial photographs were selected from Flickr, google+, or personal collections without paying attention to its visual quality deliberately, and as such are assumed to represent a typical set of photographs for a given content category. The common property between photographs of landscapes and architecture is that the focal distance for capturing the scene is always large, whereas the focal distance for our artificial photographs is much smaller. Another difference is that the scale of the scene of architecture or landscapes is much larger than that of the artificial photographs. As such, we suggest dedicating less attention to photographs of architecture or landscapes when investigating the effects of depth of field, while the variation of depth of field in photography at a close focal distance should be considered more carefully.

5.4.2 Subjective evaluation of aesthetic appeal and overall quality

We found that the subjective ratings on aesthetic appeal and overall quality were both significantly influenced by the content category, which confirms the previous literature that overall quality is affected by image content (Eckert and Bradley, 1998; Good et al., 1994). The mean scores were higher for photographs of landscapes than for other categories in both aesthetic appeal and overall quality, whereas the mean scores for street photographs were the lowest in aesthetic appeal and overall quality. This finding is in agreement with the results of the questionnaire on the participants' preference per content category: 76% of the participants liked photographs of landscapes most, whereas 38% of the participants disliked photographs of streets most.

The judgment of aesthetic appeal is a highly subjective task, and as such expected to be influenced by many individual differences such as taste and lifestyle. In addition, there is no absolutely agreed standard for measuring aesthetic appeal. Compared to aesthetic appeal, overall quality is more clearly defined. Researchers have investigated many factors that influence overall quality such as color, sharpness, and distortions (Liu et al., 2010; Redi et al., 2009). Hence, we expected that the evaluation of overall quality was more stable than that of aesthetic appeal. However, our results showed that participants achieved high agreement on aesthetic appeal, at least to the same degree or even slightly better than on overall quality.

Although depth of field was not found to be a significant main factor on aesthetic appeal in general, a significant interaction between depth of field and content category was found. Our results showed that medium depth of field was considered leading to higher aesthetic appeal in photographs of sport and lower aesthetic appeal in photographs of animals as compared to the small and large depth of field. Previous research revealed photographic characteristics that impact aesthetic appeal, such as differences in semantics or clutter (Cerosaletti and Loui, 2009), and so, based on this research we may infer the reasons for the difference in the effect of depth of field on aesthetic appeal between photographs of animals or sport. For example, for animal photographs, observers may either be attracted by the animal itself which requires a small depth of field, or they may be interested in seeing the environment where the animal is living in as well, which requires a large depth of field. Photographs of sport often contain apparent motion, and so, a medium depth of field may show the story of the scene clearly and occlude the distracting background. An alternative explanation may result from the effect of depth of field on depth perception. Previous research (Held et al., 2010; Vishwanath, 2014; Vishwanath and Blaser, 2010) concluded that depth of field influenced the perception of egocentric distance quantitatively. A small depth of field causes pictorial objects to appear closer, which is more similar to the real-world situation, in which people prefer to watch animals at a close distance. Hence photographs of animals with small depth of field were scored more beautiful. For photographs of sports, observers may prefer the scene to be either very close or very far, and therefore score photographs with small or large depth of field higher than medium depth of field.

Contrary to aesthetic appeal, depth of field was found to significantly influence overall quality. Here we infer that depth of field blur is mainly regarded as image degradation when observers judge the overall quality of photographs. This finding implies that, when designing blur quality metrics or de-blurring algorithms, the distinction between depth of field blur and "distortion" blur, as pointed out in (Liu et al., 2011), is superfluous. According to Table 5.5, it is clear that sharpness is the feature of a photograph mentioned most to affect its overall quality. The reason may be that the size of the depth of field was not controlled properly, leading to blur in areas where participants were interested in. On the other hand, it is not likely that a majority of the photographs we selected had a sub-optimal depth of field. The interaction between overall quality and content category was significant as well. Similar as for aesthetic appeal, the effect of depth of field on overall quality was different across content categories, suggesting that each content category has its own optimal size of the depth of field.

5.4.3 Correlation between aesthetic appeal and overall quality

Perceived aesthetic appeal and overall quality were found to be significantly positively correlated with each other. This corroborates results found in literature (Cerosaletti et al., 2011; Redi and Heynderickx, 2012), showing that mean aesthetic ratings increase with increasing mean overall quality ratings. This finding was consistent across content categories and also across different levels of depth of field. The criterion that photographs with higher overall quality always lead to higher aesthetic appeal is common for general content. Manipulating depth of field does not influence this conclusion.

5.4.4 Comparison between experienced and naive participants

Our current study showed that experienced participants scored aesthetic appeal differently from naive participants for variations in content category and depth of field. For experienced participants, the influence of content category on aesthetic appeal was obvious, whereas this effect was relatively small for naive participants. We found that experienced participants scored aesthetic appeal higher with increasing depth of field, whereas naive participants scored aesthetic appeal highest when depth of field was small. These findings suggest that experienced participants consider blur introduced by depth of field as degrading aesthetic appeal, which is in agreement with previous studies arguing that sharpness is a visual property that makes a photograph more beautiful (Datta et al., 2006; Tinio et al., 2011). For naive participants, the results confirm that a small depth of field may improve the aesthetic appeal in photographs (Datta et al., 2006; Dhar et al., 2011). Hence, it can be concluded that the existing and often taken for granted conclusion that a small depth of field increases image beauty is only partially true, and its applicability depends on the expertise of the observers. Furthermore, these results also suggest that the experienced participants would have chosen a different depth of field (i.e., camera aperture setting) when they would make the photograph themselves.

As to overall quality, we did not find a significant difference between experienced and naive participants for the effects of content category and depth of field. Both groups of participants considered blur introduced by depth of field as decreasing overall quality. Based on the input from the questionnaire, we noticed that sharpness was regarded as the most important attribute when judging overall quality, whereas it was not that important for the assessment of aesthetic appeal. We may conclude that there is no need to consider whether the participants are experienced or naive when judging overall quality.

5.5 Conclusion

Based on this study, we may conclude that there is no need to consider depth of field in photographs in general when (objectively) assessing their aesthetic appeal, but that depth of field is important for the assessment of overall quality. Our results indicate indeed that a smaller depth of field compromises overall quality, possibly due to the induced lack of sharpness; we can thereby conclude that blur should be detected and when possible eliminated to guarantee a high overall quality of the photograph.

Our study also revealed that depth of field plays a significant different role across content categories on both aesthetic appeal and overall quality. Thus, to researchers working in the area of computational visual quality, we suggest that depth of field is not a common criterion across content categories and that the systems or models for evaluating the aesthetic appeal or overall quality automatically may benefit a lot from weighing the impact of depth of field depending on content category.

Our results confirm that both aesthetic appeal and overall quality are influenced by content category, suggesting that it is worthwhile to investigate the visual quality of photographs or videos under consideration of the actual content. We also confirm that the subjective evaluation of aesthetic appeal is strongly correlated with the evaluation of overall quality, and this relationship is not affected by content category or depth of field. Finally, the estimation of aesthetic appeal is influenced by the background of the observers whereas the perception of overall quality is not.

Bibliography

Axelsson, s. (2007). Towards a psychology of photography: dimensions underlying aesthetic appeal of photographs. *Perceptual and Motor Skills*, 105(2):411–434.

Biemond, J., Lagendijk, R. L., and Mersereau, R. M. (1990). Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883.

Cerosaletti, C. D. and Loui, A. C. (2009). Measuring the perceived aesthetic quality of photographic images. *International Workshop on Quality of Multimedia Experience*, pages 47–52.

Cerosaletti, C. D., Loui, A. C., and Gallagher, A. C. (2011). Investigating two features of aesthetic perception in consumer photographic images: clutter and center. In *SPIE Electronic Imaging*, pages 786507–786507–8.

Crete, F., Dolmiere, T., Ladret, P., and Nicolas, M. (2007). The blur effect: perception and estimation with a new no-reference perceptual blur metric. In *SPIE Human Vision and Electronic Imaging XII*, pages 64920I–64920I–11.

Datta, R., Joshi, D., Li, J., and Wang, J. (2006). *Studying Aesthetics in Photographic Images Using a Computational Approach*, volume 3953 of *Lecture Notes in Computer Science*, chapter 23, pages 288–301. Springer Berlin Heidelberg.

de Ridder, H. (1998). Psychophysical evaluation of image quality: from judgment to impression. In *SPIE Human Vision and Electronic Imaging III*, pages 252–263. International Society for Optics and Photonics.

de Ridder, H. (2001). Cognitive issues in image quality measurement. *Journal of Electronic Imaging*, 10(1):47–55.

Dhar, S., Ordonez, V., and Berg, T. L. (2011). High level describable attributes for predicting aesthetics and interestingness. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1657–1664.

Eckert, M. P. and Bradley, A. P. (1998). Perceptual quality metrics applied to still image compression. *Signal processing*, 70(3):177–200.

Fedorovskaya, E. A. (2002). Perceived overall contrast and quality of the tone scale rendering for natural images. In *SPIE Human Vision and Electronic Imaging VII*, volume 4662, pages 119–128.

Ferzli, R. and Karam, L. J. (2009). A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *IEEE Transactions on Image Processing*, 18(4):717–728.

Good, W. F., Maitz, G. S., and Gur, D. (1994). Joint photographic experts group (jpeg) compatible data compression of mammograms. *Journal of digital imaging*, 7(3):123–132.

Held, R. T., Cooper, E. A., Obrien, J. F., and Banks, M. S. (2010). Using blur to affect perceived distance and size. *ACM Transactions on Graphics*, 29(2).

Hillaire, S., Lcuyer, A., Cozot, R., and Casiez, G. (2007). Depth-of-field blur effects for first-person navigation in virtual environments. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 203–206.

Hobfeld, T., Schatz, R., and Egger, S. (2011). Sos: The mos is not enough! In *Third International Workshop on Quality of Multimedia Experience (QoMEX)*, pages 131–136.

IJsselsteijn, W. A., de Ridder, H., and Vliegen, J. (2000). Effects of stereoscopic filming parameters and display duration on the subjective assessment of eye strain. In *SPIE Stereoscopic Displays and Virtual Reality Systems VII*, volume 3957, pages 12–22.

Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q.-T., Wang, J. Z., Li, J., and Luo, J. (2011). Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5):94–115.

Kaptein, R. G., Kuijsters, A., Lambooij, M. T., IJsselsteijn, W. A., and Heynderickx, I. (2008). Performance evaluation of 3d-tv systems. In *SPIE Electronic Imaging*, pages 680819–680819–11.

Ke, Y., Tang, X., and Jing, F. (2006). The design of high-level features for photo quality assessment. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 419–426.

Kuijsters, A., Ijsselsteijn, W. A., Lambooij, M. T., and Heynderickx, I. E. (2009). Influence of chroma variations on naturalness and image quality of stereoscopic images. In *SPIE Electronic Imaging*, pages 72401–72401–9.

Liu, H., Klomp, N., and Heynderickx, I. (2010). A no-reference metric for perceived ringing artifacts in images. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(4):529–539.

Liu, H., Wang, J., Redi, J., Le Callet, P., and Heynderickx, I. (2011). An efficient no-reference metric for perceived blur. In *Third European Workshop on Visual Information Processing (EUVIP),*, pages 174–179.

Loui, A., Wood, M. D., Scalise, A., and Birkelund, J. (2008). Multidimensional image value assessment and rating for automated albuming and retrieval. In *IEEE International Conference on Image Processing*, pages 97–100.

Luo, Y. and Tang, X. (2008). Photo and video quality evaluation Focusing on the subject. *Computer VisionECCV*, pages 386–399.

Marziliano, P., Dufaux, F., Winkler, S., and Ebrahimi, T. (2002). A no-reference perceptual blur metric. In *International Conference on Image Processing*, volume 3, pages III–57–III–60 vol. 3.

Marziliano, P., Dufaux, F., Winkler, S., and Ebrahimi, T. (2004). Perceptual blur and ringing metrics application to jpeg2000. *Signal processing: Image communication*, 19(2):163–172.

Nefs, H. T. (2012). Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing and perceiving*, 25(6):577–595.

Obrador, P., Saad, M. A., Suryanarayan, P., and Oliver, N. (2012). *Towards category-based aesthetic models of photographs*, volume 7131. Springer.

Ogle, K. N. and Schwartz, J. T. (1959). Depth of focus of the human eye. *Journal of the Optical Society of America*, 49(3):273–279.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):523–531.

Redi, J., Gastaldo, P., Zunino, R., and Heynderickx, I. (2009). Reduced reference assessment of perceived quality by exploiting color information. In *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*.

Redi, J. and Heynderickx, I. (2012). Image integrity and aesthetics: towards a more encompassing definition of visual quality. In *SPIE Electronic Imaging*, pages 829115–829115–10.

Redi, J., Liu, H., Alers, H., Zunino, R., and Heynderickx, I. (2010). Comparing subjective image quality measurement methods for the creation of public databases. In *IS&T/SPIE Electronic Imaging*, pages 752903–752903–11.

Roufs, J. A. (1989). Brightness contrast and sharpness, interactive factors in perceptual image quality. In *Human Vision, Visual Processing, and Digital Display*, volume 1077, pages 66–72.

Seuntiens, P., Meesters, L., and Ijsselsteijn, W. (2006). Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric jpeg coding and camera separation. *ACM Transactions on Applied Perception*, 3(2):95–109.

Szechter, L. E. and Liben, L. S. (2007). Children's aesthetic understanding of photographic art and the quality of artrelated parentchild interactions. *Child development*, 78(3):879–894.

Tinio, P. P., Leder, H., and Strasser, M. (2011). Image quality and the aesthetic judgment of photographs Contrast, sharpness, and grain teased apart and put together. *Psychology of Aesthetics, Creativity, and the Arts*, 5(2):165.

Vishwanath, D. (2007). Is focal blur an absolute cue to depth? *Journal of Vision*, 7(9):845–845.

Vishwanath, D. (2014). Toward a new theory of stereopsis. *Psychological review*, 121(2):151.

Vishwanath, D. and Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision*, 10(10).

Yuan, L., Sun, J., Quan, L., and Shum, H.-Y. (2007). Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics*, 26(3):1.

Zhang, T., Nefs, H. T., and Heynderickx, I. (2014a). Human discrimination of depth of field in stereoscopic and nonstereoscopic photographs. *Perception*, 43(5):368–380.

Zhang, T., O'hare, L., Hibbard, P. B., Nefs, H. T., and Heynderickx, I. (2014b). Depth of field affects perceived depth in stereographs. *ACM Transactions on Applied Perception*, 11(4):18.

Chapter 6

Conclusion and discussions

The main aim of this thesis is to investigate the role of depth of field in pictorial and solid space from a perceptual level to a higher cognitive level. To this end, four studies were conducted, of which the first two investigated depth of field on a fairly low level of perception: namely, the psychophysical discrimination of depth of field and the effects of depth of field on depth perception, respectively. Two more studies were conducted to investigate the role of depth of field on a more cognitive level: namely, on the effect of depth of field on change detection, and on aesthetic appeal and overall quality, respectively. In this chapter, I will first summarize the experimental findings in this thesis and relate them to the main research questions that were raised in Chapter 1. Then, I will proceed to discuss the limitations and contributions of my work.

6.1 Answers to the research questions

A prerequisite for studying depth of field is the fundamental knowledge on how well humans are capable of discriminating depth of field and how well these discrimination thresholds can be predicted from blur detection and blur discrimination thresholds. Since this information is surprisingly lacking from the literatures, the first question that I raised in this thesis was:

1. What is the discrimination threshold of depth of field in (stereo) pictures and do factors such as viewing condition, image content, or scale of the scene influence the threshold of the depth of field?

Three factors, namely viewing condition, image content, and scale of the scene, were manipulated for a small and large reference value of depth of field to measure the discrimination threshold in depth of field. It was found that the increment threshold for a small depth of field was much smaller than that for a large depth of field. Viewing condition and image content did not influence the discrimination of depth of field significantly, but whether the scale of the scene was smaller or larger did significantly affect the discrimination threshold. After analysing the empirical data, the measured discrimination thresholds of depth of field were compared to predicted values, which were calculated from blur discrimination thresholds by previous researchers (Chen et al., 1989; Hess et al., 1989; Watson and Ahumada, 2011). The measured discrimination threshold for a small depth of field was smaller than what was predicted from blur discrimination.

I concluded that humans are much more sensitive to changes in small depths of field than in large depths of field. This conclusion holds for both stereo and non-stereo viewing conditions. Binocular disparity does not influence the discrimination of depth of field. I further concluded that the depth structure in the picture influences observers ability to discriminate depth of field, since the increment threshold decreases with increasing scale of the scene. From the comparison between the measured and predicted discrimination thresholds, I can draw the conclusion that observers are more sensitive to changes in depth of field than what would be expected from the known values of blur discrimination. Using a single level of blur in a picture that contains limited depth of field may underestimate observers ability to discriminate depth of field.

One of the main reasons that depth of field has attracted researchers' attention is that it can be used to enhance depth perception (Nefs, 2011) and also because the depth of field can be used to construct a depth map from an image (Pentland, 1987). Notably, previous studies were mainly performed using non-stereo pictures. Therefor, it remains unclear whether the depth of field has the same effect on depth perception in stereo pictures as in non-stereo pictures. In the second study we thus raised the question:

2. How does depth of field influence perceived depth when there are other strong depth cues such as binocular disparity in pictures?

To answer this question, I conducted an experiment with a Wheatstone stereoscope and a 3D TV, while manipulating depth of field and physical depth in the scene. It was found that both depth of field and physical depth significantly influence the perceived depth under stereo viewing conditions. When compared to the effects of depth of field under non-stereo viewing conditions (Nefs, 2011), I found that binocular disparity influences the effects of depth of field on depth perception. In non-stereo pictures, perceived depth decreases monotonically with increasing depth of field, whereas perceived depth increases with increasing depth of field at small depth of field and short viewing distances in stereo pictures. When depth of field is large, perceived depth decreases with increasing depth of field, and is then not influenced by binocular disparity.

I concluded that the presence of strong depth cues such as binocular disparity may weaken the effects of depth of field as a depth cue when the depth of field is small. When depth of field is not small, perceived depth increases with increasing depth of field no matter whether binocular disparity is zero or not.

Apart from the effects of depth of field on depth perception, another feature of depth of field is that it introduces a blur gradient in pictures. Pictures with a small depth of field contain sharp areas, but also areas with different amounts of blur. This leads to the intuitive question whether observers' attention gravitates to the sharp areas. Hence, in a third study I investigated the research question:

3. How does depth of field influence change detection in pictorial scenes and does binocular stereopsis influence such effects?

In this study, participants were required to detect a difference in a series of two alternating images. Three main factors were investigated: namely, depth of field, position of the change with respect to the focal plane, and stereo versus non-stereo viewing conditions. I found that a small depth of field facilitates change blindness and that the response time for change detection is longer in a scene with a small depth of field than in a scene with a large depth of field or with uniform blur. In case of a small depth of field, a change in the sharp area is detected faster and more easily than a change in the blurred foreground or background. These findings are similar for non-stereo or stereo images.

I concluded that the presence of binocular disparity does not influence the effects of depth of field on change detection; depth of field directs viewers attention similarly in both non-stereo and stereo viewing conditions. I also concluded that binocular disparity changes the role of the distance in depth on change detection from the change object to the object closest to the observer. In non-stereo pictures, depth of field can enhance depth perception (Nefs, 2011; Pentland, 1987), in another word, showing an object is closer or further away. However, such effects do not mean that changes closer to the observer are detected faster and easier. Instead, it was found that changes in the sharp area are detected faster. In stereo pictures, changes that are closer to the observer are detected faster and closer, but such trend was weak in pictures with small depth of field. When a picture with small depth of field is observed under stereo viewing conditions, the allocation of attention is mainly influenced by the effect of depth of field instead of binocular disparity.

The previous three studies investigated the effects of depth of field in pictorial scenes from the point of view of psychology. All the pictures used in these three experiments were created by myself in the lab. Photographers and cinematographers are more interested in the effects of depth of field on viewing experience in photographs. Hence, in the fourth study, I raised the following question:

4. How does depth of field influence the subjective evaluation of aesthetic appeal and overall quality in photographs without considering other features such as color, contrast, and so on?

To answer this question, participants were asked to assess 339 photographs, differing in content category and depth of field, on aesthetic appeal and overall quality. I found that the specific content category of a photograph influences its aesthetic appeal and overall quality significantly. Depth of field does influence overall quality significantly, but not aesthetic appeal. Furthermore, depth of field plays a different role in both aesthetic appeal and overall quality across content categories.

This study provides guidelines on how to use depth of field to improve aesthetic appeal and overall quality. Generally speaking, there are no common rules on how to manipulate depth of field to make photographs more beautiful. For photographs of animals, either a small or large depth of field may yield a more beautiful photograph than a medium depth of field. For photographs of sport, a medium rather than a small or a large depth of field makes the photograph more beautiful. As to overall quality, a general conclusion is that the aesthetic appeal and overall quality improves with increasing depth of field. Furthermore, I also concluded that higher aesthetic appeal leads to higher overall quality and vice versa.

6.2 Limitations

Like all scientific research, my thesis surely has its limitations. First, the four studies that are discussed in this thesis are limited to the effects of depth of field in static pictures /photographs. We do not extend our research into the areas in films, gaming, or virtual reality, where the performance of depth of field has rich research potential. It was found in the previous study (Hillaire et al., 2007) that depth of field can influence players performance and experience in gaming. In virtual reality or gaming, it is possible to use depth of field to enhance realism, and as such to improve the immersion of observers. However, the findings in this thesis still cannot reflect the perception of depth of field in these applications.

Second, this thesis only addressed the perception and cognition of depth of field related to optical systems. For instance, this research mainly focuses on depth of field that is created by cameras and presented in pictures/photographs. Apart from the perceived depth of field, there are researches on depth of field of human eye (Campbell, 1957; Marcos et al., 1999). However, the relationship between these two areas has not been well explored. The only existing research was performed by O'Hare et al. (2013), where they found that the inconsistency between the depth of field in pictures and that in the human eyes does not cause extra viewing discomfort when observing stereo pictures. However, it is still not clear how depth of field of the human eye influences the perception and cognition of depth of field in pictures or photographs.

Third, depth of field in this thesis is mostly decided by parameters of the optical systems such as focal distance, and metadata of camera settings. Such depth of field is not always what observers expect because the value of depth of field is fixed as soon as the picture is created. Researchers investigated how to create depth of field as the observers expected, that is semantic depth of field (Kosara et al., 2001). Semantic depth of field, similar as depth of field, is found to be able to direct observers attention and enhance video saliency (Su and Takahashi, 2010). We may expect that semantic depth of field can have similar effects as found in this thesis, however, this still has to be tested.

6.3 Contributions

This thesis addressed questions related to an important picture characteristic, namely depth of field. Depth of field, as a depth cue (Pentland, 1987; Vishwanath and Blaser, 2010), has already been investigated a lot in the literature. As investigated in the previous study (Vishwanath, 2014), depth of field can affect the impression of stereopsis of pictures. Depth of field and stereopsis can exist simultaneously in a picture, but what will happen to the perception of the picture is still not clear. Additionally, as we may notice that depth of field also attracts photographers attention because it is an important photographic technique. Hence, it is interesting to look into details about how depth of field influences the quality of experience of the photographs. This thesis raised four research questions to address the characteristics of depth of field in 3D picture perception and quality of experience of photographs. In the end, I found that stereopsis influences the effects of depth of field on depth perception. Meanwhile, it was found that depth of field can also influence the effects of stereopsis on change detection. Photographers can also find useful suggestions in this thesis about how depth of field influences the aesthetic appeal and overall quality in photographs. As such, this thesis not only gives practical advices to photographers and cinematographers but also has scientific contributions.

6.3.1 Practical advice

The first practical suggestion of this thesis can benefit people who are interested in using depth of field. This thesis is one of the first to investigate how to manipulate depth of field in both non-stereo and stereo pictures effectively. Researchers and photographers are suggested to select depth of field more carefully in the range of small values because humans are very sensitive to the changes in this region.

The second practical suggestion is on when to decide depth of field is an important image attribute. The findings are beneficial to researchers in the area of computational aesthetics and image quality. For example, when they build a general model to automatically estimate overall quality of pictures, they are suggested to consider depth of field as an image attribute in the model. However, when building a model to automatically assess aesthetic appeal, there is no need to take depth of field into account. However, when building a model for a specific category of pictures such as animals, depth of field turns out to be an important image attribute for automatically assess aesthetic appeal and overall quality.

6.3.2 Theoretical contribution

The main theoretical contribution of this thesis is to reflect the relationship between depth of field and binocular disparity in depth perception and change detection. Binocular disparity is a strong depth cue that can provide accurate depth information. Depth of field, as a monocular depth cue, may compress the pictorial scene in non-stereo pictures (Nefs, 2011). This thesis investigates how perceived depth is specified when there are two depth cues: namely, binocular disparity and depth of field. When binocular disparity is small, the effect of depth of field on perceived depth is weak. When binocular disparity is large, the smaller the depth of field is, the larger the gap between perceived depth and stereoscopically presented depth is. In addition to the effect of binocular disparity on depth of field, this thesis also shows that other depth cues like height-in-the-field in a picture may weaken the effect of depth of field on perceived depth. To conclude, the weight of depth of field as a depth cue varies with other depth cues in pictures. From the point of view of pictures and stereopsis, this thesis suggests that the pictorial space is less expanded compared to the physical space with increasing depth of field when the size of depth of field is in a very small range. In other words, a very small depth of field may weaken the impression of stereopsis, that is, it compresses the impression of depth.

As reported in the literature, binocular disparity clearly illustrates the depth structure in pictorial scenes, and objects closer to the observer attract attention more quickly (Jansen et al., 2009; Mazza et al., 2005). Depth of field acts as a depth cue in both non-stereo and stereo pictures, and so also illustrates depth structure in pictorial scenes. However, objects that are perceived closer in scenes with a small depth of field do not automatically attract attention faster. Instead, objects in the sharp area of the picture attract attention faster, and this sharp area may be anywhere in the pictorial scene. Then we may be wondering that when a picture with small depth of field is observed under stereo viewing conditions, which one is the dominant factor that control the change detection. This thesis shows that a small depth of field weakens the effect of binocular disparity on change detection. In the stereo picture with small depth of field, the change in the sharp area is detected faster and closer even the area is farther away from the observer.

6.4 Future work and final thoughts

The effect of blur

It is found that the discrimination of depth of field is poorly predicted from the discrimination of blur. The prediction of discrimination threshold of depth of field is overestimated from any single level of blur contained in the blur gradients introduced by depth of field. It would be interesting to get a more detailed view on how the discrimination threshold for depth of field evolves with increasing reference values in order to get a finger behind the question why discrimination thresholds are poorly predicted from blur thresholds.

Weight of depth cues

The effects of depth of field on perceived depth is discussed in my work. However, a picture always contains multiple depth cues. Previous studies investigated how the combination of blur and disparity influences perceived depth (Langer and Siciliano, 2015; Mather and Smith, 2000; Wang et al., 2011b,a). As discussed in this thesis, depth of field blur is different from the blur mentioned in the literature, the future work can direct to the investigation of the weight of depth of field as a depth cue when there are other depth cues such as binocular disparity, motion parallax, and perspective.

Implementation of depth of field in other applications

With the findings in this thesis, we know how humans perceive depth of field and the characteristics of depth of field in pictures. One of the future directions to continue our research is to implement depth of field in applications such as virtual reality and gaming and then test how depth of field influences the immersion and performance of the participants.

In conclusion, this thesis explored the characteristics of depth of field in pictures from a low perception level to a high cognitive level. The findings in this thesis are very fundamental to understand the effects of depth of field in pictorial perception. This thesis reveals how depth of field and binocular stereopsis influences each other in depth perception and change detection, which hints at the use of depth of field in more practical applications such as gaming and virtual reality. In addition, photographers can benefit from the thesis on how to control and select depth of field in photography to improve observers' viewing experience. It also help cinematographers understand how to use depth of field properly to direct observers attention.

Bibliography

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

Chen, C.-C., Chen, K.-P., Tseng, C.-H., Kuo, S.-T., and Wu, K.-N. (1989). Constructing a metrics for blur perception with blur discrimination experiments. *IS&T/SPIE Electronic Imaging*.

Hess, R., Pointer, J., and Watt, R. (1989). How are spatial filters used in fovea and parafovea? *JOSA*, 6(2):329–339.

Hillaire, S., Lcuyer, A., Cozot, R., and Casiez, G. (2007). Depth-of-field blur effects for first-person navigation in virtual environments. *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 203–206.

Jansen, L., Onat, S., and Knig, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1).

Kosara, R., Miksch, S., and Hauser, H. (2001). Semantic depth of field. *IEEE Symposium on Information Visualization*, pages 97–104.

Langer, M. S. and Siciliano, R. A. (2015). Are blur and disparity complementary cues to depth? *Vision research*, 107:15–21.

Marcos, S., Moreno, E., and Navarro, R. (1999). The depth-of-field of the human eye from objective and subjective measurements. *Vision research*, 39(12):2039–2049.

Mather, G. and Smith, D. R. (2000). Depth cue integration: stereopsis and image blur. *Vision research*, 40(25):3501–3506.

Mazza, V., Turatto, M., and Umilt, C. (2005). Foregroundbackground segmentation and attention: A change blindness study. *Psychological Research*, 69(3):201–210.

Nefs, H. T. (2011). Depth of field affects perceived depth-width ratios in photographs of natural scenes. *Seeing and perceiving*, 25(6):577–595.

O'Hare, L., Zhang, T., Nefs, H. T., and Hibbard, P. B. (2013). Visual discomfort and depth-of-field. *i-Perception*, 4(3):156–169.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (4):523–531.

Su, Z. and Takahashi, S. (2010). Real-time enhancement of image and video saliency using semantic depth of field. *International Conference on Computer Vision Theory and Applications*, pages 370–375.

Vishwanath, D. (2014). Toward a new theory of stereopsis. *Psychological review*, 121(2):151–178.

Vishwanath, D. and Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of vision*, 10(10).

Wang, J., Barkowsky, M., Ricordel, V., and Le Callet, P. (2011a). Clarifying how defocus blur and disparity affect the perceived depth. pages 1–2.

Wang, J., Barkowsky, M., Ricordel, V., and Le Callet, P. (2011b). Quantifying how the combination of blur and disparity affects the perceived depth. *IS&T/SPIE Electronic Imaging*, pages 78650K–78650K–10.

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

Acknowledgement

In the end, I want to say that I really appreciate the help from many people during the time to finish my PhD project.

First, I would like to express my sincere gratitude to my promoter Prof. Ingrid Heynderickx for continuous effort on my PhD project. She is the person who guided me into perception area before I came to the Netherlands. I also thank her for her support and belief in me during these years of my PhD study.

My sincere thanks goes to my co-promoter and daily advisor Dr. Harold Nefs, without whom the thesis would have been less complete. It's quite difficult for me to use several sentences to express my appreciation of his supervision. I would like to thank him for the ideals that he shared with me, the research directions that he advised me. During the four years, he spent endless time in many helpful discussions to help me solve problems during developing the PhD project. In addition, I would also thank him for helping me get used to the life in the Netherlands when I arrived here and encouraged me when I came across difficulties and frustrations from the study and the life.

I also want to thank Prof. Catholijn Jonker who gave me a great support for my PhD research. As the head of the group, she was always patient to listen to the ideas from the PhD students and then took actions to make them work.

I would like to acknowledge the members of the Pi lab: I really appreciate all the useful suggestions from Judith Redi during our meetings and corporation, and I also like the delicious Italian food she brought a lot. Sylvia Pont, I thank her for organizing the meetings from which I broaden my view on the area of perception. I would thank Maarten Wijntjes, I spent quite some time in his lab to design and run my experiments. He always has cool ideas and staff to do research. I also want to thank all the other Pi lab members for the good time we had during these years: Yi, Ernestasia, Ling, Fan, Tatiana, Qian.

In addition, I would like to thank all my colleagues in II group, for all the fun we had together during these years. They also gave me different kinds of support in my research such as being the participants of my experiments. I got lots of suggestions to get used to Dutch culture from them during lunch time. Anita, as the secretory of the group, I can always get help from her for different kinds of staff. Ruud and Bart, I cannot remember how many times to go to their office to ask for help. I would like to thank all of them: Willem-Paul, Birna, Koen, Joost, Reyhan, Chris, Timea, Chang, Reyhan, Maaike, Corine, Myrthe, Vincent, Ni, Yun, Chao, Yangyang, Dwi, Wenxin, Junchao, Thomas, Alex, Dylan, Wouter, Hani, Tim, Iris, Christian, Tjerk, Vanessa, Andre, Christina, Arman, and so on.

I also want to express my gratefulness to my officemate: Iulia, Changyun, and Ernestasia. I had a lot of fun with them during my whole PhD life. We had so many discussions on research and culture difference, and shared happiness and sadness during these years.

I also want to express my gratefulness to my officemate: Iulia, Changyun, and Ernestasia. I had a lot of fun with them during my whole PhD life. We had so many discussions on research and culture difference, and shared happiness and sadness during these years.

To live in a culture totally different from the Chinese culture is a big challenge. My PhD life would not go smoothly without many of my Chinese friends: Ling Xia, Xiaojun Shi, Yun Ling, Chao Qu, Ni Kang, Junchao Xu, Yue Ding, Chang Wang, Wenxin Wang, Yangyang Shi, Hantao Liu and so on. Additionally, I would like to thank my housemate: Ling Xia, Danqian Shen, Tian Xu, Adeline, Linna Wei, Kun Yan. I had a lot of fun with them during the spare time which greatly supported me to live happily in the Netherlands. I also want to thank my friend, Jin Li, Jun Xu, Dan Ye, Yuan Jiang. Although we didn't not meet a lot, but I think I gained a lot of support from you via many voice and video calls. Besides, I would like to thank my friend Shan Jiang and her daughter Xuanxuan to contribute her toys to my study.

In the end, I especially want to thank my parents, my husband and sister, who have always supported my decisions and believed in me in this long journey in the Netherlands. I believe I cannot finish my PhD project without their love.

Tingting, 2015

Curriculum Vitae

Tingting Zhang was born in Liyang, China on December 3, 1986.

She finished the middle school in 2001 and high school in 2004 in Liyang. From 2004 to 2008, she was studying in Electronic Science and Technology in Southeast University in China for her bachelor's degree. From 2008 to 2011, she was studying in Optical Engineering in Southeast University for her master's degree. The title of her master thesis is "Numerical simulation of the gas discharge lamps based on fluid model". Since February of 2011, she started working as a PhD student in the Interactive Intelligence Group, Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, the Netherlands. Her research interest is on visual perception, 3D perception, visual attention, and Quality of Experience. During the time to pursue the doctor's degree, she mainly worked on the project "Perception and cognition of depth of field". She conducted four subjective studies to investigate depth of field from a low perception level to a relatively high cognitive level.