



Study of depth bias of observers in free viewing of still stereoscopic synthetic stimuli

Junle Wang, Patrick Le Callet, Sylvain Tourancheau, Vincent Ricordel,
Matthieu Perreira da Silva

► To cite this version:

Junle Wang, Patrick Le Callet, Sylvain Tourancheau, Vincent Ricordel, Matthieu Perreira da Silva.
Study of depth bias of observers in free viewing of still stereoscopic synthetic stimuli. *Journal of Eye
Movement Research*, 2012, 5 (5), pp. 1-11. hal-00730667

HAL Id: hal-00730667

<https://hal.science/hal-00730667>

Submitted on 10 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Study of depth bias of observers in free viewing of still stereoscopic synthetic stimuli

Junle Wang¹

Patrick Le Callet¹

Sylvain Tourancheau²

Vincent Ricordel¹

Matthieu Perreira Da Silva¹

1. LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, France

2. Mid Sweden University, Sundsvall, Sweden

Observers' fixations exhibit a marked bias towards certain areas on the screen when viewing scenes on computer monitors. For instance, there exists a well-known "center-bias" which means that fixations are biased towards the center of the screen during the viewing of 2D still images. In the viewing of 3D content, stereoscopic displays enhance depth perception by the mean of binocular parallax. This additional depth cue has a great influence on guiding eye movements. Relatively little is known about the impact of binocular parallax on visual attention of the 3D content displayed on stereoscopic screen. Several studies mentioned that people tend to look preferably at the objects located at certain positions in depth. But studies proving or quantifying this depth-bias are still limited. In this paper, we conducted a binocular eye-tracking experiment by showing synthetic stimuli on a stereoscopic display. Observers were required to do a free-viewing task through passive polarized glasses. Gaze positions of both eyes were recorded and the depth of eyes' fixation was determined. The stimuli used in the experiment were designed in such a way that the center-bias and the depth-bias affect eye movements individually. Results indicate the existence of a depth-bias: objects closer to the viewer attract attention earlier than distant objects, and the number of fixations located on objects varies as a function of objects' depth. The closest object in a scene always attracts most fixations. The fixation distribution along depth also shows a convergent behavior as the viewing time increases.

Keywords: Visual attention, 3DTV, stereoscopy, eye-tracking, depth bias

Introduction

Visual attention is one of the most important mechanisms deployed in the human visual system to reduce the complexity of scene analysis (Wolfe, 2000). Driven by visual attention, viewers can selectively focus their attention on specific areas of interest in the scene. In the last decades, great efforts have been put into the research of visual attention for 2D content viewing conditions. More recently, studies about visual attention in the viewing of stereoscopic 3D content have been gaining an increasing amount of attention, because of the emergence of 3D products (in cinema and home) and recent availability of high-definition 3D-capable equipments to acquire and display stereoscopic content (Huynh-Thu, Barkowsky, & Le Callet, 2011).

In the studies about the deployment of visual atten-

tion on planar screen, it has been found that observers' fixations exhibit a marked bias towards certain areas on the screen. In the viewing of 2D images or videos, a so-called "center-bias" (or "central fixation bias") has been demonstrated: gaze fixations are biased towards the center of the scene (Tseng, Carmi, Cameron, Munoz, & Itti, 2009; Tatler, 2007). However, in the viewing of 3D content on stereoscopic displays, the viewing behavior of observers is changed due to the variation of depth perception. In this viewing condition, depth perception is enhanced by binocular depth cues (e.g. binocular disparity), disparity information is exploited by brain to retrieve the 3D layout of the environment and lead to a stereoscopic perception of depth (Neri, Bridge, & Heeger, 2004; Howard & Rogers, 1995; Wheatstone, 1838). It has been shown recently that observers' fixations are biased not only towards the center area on the screen but also towards certain

depth planes in the scene (Jansen, Onat, & König, 2009; Wang, Le Callet, Ricordel, & Tourancheau, 2011; Ramasamy, House, Duchowski, & Daugherty, 2009). It is thus reasonable to suppose the existence of a so-called “depth-bias”.

In the area of developing computational models of visual attention in stereoscopic visualization, a set of models containing some hypotheses of depth-bias have been proposed. Several of them consist of a similar architecture: saliency is computed by using 2D visual features and is then weighted according to depth information. Most of these studies assumed that areas or objects close to the viewer were more salient than distant ones (Maki, Nordlund, & Eklundh, 1996, 2000; Zhang, Jiang, Yu, & Chen, 2010; Chamaret, Godeffroy, Lopez, & Le Meur, 2010). Verifying the existence of a depth-bias and quantifying it would be beneficial for the development of these 3D visual attention computational models.

However, the studies exploring this depth-bias are still limited. Jansen et al. (Jansen et al., 2009) investigated the viewing behavior in the observation of 2D and 3D still image. They conducted a free-viewing task on 2D and 3D version of the same set of images with natural content, pink noise or white noise. They found that viewer fixated closer locations earlier than more distant locations in both 3D images and 2D images. This result is also consistent with the works of Wang et al. (Wang, Le Callet, et al., 2011), who found that the closest object in a scene always attracted most fixations. On the contrary, Ramasamy et al. (Ramasamy et al., 2009) showed that observers’ gaze points were more concentrated at the far end (in terms of depth) when viewing a scene containing long deep hallway. The inconsistency between the conclusions of Jansen et al. and those of Ramasamy et al. might be due to the stimuli they used in their experiments. The use of images with natural content brought in many visual features other than depth (e.g. color, intensity contrast, orientation, center-bias). These features might affect the distribution of observers’ visual attention in both a bottom-up way and top-down way (Wolfe & Horowitz, 2004; Itti, Koch, & Niebur, 1998; Le Meur, Le Callet, Barba, & Thoreau, 2006; Bruce & Tsotsos, 2009; Wang, Chandler, & Callet, 2010). Therefore, it is important to get rid of the influence of 2D visual features on visual attention in order to investigate only the effect of the depth-bias. Using synthetic stimuli which are properly designed and controlled may be helpful to avoid any other side effects beside the bias under study as opposed to less controlled natural stimuli.

Studies of investigating the depth-bias on planar stereoscopic display by using synthetic stimuli are still relatively lacking. In the present study, we conducted

a binocular eye-tracking experiment by showing synthetic stimuli on a state-of-the-art stereoscopic display. Our results showed that the closer objects in a scene could attract more fixations, especially at the very beginning of observation. The number of fixations varied as a function of objects’ depth. This kind of distribution of fixation was time dependent.

Methods

We conducted a binocular eye-tracking experiment by showing synthetic stimuli on a stereoscopic display. Observers were required to do a free-viewing task. Gaze positions of both eyes were recorded, and both the location and the depth of fixations were computed. Stimuli presented during this experiment were designed in such a way that depth would affect eye movements independently from other visual features.

Participants

Twenty-seven subjects participated in this experiment (12 males and 15 females). The subjects ranged in age from 18 to 44 years. The mean age of the subjects was 22.8 years old. All of them were naive to the purpose of the experiment, and were compensated for the participation of the experiment. All the subjects had either normal or corrected-to-normal visual acuity. The vision (corrected if necessary) of each observer was checked, thanks to three normalized tests:

- Monoyer chart test was performed to check the acuity. Only the subjects who got a result higher than 9/10 took part in the experiment.
- Ishihara test was performed to check color vision. Only the subjects without any color troubles took part in the experiment.
- Randot stereo test was performed to check the 3D acuity. Only the subjects who got a result higher than 7/10 took part in the experiment.

Viewing conditions

Stimuli were displayed on a 26-inch (552×323 mm) Panasonic BT-3DL2550 stereoscopic LCD screen. Stereoscopy was achieved thanks to a pair of passive polarized glasses. The screen had a resolution of 1920×1200 pixels, and the refresh rate was 60 Hz. The maximum luminance of the display was 180 cd/m^2 , which yielded a maximum luminance of about 60 cd/m^2 when watched through glasses. The environment luminance was adjusted according to each observer, in order to let the pupil have an appropriate size for eye-tracking. SMI RED 500 remote eye-tracker was used to record the eye movements. The accuracy of this eye-tracker is 0.4 degree.

The viewing distance has been set to 93 cm. In this condition, each pixel subtends about 62 arcsec and the whole screen subtends 33.06×18.92 visual degrees in the observers' field of view. All the objects were displayed in an area within 10.32×5.91 degrees. A chin-rest was used to stabilize observer's head, and the observers were instructed to "view anywhere on the screen as they want".

All 118 scenes were presented in random order. Each scene was presented for 3 seconds. After each scene, a point located in the center of the screen and with no disparity was presented for 500 ms. A nine-point calibration of the eye-tracker was performed at the beginning of the experiment, and repeated every twenty scenes. The quality of the calibration was verified by the experimenter on another monitor. Participants could allow themselves a rest before each calibration.

Stimuli

The experiment consisted in the presentation of stereoscopic scenes in which some identical objects were displayed at different depth planes.

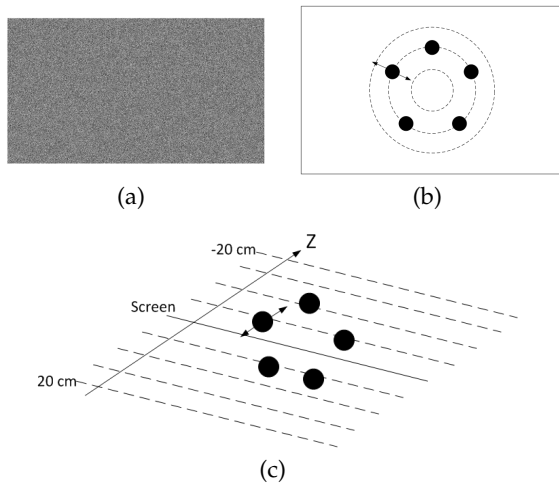


Figure 1. Composition of the stimuli. (a) The background of the stimuli. Only white-noise is contained in the background, which is positioned behind all the objects at -20 cm. (b) Positions of the objects' projections on the screen plane. Projected objects laid uniformly on a circle which centers at the screen center. (c) Allocation of objects in depth, potential depth values range from -20 cm to 20 cm by a step of 5 cm.

The background was a flat image consisting of white noise as shown in figure 1(a), which was placed at a depth value of -20 cm (20 cm beyond the screen plane). In each scene, the objects consisted of a set of black disks of the same diameter S . They were displayed at different depth values randomly chosen among $\{-20, -15, -10, -5, 0, 5, 10, 15, 20\}$ cm. Though the objects were

placed at different depths (figure 1(c)), the positions of the objects' projection on the screen plane uniformly laid on a circle centered on the screen center (figure 1(b)). Thus, it can be assumed that no "center-bias" was introduced in the observation.

For the stereo viewing, the perceived depth was achieved by horizontally shifting the object towards different directions in the left view and the right view to simulate the binocular disparity. The relationship between disparity (in pixel) and the perceived depth (in cm) was modeled by Equation 1:

$$D = V / (1 + \frac{I \cdot W}{P \cdot R_x}) \quad (1)$$

where D represents the perceived depth, V represents the viewing distance between observer and screen plane, I represents the interocular distance, W and R_x represents, respectively, the width (in cm) and the horizontal resolution of the screen, P is the disparity in pixels. Note that a positive disparity value ($P > 0$) indicates a crossed disparity, and a negative disparity value ($P < 0$) indicates an uncrossed disparity. The objects at positive depth value are between the viewer and the display, while the objects at negative depth values are beyond the display. A depth value of 0 corresponds to the screen plane.

The depth range (from -20 cm to 20 cm) was chosen in order to match to the comfortable viewing zone (Chen, Fournier, Barkowsky, & Le Callet, 2010), in the particular viewing conditions of this experiment. Therefore, it could be assumed that the conflict between accommodation and vergence in our experiment would not cause unacceptable level of visual discomfort or visual fatigue (Hoffman, Girshick, Akeley, & Banks, 2008).

To generate different stimuli, three parameters were independently varying from one scene to another:

1. The number of objects, $N \in \{5, 6, 7, 8, 9\}$.
2. The radius R of the circle on which the objects were projected on the screen plane, $R \in \{200, 250, 300\}$ pixels.
3. The size of the objects, which is represented by the diameter of the disk S . There were ten candidate values of S varying from 48 pixels to 168 pixels by a step of 12 pixels. Given a combination of N and R , S was selected from the range of $[\frac{\pi R}{N\sqrt{2}}, \frac{2\pi R}{N\sqrt{2}}]$, which was used to avoid any overlap between objects.

Derived from the combinations of this set of parameters, 118 scenes were presented to each observer. We had 30 five-object stimuli, 26 six-object stimuli, 23 seven-object stimuli, 21 eight-object stimuli, and 18 nine-object stimuli. Figure 2 gives some examples

of these different scenes. Note that the set of three independent parameters enabled the potential studies of the impact of different factors on depth-bias. However, in the scope of this paper, we particularly focus on the impact of objects number on depth-bias. We thus separated all the stimuli into five groups only depending on the number of objects (regardless the other two parameters).

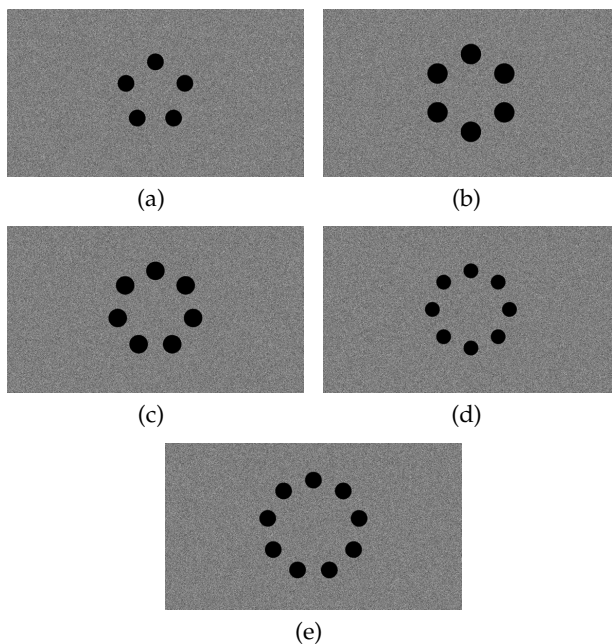


Figure 2. Examples of the five stimulus types (in terms of the number of objects contained) used in the eye-tracking experiment.

There were several advantages of using this type of synthetic stereoscopic stimuli for the investigation of depth-bias:

- Firstly, it allowed a precise control of the influence of 2D visual features on visual attention. Even in 3D viewing, human's eye movements are still affected by many bottom-up 2D visual features of the stimuli, such as color, intensity, object's size, and the center-bias. These factors could contaminate our evaluation of the influence of depth on visual attention. In our experiment, all the objects were with a constant shape, a constant size, and were positioned at a constant distance to the center of the screen. This set up let the stimuli get rid of as many bottom-up visual attention features as possible. The white noise background and the simple allocation of the objects allowed to avoid, as much as possible, the potential influence of top-down mechanisms of visual attention.

- Secondly, it allowed a precise control of the influence of depth cues on depth perception. Disparity was the only depth cue we took advantage of in

this experiment. The reason of choosing binocular disparity was that the relationship between this depth cue and the perceived depth could be well modeled (see Equation 1). While for some other (monocular) depth cues, such as blur, perspective, occlusion, (Wang, Barkowsky, Ricordel, Le Callet, et al., 2011), the influence on perceived depth was more difficult to be quantitatively measured.

- Thirdly, the white noise background and the simple allocation of objects limited the complexity of the scenes presented to the observers to a low level, which made a shorter observation duration feasible. The viewing time in our experiment was short (3 seconds for each trail). Nevertheless, it was still long enough for participants to subconsciously position their fixations on objects and explore the scene as they want. Hence, using these simple stimuli allowed experimenters to collect more data, as well as to learn the evolution of depth-bias over time.

Post processing of eye tracking data

The first step of processing was to identify the fixations and filter out the saccades. The recorded eye movement data were processed by the event detection software "BeGazed" provided by SMI. This software selected saccades as primary events using a velocity-based algorithm (Salvucci & Goldberg, 2000). Fixations (and blinks) were computed and derived from the primary saccade events:

1. The velocities of all the recorded gaze points were first calculated. The peaks were then detected from all these velocities. Note that a "peak" was defined as the peak values of velocity above the Peak Threshold (i.e. 40 degree/s in our experiment). Given the stream of velocity values, for each peak, we searched to the left for the first velocity which was lower than the fixation velocity threshold, in order to detect the start of a saccade-like event. Similarly, we searched also to the right for the first velocity lower than the fixation velocity threshold, in order to detect the end of the saccade-like event.

2. We assumed the saccade-like event a real saccade, if (1) the distance between start and end exceeded the Minimum Saccade Duration (22 ms), and (2) the single peak value lied in the range of 20% to 80% of the distance between start and end.

3. Finally, a fixation event was created between the newly and the previously created blink or saccade. All fixations below the minimum fixation duration (50 ms) were rejected.

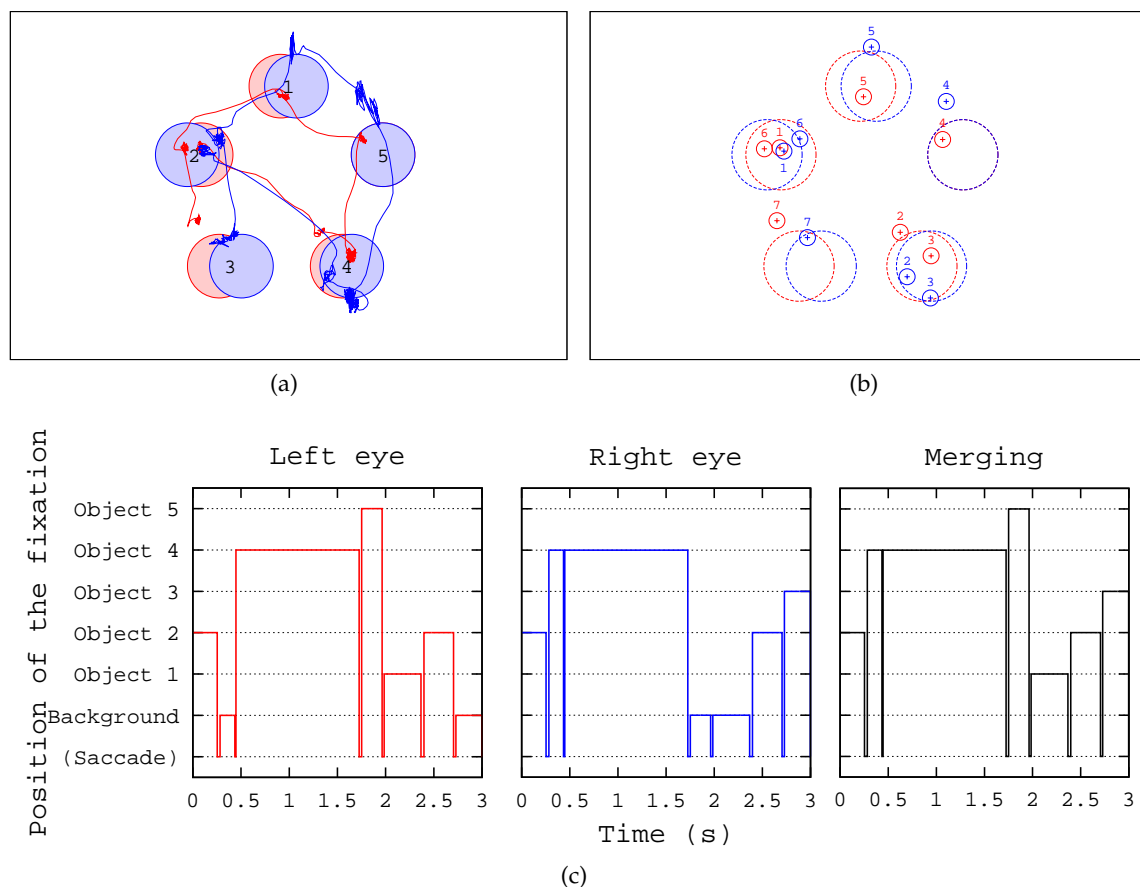


Figure 3. Illustration of the process of relating a fixation to an object. In (a), the curves illustrate the path of each eye, and disks illustrate the objects in each view. Red color represents the left eye's path and the left view stimuli, while blue color represents the right eye's path and the right view stimuli. The numbers inside each object indicate objects' ID. (b) shows the fixations of each eye (same color as introduced previously) along with the area of detection for each object and each view. The detection areas are 10% larger than the size of the objects. The numbers indicate the temporal order of the fixations. (c) represents the position of each fixation as a function of time. Results are given for each eye, and for the merging of both eyes.

The second step was to determine the spatial position of each fixation in order to relate it to the objects present in the scene. We found that directly computing the depth of a fixation based on the left and the right fixation's coordinates on the screen plane was difficult, due to the insufficient accuracy of the eye-tracker (see Figure 3(a)). Therefore, we adopted an indirect method to compute each fixation's depth. The computation was done independently for both eyes. Left eye positions were matched with left eye stimuli, and right eye positions were matched with right eye stimuli. It was then checked if a fixation was located on one of the objects or not. For each eye, a fixation was considered to be located on an object if it was positioned inside the object or in a surrounding area 10% larger than the object (to compensate for potential inaccuracy of the eye-tracker). Otherwise, the fixation was considered to be located on the background. Figure 3 illustrates the processus.

Both eyes' fixations were then merged by the following rule: a given object was considered as being fixated if at least one eye's fixation was inside this object (Figure 3(c)). Because each object's depth was known, the depth of a fixation could be deducted from its position. Note that only the fixations located on an object were considered in the following analysis.

Results

Fixation distribution in depth

The numbers of fixations located on each object were calculated for each observer and each scene. The result was then transformed into a frequency distribution: for each observer, we divided the number of fixations on each object by the total number of fixations. We considered the uniform probability

distribution $P_r = 1/N$ as a reference, based on the assumption that each object would attract the same amount of fixations if there was no depth-bias on the distribution of fixations (N represents the number of objects contained in a scene). This process was done repeatedly for all the five types of stimuli which contained different numbers of objects in the scene.

Figure 4 shows how the fixations are distributed on objects located at different depth planes in a scene. As we can see in the different plots, regardless the number of objects contained in the scene, the object closest to the observer always attracts most fixations (more than 20% of the total amount). The percentage of fixations then decreases as the depth order increases in the front half part of the scene. The curves generally follow a very similar shape in all five conditions. The frequency of the fixations on the objects located in the front range of the scene is significantly higher than predicted by the uniform probability distribution. The observation from these curves having similar trend supports the existence of depth bias.

Table 1
Results of the ANOVA performed on the fixation distributions presented in Figure 4.

Num. of objects	ANOVA result
5	$F(4,130) = 11.73, p < 0.05$
6	$F(5,156) = 12.42, p < 0.05$
7	$F(6,182) = 13.22, p < 0.05$
8	$F(7,208) = 13.8, p < 0.05$
9	$F(8,234) = 11.85, p < 0.05$

A one-way ANOVA has been performed to check the statistical significance of the values for each type of the scene. The results (presented in Table 1) confirm that there exists a significant effect of fixation's depth order on fixation distribution. A post hoc paired t-test with Bonferroni correction has been then performed to check the significant difference between each pair of ordinal fixations in depth. For all the conditions, the percentage of fixations on the first object is significantly higher than the others, while the fixation percentage from the third to N^{th} ordinal objects are not significantly different from each other.

Variation of fixation's depth as the function of fixation's temporal order

The curves in Figure 5 show how the first fixation of all observers is distributed on the objects in each type of stimuli. These curves were computed by the same way as introduced in the previous section,

except that only the first fixation of each observation were considered. These curves indicate the degree of depth-bias in a short viewing duration at the very beginning of observation. If we compare each curve in Figure 5 to the corresponding one in Figure 4, we find that the first fixations are more likely located on the closest object in a scene. These distributions of the first fixations demonstrate that the first fixation on each stimuli is more often located on the closest object than the following fixations.

To further evaluate the temporal evolution of the fixations' average depth, we investigated how the average depth of fixations varied as the temporal order of fixation increased. The relative depth position of each fixation in the scene's depth range was first computed by equation 2:

$$D_{r_i} = (D_i - D_{min}) / (D_{max} - D_{min}) \quad (2)$$

where D_i is the absolute depth of the i^{th} fixation, D_{min} and D_{max} are the minimum and maximum absolute depth of objects in a scene, respectively. Relative depth of the first seven fixations that are located on objects are computed and plotted in Figure 6.

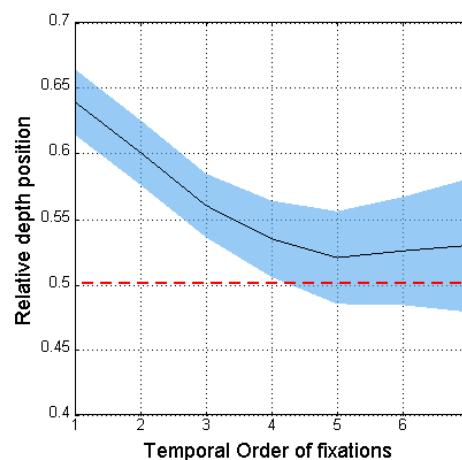


Figure 6. Relative depth position of fixations as a function of temporal order. The red dash line indicates the average value of relative depth of all objects displayed in the experiment. The light blue region represents the 95% confident interval.

An initial front response upon a new scene is observed in every participant and revealed in Figure 6. In this figure, the red dash line is plotted to indicate the average value of relative depth of all objects displayed in the experiment. If there was no depth-bias, the observers would explore the scene uniformly in depth during the observation, and each object in the scene should have the same probability to be fixated. That means the average depth value should vary little throughout the fixation sequence, and stay around the

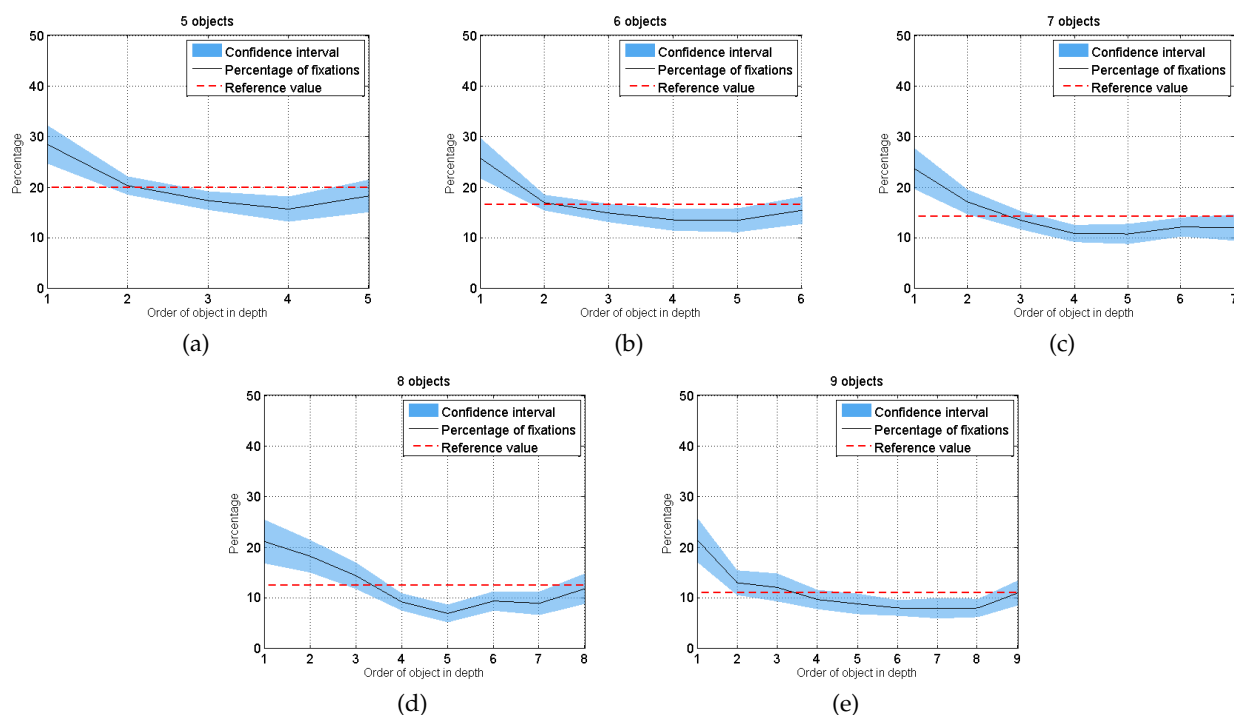


Figure 4. Fixation distribution (all the fixations were considered for all the conditions) as a function of the order of object's depth for the scenes containing different number of objects ($N \in 5, 6, 7, 8, 9$). X axis is the order of objects; sub-figure (a) to (e) represents the group of scenes that contain 5 to 9 objects, respectively. Y axis represents the percentage of number of fixations. The blue area represents the 95% confidence interval. The dash line represents the uniform probability distribution ($1/N$).

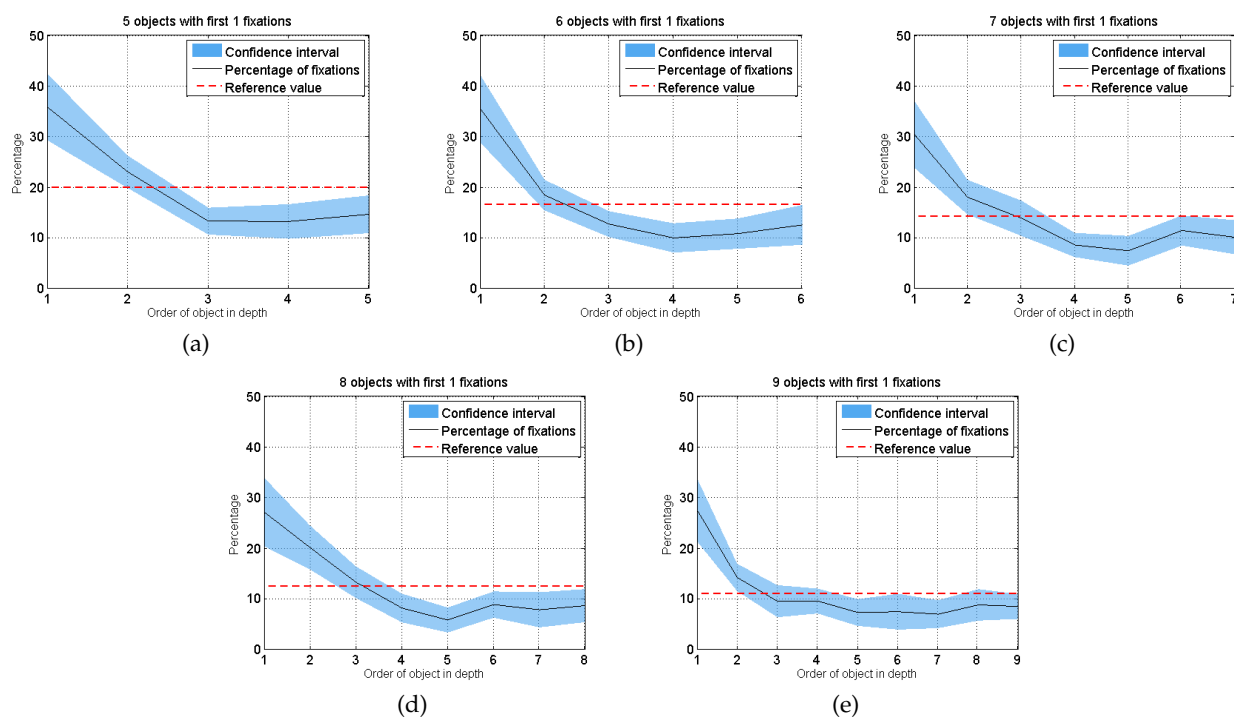


Figure 5. Fixation distribution (only the first fixation was considered) as a function of the order of object's depth for the scenes containing different numbers of objects ($N \in 5, 6, 7, 8, 9$). X axis is the order of objects; sub-figure (a) to (e) represent the group of scenes that contain 5 to 9 objects, respectively. Y axis represents the percentage of number of fixations. The blue area represents the 95% confidence interval. The dash line represents the uniform probability distribution ($1/N$).

value of 0.5. However, a clear decrease of the average depth can be observed in figure 6. The first fixations are also found to be more often located on the objects close to the viewer. A one-way ANOVA has been performed to check the significant difference among the relative depth of each temporally ordinal fixation. The result shows an effect of fixation sequence on depth ($F(6,860) = 7.94, p < 0.01$). A post hoc paired t-test with Bonferroni correction shows that the depth of the first and second fixation is significantly higher than the following fixations ($p < 0.01$).

This curve of fixations' average depth as a function of fixations' temporal order shows a viewing strategy in which observers tend to explore a scene from the closest objects. The average depth values of all fixations are higher than the average depth of objects, which means that observers pay more attention to the objects in the front part of the scene than the objects in the back part. All these observations support the existence of depth-bias.

Time dependence of fixation distribution in depth

The analyses in the previous sections reveal a variation of fixations' average depth according to the temporal ordinal of fixations. This variation implies that the level of depth-bias may be time dependent. In order to verify the time-dependence, we uniformly separated the 3-second observation time into six slices, then the fixation distribution as a function of depth (as introduced in Section *Fixation distribution in depth*) was computed for each slice of time. This processing was done repeatedly for all the five types of stimuli. The results are illustrated in Figure 7.

In all the five types of scene, a clear depth-bias is found within the first 1000 ms observation time. As the observation time increases, the number of fixations located on the closest object becomes smaller. The distribution of fixations on all the objects in a scene becomes more uniform. This tendency holds for all the five types of stimuli regardless the numbers of objects contained in the scenes. However, even if it is clear that the depth bias occurs at the beginning of the presentation, it is still hard to draw a conclusion of the time-dependence of depth-bias. Since once an object has been looked at, it is less likely to be looked at again due to the inhibition of return (Klein, 2000).

Discussion

The main goal of the present study is to determine if there exists a so-called "depth-bias" in the viewing of 3D content on planar stereoscopic display. We examined how the depth order and the relative depth of the objects influenced observers' viewing behavior.

In terms of depth order, experimental results clearly showed that observers paid more attention to the objects close to them than to the other objects. This phenomenon could be caused by a viewing strategy in which people prefer to explore a scene from the objects with least distance, since this kind of objects might mean some potential dangers in nature (Bowler, 1989). We found that this depth-bias was obvious at the very beginning of observation (i.e. the first fixation of each observation). The initial bias and the short reaction time implied that depth-bias might be the result of a bottom-up mechanism. Results also showed that the preference of looking at the closest objects decreased as the observation time increased. After one second of observation, fixations were distributed almost equally on all the objects in the scene regardless of the depth order of these objects. However, the sparse nature of stimuli makes it hard to draw conclusions on the time-dependence of depth-bias. In the experiment, once the closest object had been looked at, other objects would be looked at and it appeared that they were selected in no particular depth order. This variation of fixations' average depth could be caused by the inhibition of return (Klein, 2000).

In the present study, the synthetic stimuli used were designed to get rid of (as much as possible) the effect of the visual features other than depth. However, this is not the only approach. Despite the advantages of using the proposed synthetic stimuli, there exists another approach to measure the effect of depth: using stimuli taking into account also the other sources of visual information. This alternative approach can demonstrate the contribution of depth above and beyond that of other factors. The present study shows that the depth-bias seems to be a bottom-up process. Therefore, this alternative approach might help to verify if the effect of depth-bias is still significant in presence of top-down information (i.e. the semantics of picture).

In terms of applications, studying the depth-bias can be beneficial to the development of computational models of 3D visual attention for both still images and videos. In the area of computationally modeling of 3D visual attention, the depth-bias is usually considered to be linear to the disparity and time-independent (Maki et al., 1996, 2000; Zhang et al., 2010; Chamaret et al., 2010). Our results about the fixation distribution in depth and the time-dependence of the depth-bias might be used to improve these computational models in predicting salient areas of a 3D scene.

In the literature, studies have shown that stereoscopic vision relies mainly on relative depth difference between objects rather than on their absolute distance in depth from where the eyes fixate (Neri et al., 2004). In the present study, the influence of depth on the

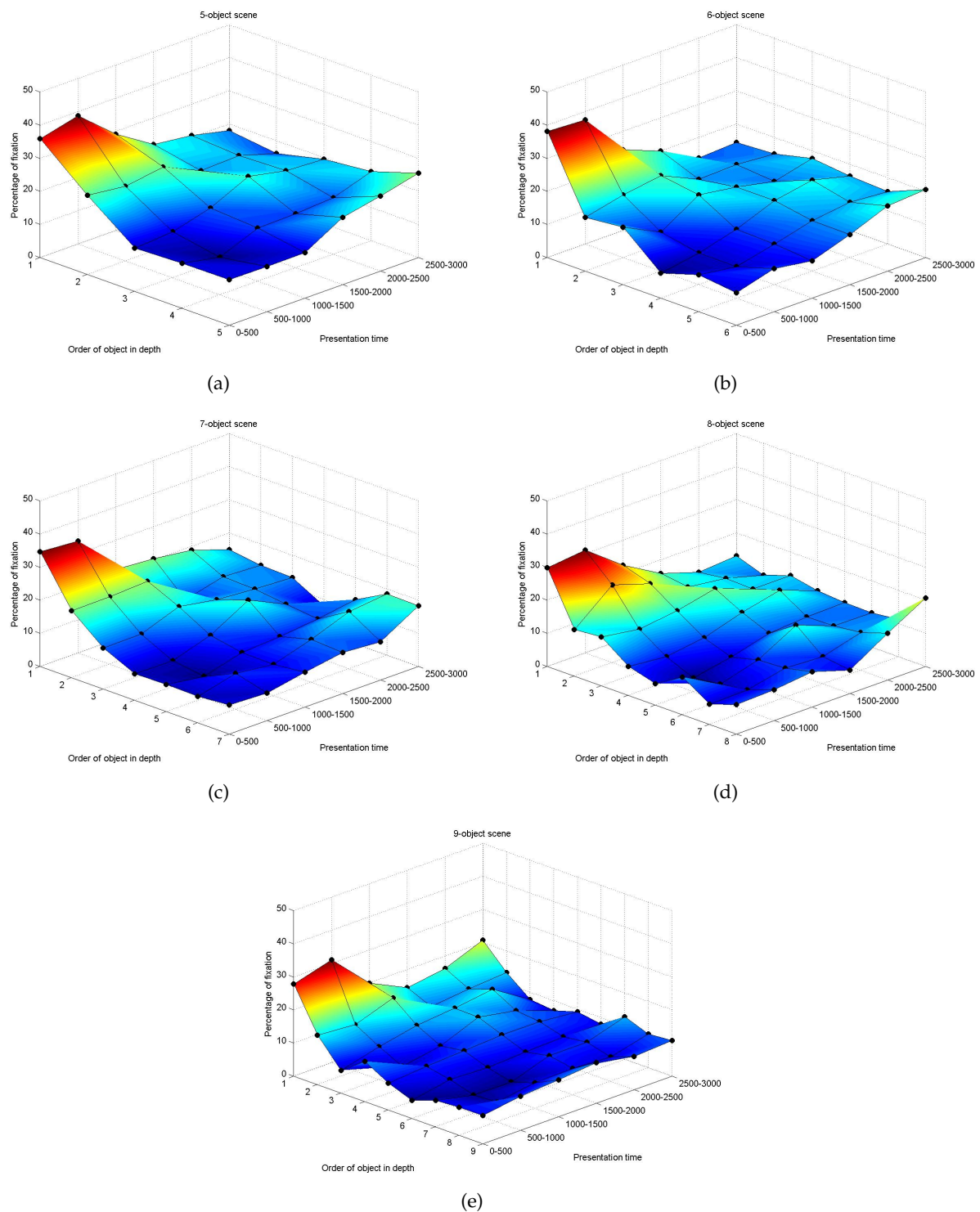


Figure 7. Variation of fixation distribution as a function of presentation time. The five sub figures represent five type of scenes which contain different numbers of objects.

distribution of visual attention was examined based on relative depth. However, the study of the influence of absolute depth information is not included in the scope of this paper. Absolute depth, which is linear

to the binocular disparity provided by stereoscopic displays, can be another important factor affecting the depth-bias, since it has been demonstrated the existence of disparity-selective neurons in the primary

visual cortex (V1) (Neri, Parker, & Blakemore, 1999; Barlow, Blakemore, & Pettigrew, 1967; Nikara, Bishop, & Pettigrew, 1968; Poggio, Fischer, et al., 1977). Moreover, we also know that the conflict between accommodation and vergence, which is caused by binocular disparity, affects the viewing behavior of stereoscopic content. The existence of areas with disparity larger than a threshold can cause problems of visual fatigue and visual discomfort. Obviously, this type of area will not attract too much attention, even if it is salient in terms of visual feature. All these evidences testify to the influence of the absolute depth (i.e. disparity) on visual attention in the viewing of stereoscopic 3D stimuli.

Conclusion

In the present study, we conducted an eye-tracking experiment using state-of-the-art stereoscopic display and eye-tracker. A large number of synthetic stimuli were designed for the experiment in order to get rid of the effect of 2D visual features, and let the visual attention of observers be influenced by only depth information. Results demonstrate the existence of the depth-bias in the viewing of 3D content on a planar stereoscopic screen.

Acknowledgments

This work was supported by the French ANR-PERSEE project (project reference: no ANR-09-BLAN-0170).

References

- Barlow, H., Blakemore, C., & Pettigrew, J. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, 193(2), 327-342.
- Bowler, P. (1989). *Evolution: the history of an idea*. Berkeley: University of California Press.
- Bruce, N., & Tsotsos, J. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, 9(3), 1-24.
- Chamaret, C., Godeffroy, S., Lopez, P., & Le Meur, O. (2010). Adaptive 3d rendering based on region-of-interest. *Proceedings of SPIE- The International Society for Optical Engineering*, 7524, 75240V.
- Chen, W., Fournier, J., Barkowsky, M., & Le Callet, P. (2010). New requirements of subjective video quality assessment methodologies for 3dtv. *Proceedings of the Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, F1b, 1-6.
- Hoffman, D., Girshick, A., Akeley, K., & Banks, M. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3), 33, 1-30.
- Howard, I., & Rogers, B. (1995). *Binocular vision and stereopsis*. Oxford University Press, USA.
- Huynh-Thu, Q., Barkowsky, M., & 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.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254-1259.
- Jansen, L., Onat, S., & König, P. (2009). Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*, 9(1), 29, 1-19.
- Klein, R. (2000). Inhibition of return. *Trends in cognitive sciences*, 4(4), 138-147.
- Le Meur, O., Le Callet, P., Barba, D., & Thoreau, D. (2006). A coherent computational approach to model bottom-up visual attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5), 802-817.
- Maki, A., Nordlund, P., & Eklundh, J. (1996). A computational model of depth-based attention. *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, 4, 734-739.
- Maki, A., Nordlund, P., & Eklundh, J. (2000). Attentional scene segmentation: integrating depth and motion. *Computer Vision and Image Understanding*, 78(3), 351-373.
- Neri, P., Bridge, H., & Heeger, D. (2004). Stereoscopic processing of absolute and relative disparity in human visual cortex. *Journal of Neurophysiology*, 92(3), 1880-1891.
- Neri, P., Parker, A., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401(6754), 695-698.
- Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research*, 6, 353-372.
- Poggio, G., Fischer, B., et al. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J Neurophysiol*, 40(6), 1392-1405.
- Ramasamy, C., House, D., Duchowski, A., & Daugherty, B. (2009). Using eye tracking to analyze stereoscopic filmmaking. *SIGGRAPH'09: Posters*, 28.
- Salvucci, D., & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the 2000 symposium on Eye tracking research & applications*, 71-78.
- Tatler, B. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 4, 1-17.
- Tseng, P., Carmi, R., Cameron, I., Munoz, D., & Itti, L. (2009). Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of Vision*, 9(7), 4, 1-16.
- Wang, J., Barkowsky, M., Ricordel, V., Le Callet, P., et al. (2011). Quantifying how the combination of blur and disparity affects the perceived depth. *Proceedings of SPIE- The International Society for Optical Engineering*, 7865, 78650K.
- Wang, J., Chandler, D. M., & Callet, P. L. (2010). Quantifying the relationship between visual salience and visual importance. *Proceedings of SPIE- The International Society for Optical Engineering*, 7527, 75270K.
- Wang, J., Le Callet, P., Ricordel, V., & Tourancheau, S. (2011). Quantifying depth bias in free viewing of still stereoscopic synthetic stimuli. In Vitu, F., Castet, E., & Goffart, L. (Eds.) *Abstracts of the 16th European Conference on Eye Move-*

- ments, Marseille, 21 - 25 August 2011. *Journal of Eye Movement Research*, 4(3), 92.
- Wheatstone, C. (1838). Contributions to the physiology of vision.–part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical transactions of the Royal Society of London*, 128, 371–394.
- Wolfe, J. (2000). Visual attention. *Seeing*, 2, 335–386.
- Wolfe, J., & Horowitz, T. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501.
- Zhang, Y., Jiang, G., Yu, M., & Chen, K. (2010). Stereoscopic visual attention model for 3d video. *Advances in Multimedia Modeling*, 314–324.