Contact: junle.wang@univ-nantes.fr

Quantifying depth bias in free viewing of still stereoscopic synthetic stimuli

Junle Wang[1], Patrick Le Callet[1], Vincent Ricordel[1], Sylvain Tourancheau[2]

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

2- Mid Sweden University, Sundsvall, Sweden

Introduction

Main purpose: Quantify a so-called “depth-bias” in the viewing of 3D content.

Rationale: In the studies of 2D visual attention, eye-tracking data shows a center-bias: fixations are biased towards the center of 2D still images. In the stereoscopic visual attention, depth is another feature having great influence on guiding eye movements. Relative little is known about the impact of depth. Several studies mentioned that people tend to look at the objects at certain depth planes.

Hypothesis: For stereoscopic condition, the existence of a "depth-bias". Studies proving or quantifying this depth-bias are still limited.

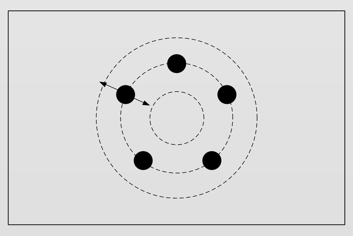
Methods: Investigating depth-bias by setting up a binocular eye-tracking experiment.

Experiment

Stimuli: Synthetic stereoscopic stimuli which contained a background and several black disks as salient objects to attract attention.

Depth generation: Horizontally shifting the objects to simulate the binocular disparity: the only depth cue we used.

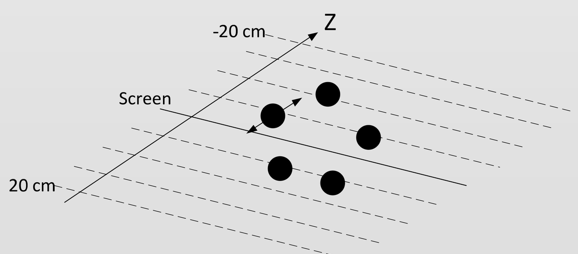
Shape



5 to 9 objects were contained in one scene. They were supposed to be the salient objects that attract attention. Their projections on the screen uniformly laid on a circle center at the screen center to avoid the effect of center-bias.

\+

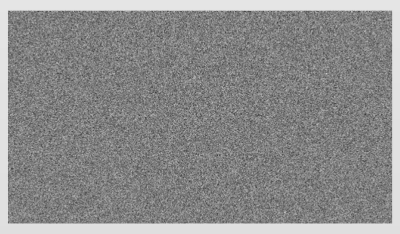
Depth



These objects were deliberately placed at different depth positions. The depth values were randomly chosen among (- 20, -15, -10, -5, 0, 5, 10, 15, 20) cm.

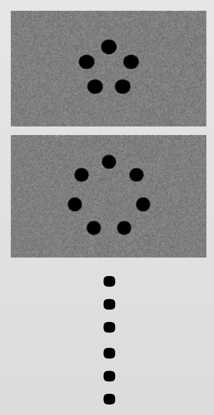
\+

Background



The background was placed at 100cm beyond the screen. It was a 2D image containing only white-noise.

=



118 conditions, 3 seconds for each

Constraints: Compared to natural content, synthesis stimuli is easier to control. Shape, position and depth (i.e. disparity) of each object can be well selected:

- Constant shape, size, color;
- Constant distance to the screen center;
- Random selection of depth.

➡ Get rid of as many other attention feature as possible: color, size, intensity, center-bias, depth contrast

Apparatus:

Display: Panasonic BT-3DL2550 LCD screen (1920 * 1200 at 60Hz).

Viewing distance: 93 cm, corresponding to a 33.06 * 18.92 degrees field of view. Stimuli were displayed in an area within 10.32 * 5.91 degrees.

Eye-tracker: SMI RED 500 remote eye-tracker. A chin-rest was also used. Data was pre-processed by Begaze accompanied by the eye-tracker to filter out the saccades.

Procedure:

1. 27 subjects naive to the purpose of the experiment. participated in the experiment. All had either normal or corrected-to-normal visual acuity.
2. Prior experiment test for visual acuity using a Snellen Chart, and for depth acuity using a Randot Stereo Test.
3. A nine-point calibration was done at the beginning and every 20 scenes. The quality of calibration was verified after each calibration.
4. A center point was showed at the screen center without any disparity for 0.5 second between two scenes.

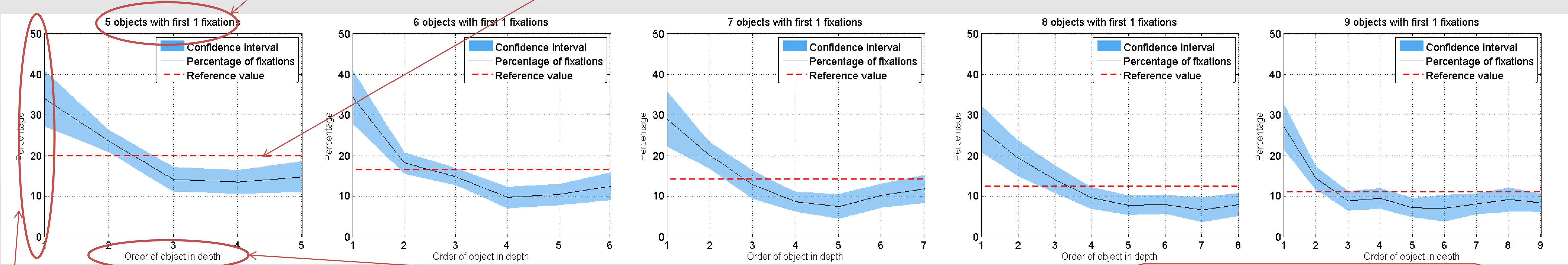
Data Analysis and Result

Fixation distribution in depth

Object’s order in depth, instead of the absolute depth value, was considered. 1 being the closest to the observer. Data was separated into different groups based on the objects number N in the scene.

Number of objects per condition — *Reference value: $P_r = 1/N$*

For only the 1st fixation

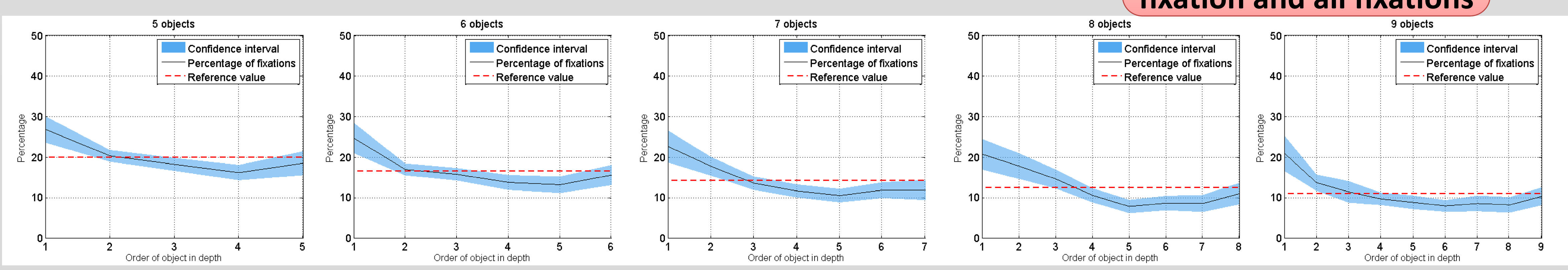


Objects’ orders in the scene

Percentage of fixations located on each object

Strong depth-bias is showed for first fixation and all fixations

Data form all the 3 seconds



Our data showed that convergence was achieved after 6 fixations for all the five types of condition.

1. A one-way ANOVA confirmed that there exists an effect of fixation's depth order on fixation distribution.
2. A post hoc paired t-test with Bonferroni correction showed that in a certain ordinal range (about the front 20%), the percentage is significance higher than the others.

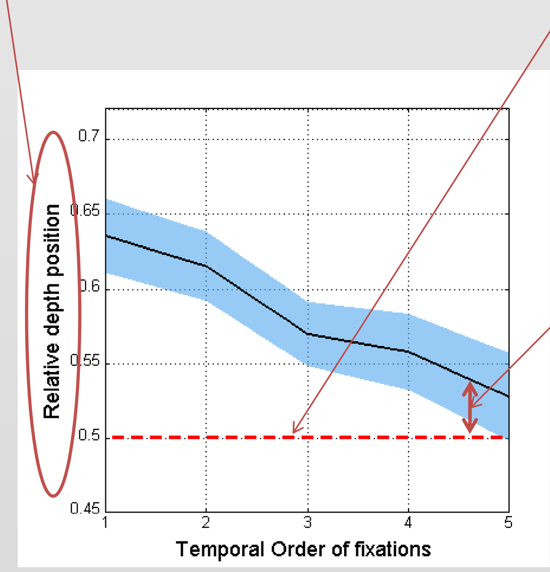
Num. of objects	ANOVA result
5	F(4, 130) = 11.73, p < 0.05
6	F(5, 156) = 12.42, p < 0.05
7	F(6,182) = 13.22, p < 0.05
8	F(7, 208) = 13.8, p < 0.05
9	F(8, 234) = 11.85, p < 0.05

Fixation distribution in time

We tried to find which part of the scene people would fixate first.

Average relative depth for each fixation.

Average relative for all objects: $D_r = 0.5005$



All the first 5 fixations located in the front half of the scene

Average relative depth D_r of each object was computed by:

$$D_{r(i)}=(D_i-D_{min})/(D_{max}-D_{min}),$$

where D_i is the absolute depth value of the i^{th} object. D_{max} and D_{min} is the depth of the front limit and the back limit of the scene respectively.

A one-way ANOVA showed an effect of fixation sequence on fixations’ depth. (F(4,670)=11.6, p<0.05)

Conclusion

1. Fixation distribution was biased to the objects locating in a certain front range (about 20%). The curves show that, the methods in the literature that considered the saliency distribution along the depth as a monotonic function or step function should be improved.
2. Strong depth-bias of the first fixtation implies that depth-bias might be a bottom-up process.
3. The convergent behavior of fixation distribution shows that the 3-second observation time was long enough. We can obtain the fixation distribution from only the first five fixations
4. The curve of average depth as a function of temporally ordinal fixations shows a viewing strategy that observers tended to explore the scene from the objects closed to them. This curve shows also the depth-bias.

Discussion and Future work

1. The similarity among the fixation distributions shows the possibility of a curve fitting for the development of a 3D visual attention model. We can analysis the scene by a 2D computational model, then combine the depth information as the complementary information.
2. In this work, we considered only the relative depth. However, it has been shown the existence of disparity-selective neurons in Area V4. And we also know that the conflict between accommodation and vergence affects the viewing behavior of stereoscopic content. These evidences show that the absolute depth value (i.e. disparity) might affect the fixation distribution.
3. We also found that, a part of observers started their observation from the objects close to the middle of the depth range, instead of the closest object. This phenomenon implies a different view strategy. A observer classification will be done to investigate this.

L'UNAM l'université nantes angers le mans — PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR
IRCCyN
UNIVERSITÉ DE NANTES
POLYTECH NANTES
Mittuniversitetet MID SWEDEN UNIVERSITY
AGENCE NATIONALE DE LA RECHERCHE ANR