Gaze Guidance in Immersive Environments

Steve Grogorick* Georgia Albuquerque† Marcus Magnor‡

Computer Graphics Lab, TU Braunschweig, Germany

Figure 1: Gaze distribution (heatmap) with (left) and without (right) gaze guidance. The red circle marks the target region.

Abstract

We investigate the efficiency of five different gaze guidance techniques for immersive environments, probing our peripheral vision's sensitivity to different stimuli embedded in complex, real-world panorama still images. We conducted extensive user studies for a commercially available headset as well as in a custom-built dome projection environment. The dome enables us to create true 360° visual immersion at high-resolution, akin to what may be expected of future-generation VR headsets. Evaluation with high-quality eye tracking shows that local luminance modulation as proposed by Bailey et al. is the most effective technique, eliciting saccades to the target region with up to 40 % success rate within the first second.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Virtual reality; Computing methodologies—Computer graphics—Graphics systems and interfaces—Perception;

1 Introduction

Virtual reality (VR) has been predicted as one of the most promising recent technologies. 360° images allow for a great immersive experience for arbitrary content, are easy to generate even with consumer-grade equipment like smart phones and easy to distribute via popular online platforms, e.g. YouTube and Facebook. However, consuming these media formats includes the risk of missing interesting parts of a captured scene, due to the unfamiliar wide range of the content — all around the viewer. The image frame that typically restrains the attention to the desired content is missing and has to be replaced by some other means. Different post-processing approaches to *unobtrusively guide* user attention to regions of interest in images or videos have been proposed in the last years, for traditional desktop applications [1, 2, 4], and recently for head-mounted displays [3, 5].

We propose an adaptation of different existing gaze guidance stimuli to wide field of view VR environments.We use eye tracking data to analyze their effectiveness in a perceptual study with 102 participants and 12 distinct 360° pictures in a VR headset and a dome projection system. In summary, we contribute the following:

- adaptation of five gaze guiding stimuli to immersive displays
- extensive user studies in wide–field of view (FOV) VR systems
- comparative evaluation of gaze guidance efficiency

*e-mail: grogorick@cg.cs.tu-bs.de
†e-mail: georgia@cg.cs.tu-bs.de
‡e-mail: magnor@cg.cs.tu-bs.de

2 Evaluated Gaze Guidance Techniques

To date, there is no gaze guidance technique that claims to be 100 % subtle, but we would like to avoid obvious approaches like the green arrow from Lin et al. [5] that would distract viewers from the actual task. This resulted in a selection of 5 stimuli which will be referred to as *ColorDot*, *SGD*, *ZoomRect*, *ZoomCircle* and *SpatialBlur*.

ColorDot refers to small red squares proposed by Dorr et al. [2]. Using the suggested duration of 120 ms exploits saccade masking for subtleness. The stimulus is shown repeatedly with a break of 2 s in between to allow free exploration of a virtual environment but still eventually achieve guidance, when the target comes into the viewers FOV. In accordance with previous work we add gaze–contingent deactivation, i.e. the stimulus is switched off as soon as a saccade towards the stimulus is detected [1]. To counteract acuity degradation of the human eye towards the periphery we apply eccentricity–based scaling [3]. Scaling is adjusted to retain a size of 1° visual angle at 12° eccentricity as in the original implementation. The final stimulus can be seen in Figure 2 (ColorDot).

Subtle Gaze Direction (SGD) refers to the dark–bright modulation proposed by Bailey et al. [1]. Except the introduction of eccentricity–based scaling, this stimulus is kept as in the original implementation, as depicted in Figure 2 (SGD).

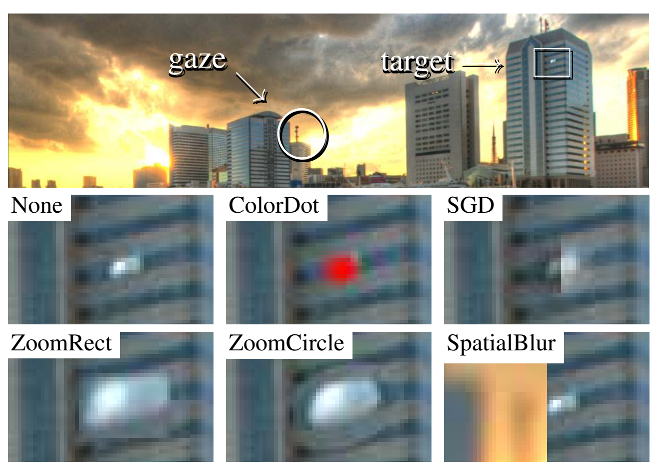

Figure 2: 360° picture with current gaze point (◯) marked. Six closeups below show the different stimuli at the target area (□). For illustration, SGD stimulus intensity was increased and shows both states (dark and bright) side by side. The left half of the SpatialBlur ed region at the current gaze point.

ZoomRect refers to the magnification stimulus proposed by Dorr et al. [2]. Similar to the ColorDot stimulus, it is extended with repeated presentation, gaze–contingent deactivation and eccentricity–based scaling. Original square shape and presentation time of 120 ms are kept and eccentricity–based scaling is adjusted to retain the suggested size of 2° visual angle at 12° eccentricity. The final stimulus can be seen in Figure 2 (ZoomRect).

ZoomCircle is a modified version of the ZoomRect stimulus. Similarly, it uses a magnification effect and is shown repeatedly. However, the stimulus is presented as a circular shape instead of a rectangular and the presentation time is increased to 500 ms. Slowing down the zoom effect shall strengthen the perception of optical expansion known from objects moving towards the observer. Finally, to prevent introduction of suddenly appearing sharp edges around the stimulus, the zoom factor to enlarge the content follows a radial gaussian fall–off from the center to the border.

Spatial Blur refers to the blur filter from Hata et al. [4] that applies spatial filtering to reduce details in non–target image regions. Our implementation follows the details of the original work.

Additionally, the shape of all stimuli is adjusted to prevent perception of perspective distortions which occur in peripheral regions for large FOVs, as suggested by Grogorick et al. [3].

3 System Setup

The dome is a 5 m tilted fulldome real–time video projection system, as in Figure 3 (a). It features six WQXGA projectors with image rates up to 120 Hz, each driven by a dedicated render node. A four camera *Smart Eye Pro*[1] setup enables real-time binocular eye tracking at 120 Hz .

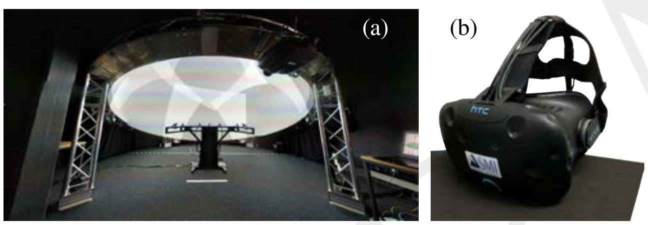

Figure 3: Dome projection system (a) and HTC Vive headset (b)

The VR headset is an *HTC Vive*, as in Figure 3 (b). It offers a resolution of 1080 × 1200 pixel per eye at 90 Hz. It is upgraded with an integrated binocular eye tracking solution from *SensoMotoric Instruments (SMI)*[2] running at 250 Hz.

4 Experiment Design

For each participant, each stimulus is shown once per VR system, with distinct still 360° pictures. All stimulus–picture presentations take 20 s while head and gaze movements are logged. A break of 10 s and a fixation marker before each trial ensures a common initial gaze point across all participants.

5 Results

A total of 102 participants, 31 female and 71 male, in the age of 19 to 41, took part in the experiment. One participant was noticeably effected by simulator sickness and thus excluded from the dataset.

Figure 1 shows exemplary gaze distributions for a sample image with and without gaze guidance. Without guidance attention is spread broadly all over the frontal hemisphere (right) while with active gaze guidance attention is notably concentrated at the target region (left). Sudden onset of a stimulus in the viewer's visual field, especially in the periphery, should trigger a saccade almost instantly. This implies that later saccades most probably do not occur due to the stimulus. Therefore, we inspect gaze data only from within the first second after the initial stimulus onset. Within this period SGD clearly outperforms the other stimuli in terms of successfully triggered target–directed saccades, as can be seen in Figure 4 (right).

In contrast, SGD reveals longer reaction times on average, which can also be seen in Figure 4 (left). However, this is most probably due to the longer presentation time, which results in saccades that were triggered later on.

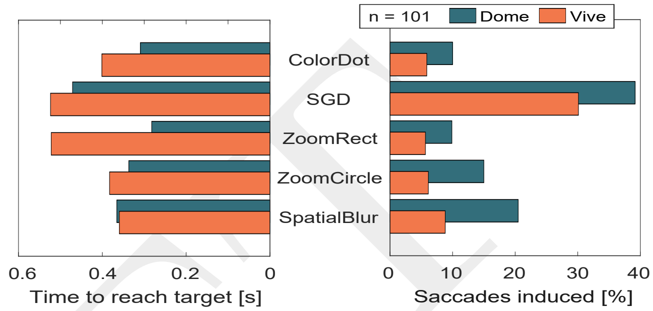

Figure 4: Reaction times and success rates of all stimuli within 1 s after stimulus onset, separated for both virtual environments.

6 Conclusion

We compare the effectiveness of several adapted subtle gaze guidance stimuli in a VR headset and — with regard to future–generation VR headsets — in a virtual dome environment. In particular, we analyze eye tracking data recorded in a psychophysics study with 102 participants and 12 distinct 360° images. Our results demonstrate that the adapted SGD stimulus outperforms the other stimuli in such wide–FOV VR environments, inducing target–directed saccades in up to 40 % of the trials within the first second.

In future work, subtleness of the different stimuli will be part of our next study. Also repetitive stimuli are to be investigated in more detail, especially with respect to priming effects between successive stimulus presentations. Furthermore, we plan to extend our study to investigate subtle gaze guidance in 360° video sequences.

Acknowledgments

The authors gratefully acknowledge funding by the German Science Foundation (DFG MA2555/15-1 “Immersive Digital Reality” and DFG INST 188/409-1 FUGG “ICG Dome”). Background images in Figure 1 and 2 CC-BY Igors Jefimovs and heiwa4126. We thank Linda Eckardt (wi2, TU Braunschweig) for recruiting participants.

References

[1] R. Bailey, A. McNamara, N. Sudarsanam, and C. Grimm. Subtle gaze direction. *ACM Transactions on Graphics (TOG)*, 28(4):100, 2009.

[2] M. Dorr, T. Martinetz, K. Gegenfurtner, and E. Barth. Guidance of eye movements on a gaze-contingent display. In U. J. Ilg, H. H. Bülthoff, and H. A. Mallot, eds., *Dynamic Perception Workshop of the GI Section “Computer Vision”*, pp. 89–94, 2004.

[3] S. Grogorick, M. Stengel, E. Eisemann, and M. Magnor. Subtle gaze guidance for immersive environments. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 4:1–4:7. ACM, 2017.

[4] H. Hata, H. Koike, and Y. Sato. Visual guidance with unnoticed blur effect. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 28–35. ACM, 2016.

[5] Y.-C. Lin, Y.-J. Chang, H.-N. Hu, H.-T. Cheng, C.-W. Huang, and M. Sun. Tell me where to look: Investigating ways for assisting focus in 360 video. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2535–2545. ACM, 2017.

[1] http://smarteye.se/wp-content/uploads/2016/10/Smart-Eye-Pro.pdf

[2] https://smivision.com/wp-content/uploads/2016/11/smi_prod_eyetracking_hmd_HTC_Vive.pdf