Visual Search in the (Un)Real World: How Head-Mounted Displays Affect Eye Movements, Head Movements and Target Detection

Tobit Kollenberg
Faculty of Technology
Bielefeld University
Germany

Alexander Neumann
Faculty of Technology
Bielefeld University
Germany

Dorothe Schneider
Faculty of Technology
Bielefeld University
Germany

Tessa-Karina Tews
Faculty of Technology
Bielefeld University
Germany

Thomas Hermann
SFB 673
Bielefeld University
Germany

Helge Ritter
SFB 673
Bielefeld University
Germany

Angelika Dierker
SFB 673
Bielefeld University
Germany

Hendrik Koesling*
SFB 673
Bielefeld University
Germany

Abstract

Head-mounted displays (HMDs) that use a see-through display method allow for superimposing computer-generated images upon a real-world view. Such devices, however, normally restrict the user's field of view. Furthermore, low display resolution and display curvature are suspected to make foveal as well as peripheral vision more difficult and may thus affect visual processing. In order to evaluate this assumption, we compared performance and eye-movement patterns in a visual search paradigm under different viewing conditions: participants either wore an HMD, had their field of view restricted by blinders or could avail themselves of an unrestricted field of view (normal viewing). From the head and eye-movement recordings we calculated the contribution of eye rotation to lateral shifts of attention. Results show that wearing an HMD leads to less eye rotation and requires more head movements than under blinders conditions and during normal viewing.

CR Categories: B.4.1 [Input/Output and Data Communications]: Input/Output Devices—Image Display H.1.2 [Models and Principles]: User/Machine Systems—Human Information Processing H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities;

Keywords: Field of view, restriction, head movements, eye movements, head-mounted display, augmented reality, visual search

1 Introduction

Recent advances in technology have led to significant improvements of head-mounted displays (HMDs) that use a video see-through method and allow for superimposing computer-generated images upon a real-world view. Improvements concern, for example, the size and robustness of such devices. In contrast, HMDs still present some restrictions such as low spatial resolution that users do not experience under normal viewing conditions [Azuma et al. 1997]. When, for example, HMDs are used for augmented reality in human-human or human-machine interaction, HMDs tend to reduce the users' field of view [Arthur 2000]. Also, the interaction partners cannot see each other's eyes. Gaze contact is thus not available as a means of communication any more (e.g. [Gibson and Pick 1963], [Kleinke 1986]). Equally important, partners can no longer determine the other person's focus of attention which often is essential for successful collaboration in shared environments (e.g., [Brennan et al. 2008], [Kaplan and Hafner 2006], [Velichkovsky 1995]).

*e-mail: ihkoesli@techfak.uni-bielefeld.de

As a consequence, according to Dierker et al. [2009], wearing an HMD can affect inexperienced HMD users to the extent of impairing their ability of performing everyday tasks, such as reaching for objects or shaking hands with others. While seeing only the HMD's projection of their surroundings, users apparently perceive objects to be closer than they actually are so that hand-eye coordination becomes difficult. Dierker et al. [2009] also noticed that users tend to alter their usual communication pattern with a partner to adapt to the new circumstances, especially concerning the frequency and strength of head gestures such as nods. While users seemed to reduce their head movements because of the heavy weight of a previous HMD system, this effect was much less observed with a newer more lightweight system.

The present study now aims at further investigating these conflicting results and at clarifying how the restrictions of the field of view (when wearing an HMD) affect eye and head movements (c.f., [Simonet and Bonnin 2003]). Furthermore, we have to take into account that the low resolution and curvature (causing pronounced peripheral blur) of such displays may hinder the peripheral vision in particular. This would affect visual processing and should show changes – in comparison to normal viewing – in eye-movement parameters and in the contribution of eye rotation to shifts of attention. Expectations are that more head movements with a higher amplitude and fewer eye movements with a lower amplitude occur when an HMD is worn than under normal viewing conditions (c.f. [Mertes 2009]). If HMDs' low resolution and peripheral blur effects indeed cause additional problems, we have to expect eye and head-movement amplitudes to differ less from normal viewing when restricting the field of view by "conventional methods" (for example by using blinders). If this hypotheses would hold true, gaze direction could be approximated with reasonable accuracy from head position measurements when an HMD is worn. At least for current HMD devices with their deficits in resolution and peripheral blur, this would eliminate the need to track the eyes separately to determine a user's focus of attention. This would save weight, make data processing easier and thus facilitate developing and evaluating new communication and interaction technologies with HMDs.

2 Experiment

We tested the hypotheses using a visual search paradigm (c.f., [Wolfe 1998]) while comparing target detection times and eye-movement patterns for three different viewing conditions: (a) participants either wore an HMD, had (b) their field of view restricted by blinders or (c) could avail themselves of an unrestricted field of view (normal viewing). From the head and eye-movement recordings we calculated the contribution of eye rotation to lateral shifts of attention, which we then compared between the different viewing conditions. In addition, we compared a number of other relevant oculomotor parameters, such as number of fixation and fixation duration, complemented by search times.

2.1 Method

Participants: Six students of Bielefeld University, 5 male and 1 female, participated in the experiment. Their average age was 24.3 years and all participants were fluent in German and had normal or corrected-to-normal vision.

Apparatus: Stimuli were projected onto a white display screen by a Sharp Notevision XG-C455W LCD projector. Stimulus projections measured 275 cm x 205 cm (width $\times$ height), subtending a visual angle of $69^o \times 54^o$. The stimulus' spatial resolution was set to 1280 $\times$ 800 pixels. Participants were seated approximately 200 cm from the screen. We used an EyeLink II eye-tracking system to record the participants' eye movements during the experiment at a sampling rate of 500 Hz. A calibration was performed before each trial to ensure accurate recordings. During the HMD trials, participants wore a customised version of an ARvision-3D video see-through head-mounted display manufactured by Trivisio. It contains two Point Grey Firefly MV cameras that deliver an uncompressed 640 x 480 pixels video stream. For each eye, the Trivisio system features one 800 x 600 pixels display. The HMD was attached to a Lenovo ThinkPad T61 that ran *LAFORGE* software, developed by [Mertes 2009], to display the video stream of one camera on both Trivisio displays (see Figure 1, left).

The blinders consisted of a dark frame that was attached to the EyeLink II and a smaller adjustable part that confines the view window (see Figure 1, right). The frame was bent on the sides to limit the field of view to the central view window. The resulting field of view was about equal to the one offered by the Trivisio HMD.

Stimuli and design: Stimuli consisted of a 8 $\times$ 6 grid of single-digit numbers (0 - 9). The target number was present only once in each stimulus display and was shown at the same grid position to all participants. Target numbers and positions varied between trials. All other numbers had random values and appeared in arbitrary frequency and position on the grid. Numbers were displayed in white ((R, G, B) = (255, 255, 255)) on a black background ((R, G, B) = (0, 0, 0)) and measured approximately 1.3^o $\times$ 2.0^o in width and height, respectively. Digits were spaced apart at approximately 8.0^o horizontally and 5.0^o vertically (see Figure 2).

Participants had to view stimuli under three different viewing conditions: normal viewing, blinders, HMD. In all three conditions, participants wore the eye tracker in order to allow for the recording of eye movements. In the normal viewing condition, the participants' field of view was not restricted. In the blinders and HMD viewing condition, blinders or HMD were additionally worn and restricted the field of view to approximately 25^o of visual angle (also see "*Apparatus*").

Procedure: The participants' task was to detect a specified target number within the number grid. Each participant solved the target number detection task for all three viewing conditions. Trials were blocked by viewing condition and participants had to accomplish 10 trials within each block. Each block started with one practice trial. The sequence of blocks was permuted between subjects so that all possible sequences of viewing conditions were accounted for. Trial order within blocks was randomised.

Prior to the experiment, we set up the eye-tracker and, depending on the viewing condition, blinders or HMD device and ran a multi-point system calibration. Before each trial, a single-point calibration was performed in order to compensate for possible drift of the

Figure 1: *Left: Participant wearing Trivisio HMD and eye tracker. Right: Participant wearing blinders and eye tracker.*

Figure 2: *Participant in front of display screen.*

eye-tracker headset. Each trial started with the display of the target number in text form, e.g., *"Two"*, on a blank black screen. When participants proceeded (response-button press, self-paced), a central fixation cross appeared for 500 ms, followed by the number grid stimulus. When participants had detected the target number, they pressed the response button to complete a trial. Figure 2 shows the experimental setting. After the experiment, participants completed a short questionnaire.

Data Analysis: To compute the contribution of eye movements to lateral shifts of attention, we used the "HREF" (head-referenced) data recording option of the EyeLink II System. Along with display-based coordinate recordings for every entry in the output file, the system writes an HREF pair of coordinates. These coordinate pairs define a point in a plane at a distance d_0 of 15000 arbitrary units from the eye (see Figure 3).

Head movements do not affect HREF coordinates and HREF coordinates do not relate to the gaze position on the stimulus display screen. Instead, the point of origin of HREF coordinates is arbitrary, so that the difference between two subsequent coordinate pairs measures their relative distance. The change of eye rotation ϕ in degrees of visual angle between two HREF points can be directly calculated as

$$\phi = c \cdot \arccos\left(\frac{d_0^2 + x_1x_2 + y_1y_2}{\sqrt{(d_0^2 + x_1x_1 + y_1y_1)(d_0^2 + x_2x_2 + y_2y_2)}}\right)$$

where (x_i, y_i) defines a coordinate pair and the constant c is set to 57.296 (for details see [SR 2007]). The calculated difference

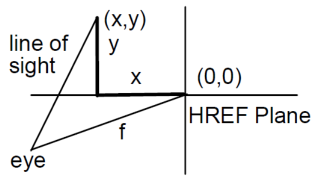


Figure 3: *Visualisation of HREF coordinates.*

between ϕ (taking into account only eye rotation) and the amplitude of the corresponding saccade measured on the display screen (eye rotation plus head rotation, also in degrees of visual angle) then indicates the amount of head movement involved in that particular shift of attention.

To compare eye rotations between the different viewing conditions where different average saccade amplitudes (eye rotation plus head rotation) had been recorded, eye rotation angles were normalised. Data from the normal viewing condition served as the standard so that mean eye rotations of all viewing conditions are reported in relation to the same overall saccade amplitude.

Statistical data analyses were computed using SPSS 16.0. We used the one-way ANOVA and general linear model for repeated measures to compare means between viewing conditions. The α-level was set to 0.05 and Bonferroni-adjusted when t-tests were computed in multiple pairwise post-hoc comparisons. Apart from eye rotation, we also analysed search time, number of fixations and fixation duration.

3 Results

Data showed that search times (ST) differed significantly between viewing conditions $(F(2;10) = 9.996; p = 0.004)$ and measured 2045 ms in the normal viewing condition, 2165 ms in the blinders condition and 3022 ms in the HMD condition. Pairwise post-hoc comparisons demonstrated that STs in the HMD condition were significantly longer than in the normal viewing $(T(5) = 4.461; p = 0.007)$ and blinders condition $(T(5) = 3.765; p = 0.013)$ while STs did not differ significantly between the normal viewing and blinders conditions (see Figure 4(a)).

The comparison of eye rotations (ER) resulted in a significant effect $(F(2;10) = 8.436; p = 0.007)$ for the viewing condition. ERs measured 15.32^o for normal viewing, 12.65^o for blinders and 11.46^o for HMD conditions. Post-hoc tests demonstrated, however, that only ERs in the HMD condition were significantly smaller than during normal viewing $(T(5) = -9.276; p < 0.001)$. ERs in the blinders condition were only smaller in tendency than those during normal viewing $(T(5) = -2.083; p < 0.092)$ while no difference existed in ERs between the HMD and blinders conditions (see Figure 4(b)).

The numbers of fixations (NF) within a trial also varied significantly between the viewing conditions $(F(2;10) = 5.631; p = 0.023)$, yielding a low NF value of 10.22 for normal viewing and higher values of 12.43 and 18.57 in the blinders and HMD conditions, respectively. Post-hoc tests again revealed a significant difference only between the normal viewing and HMD conditions $(T(5) = -3.202; p = 0.024)$. Mean NFs are charted in Figure 4(c).

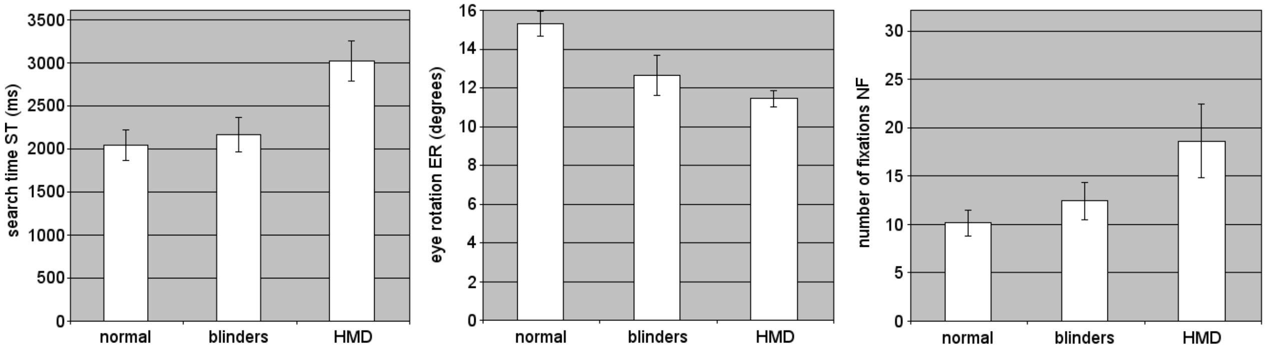
We could observe no significant difference between the viewing conditions with regard to mean fixation durations (FD). FDs remained almost unchanged in the normal viewing (158 ms), the blinders (160 ms) and the HMD condition (154 ms).

Most relevant findings from the analysis of the questionnaire data revealed that participants considered the eye tracker less restricting to their field of view than the HMD. Most participants stated that they were conscious about that they moved their head more when wearing the HMD or the blinders than during normal viewing. The questionnaires also revealed considerable individual differences between participants in evaluating the extent of the field of view restriction in the HMD and blinders conditions. Two participants experienced a stronger restriction of their field of view while wearing the blinders than the head-mounted display while four participants noted no difference between these two conditions. All participants agreed that the view through the HMD was slightly blurred.

4 Discussion and Conclusions

Results show that wearing an HMD leads to less eye rotation and consequently requires more head rotation to achieve the same lateral shift of gaze direction – and thus, shift of attention – than under normal viewing conditions. In tendency, eye rotation contributes less to shifts of attention when wearing blinders, too. This finding is not unexpected when the field of view is restricted and observers have to move their focus of attention to locations outside their current field of view. At first sight, since no significant differences exist in eye rotation between the HMD and blinders viewing conditions, wearing an HMD does not seem to affect visual processing differently than wearing blinders. Although not significantly so, eye rotations in the HMD viewing condition are somewhat smaller than in the blinders viewing condition. Furthermore, when taking into account search times, it takes significantly longer to successfully accomplish the search task under HMD viewing conditions than when wearing blinders. These two observations clearly demonstrate that equivalent restrictions of the field of view as achieved by HMD and blinders do not necessarily have the same effects on the performance in a specific task, for example, visual search. When participants view stimuli through a head-mounted see-through device, their search performance deteriorates significantly in comparison to when using blinders or indeed in comparison to normal viewing. In contrast, similar search times in blinders and normal viewing conditions indicate that search performance does not suffer significantly from the restricted field of view caused by the blinders. The blurred image that the HMD device provides in particular in the peripheral field of view and possibly also the distortion of the view due to the curvature of the display do thus seem to be likely causes for longer search times. This is confirmed by the subjective ratings and comments that users gave after the experiment who consistently reported to have noted display blur that affected their view.

When further taking into account the numbers of fixations that were recorded during the different viewing conditions, we obtain valuable hints towards details of visual processing during the visual search task. Significantly more fixations are being made in the HMD viewing condition than during normal viewing. In addition, the means of fixation numbers are notably higher – even though not significantly so – in the HMD viewing condition than in the blinders condition. These observations in conjunction with the findings discussed above may suggest particularly for the HMD condition that fixation points are spaced more closely than under normal viewing conditions when gaze shifts often involve both head and eye rotation. Furthermore, some “orientation process” in an attention area that has been newly attended to may require one or more (low-amplitude) corrective saccades. It appears to be likely that this is a consequence of the blurred periphery in HMD viewing conditions. The selection of targets for shifts of attention becomes less clear, even when target areas are relatively close. Interestingly, this effect does not recur in fixation durations. Rather than increase in the HMD viewing condition, they remain stable in all viewing condi-



(a) Mean search times (in ms).

(b) Mean eye rotations (in degrees of visual angle).

(c) Mean numbers of fixations.

Figure 4: *Results from the eye tracking analysis for the different viewing conditions (normal viewing, blinders, HMD), respectively.*

tions.

Nevertheless, the amplitude of eye rotation angles still reaches considerable magnitudes in the HMD viewing condition. Our findings do thus yield only limited support for an approach that predicts the focus of attention based on head orientation in practical HMD applications. Although eye rotation is reduced in comparison to normal viewing, head orientation does not serve as a reliable approximation of the focus of visual attention when wearing an HMD. When using an HMD device, for example in an augmented reality scenario, the focus of attention needs to be determined for superimposing additional information for a scene object that is currently at the centre of (visual) attention. Results from this study clearly suggest that the obligation remains to track the eyes separately from the head to accurately determine the user's focus of attention. Unfortunately, this means that HMD applications still need to accommodate an eye tracker, ideally integrated into the HMD device. Of course, this will make the entire device even more heavier and obtrusive. Furthermore, integration of eye movement, head movement and display data has to be accomplished which may add to computational costs.

In conclusion, the present study has shown that current head-mounted video see-through devices still present considerable restrictions to users, in particular with regard to the field of view that is available and accessible without problem. Findings also highlight the urge to improve presentation display technology so that the width and quality of the field of view better resemble normal viewing conditions. This would make the use of HMDs more intuitive and provide more natural viewing and interaction conditions. With advances in technology and miniaturisation, the integration of an eye tracker in an HMD may be a problem that we can solve in the near future. This should significantly increase the accuracy of attention-based augmented reality displays and thus improve user-friendliness and acceptance of such advanced interaction technology.

Acknowledgements

This research was supported by the German Research Foundation (DFG Sonderforschungsbereich 673 “Alignment in Communication”, Projects A5 and C5).

References

ARTHUR, K. 2000. *Effects of field of view on performance with head-mounted displays*. Dissertation, University of North Carolina.

AZUMA, R., ET AL. 1997. A survey of augmented reality. *Presence-Teleoperators and Virtual Environments 6*, 4, 355–385.

BRENNAN, S., CHENA, X., DICKINSONA, C., NEIDERA, M., AND ZELINSKY, G. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition 106*, 1465–1477.

DIERKER, A., BOVERMANN, T., HANHEIDE, M., HERMANN, T., AND SAGERER, G. 2009. A Multimodal Augmented Reality System for Alignment Research. *Proceedings 13th International Conference on Human-Computer Interaction*, 422–426.

GIBSON, J., AND PICK, A. 1963. Perception of another persons looking behavior. *American Journal of Psychology 76*, 386–394.

KAPLAN, F., AND HAFNER, V. 2006. The challenges of joint attention. *Interaction Studies 7*, 135–169.

KLEINKE, C. 1986. Gaze and eye contact: a research review. *Psychological Bulletin 100*, 78–100.

MERTES, C. 2009. *Multimodal augmented reality to enhance human communication*. Diploma thesis, Bielefeld University.

SIMONET, P., AND BONNIN, T. 2003. *Coordination Oeil-Tete chez le presbyte*. Ecole d'optomtrie, University of Montreal.

SR RESEACH LTD. 2007. *EyeLink II User Manual, Version 2.11*.

VELICHKOVSKY, B. 1995. Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition 3*, 199–222.

WOLFE, J. 1998. Visual search. In *Attention*, University College London Press, London, H. Paschler, Ed., 13–71.