第 27 回日本バーチャルリアリティ学会大会論文集（2022 年 9 月）

融合身体におけるユーザ間のビューマッチングに関する基礎研究

Basic study on view matching between users in virtual co-embodiment

Chatziantoniou Nikolaos[1), 雨宮智浩 [1,2), 小柳陽光 [2)，伊藤研一郎 [2)，青山一真 [2)，葛岡英明 [1)

Chatziantoniou Nikolaos, Tomohiro Amemiya, Akimi Oyanagi , Kenichiro Ito , Kazuma Aoyama, Hideaki Kuzuoka

1) 東京大学大学院情報理工学系研究科（〒 113-8656　東京都文京区本郷 7-3-1, {n.chatz,kuzuoka}@cyber.t.u-tokyo.ac.jp）

2) 東京大学バーチャルリアリティ教育研究センター（〒 113-8656　東京都文京区本郷 7-3-1, {amemiya,oyanagi,ito,aoyama}@vr.t.u-tokyo.ac.jp））

概要: In a virtual co-embodiment system between a trainer and a trainee, being able to visually redirect the focus and attention of the trainee to a specific point is imperative for the training process. In such systems viewpoints are altered dynamically and in quick succession, making a direct head movement sharing approach unsuitable, due to motion sickness. We propose a technique that blurs part of the follower's viewpoint to visually guide it towards and match the viewpoint of the leader in a subtle way without significant cognitive load. In this preliminary study we evaluate and compare our technique with three other visual techniques on usability, mental workload, user presence and time performance using static view targets. We find that the technique can be a subtle, quick and low workload visual guidance system with comparable levels of presence and sickness.

キーワード： Human-Centered Computing, Visual Guidance, Virtual Co-Embodiment

1. Introduction

Virtual Reality (VR) and the current Head-Mounted Displays (HMDs) allow for near natural Field of View (FOV). Cybersickness (CS) limits the way we can manipulate the user's visual field, necessitating a guided approach instead of arbitrary camera motions. Attention guidance methods aim to shift our focus towards desired Points of Interests (POI) in the immersive environment. Visual guidance has been proposed in numerous approaches for VR [1]. A major focus has been given towards general post-processing approaches that are not tied with the scene. Especially attractive are subtle methods which aim to guide users without distracting from the actual content or task [2].

1.1 Virtual co-embodiment

In virtual co-embodiment, a state in which a user and another entity (another user, robot or autonomous agent) are embodied in the same virtual avatar in VR while sharing control [3], several techniques limit the shared control to the body below the neck. Even though both users share a first person perspective, the control of the head movement is not shared and intentionally kept independent for each user, mainly to prevent CS.

In such systems between a trainer and a trainee, being able to redirect the focus and attention of the trainee to a specific part of the body or the environment is crucial for the training process, but should not hinder with it. Therefore, there is a need to unobtrusively guide the user's view to the desired POI, without inducing CS or disrupting the main focus from the task at hand. To achieve this, a view guidance technique based on visual effects has been presented and initially investigated in static target locations. This solution aims to simulate a synchronous viewpoint by guiding the trainee to the trainer's viewpoint using post processing visual effects, particularly blur effects.

1.2 Usage of Blur

Blur has previously been used in various forms to both guide attention and mitigate CS. A method that manipulates the resolution of specific areas in images has shown that the attention of viewers can be guided to the higher resolution parts without them realizing it [4]. Furthermore, blur-related techniques have been proposed by researchers to reduce the induced CS such as incorporating spatial or defocus blur [5], or by reducing the image quality of non-foveal regions to mitigate CS by applying blur [6]. Therefore, we decided to utilize blur into our proposed technique.

図 1: The four conditions in the virtual apartment: (a)Blur, (b)Overlay, (c)Arrows and (d)Crosshair.

2. The User Study

2.1 Participants and Experimental Setup

We recruited 12 student participants via SNS (two females and ten males). All of them were between 22 and 35 years old (M=26.5, SD=3.93) with normal or corrected-to-normal vision. Five of them had prior experience in VR with three of them using it multiple times a week. An HTC Vive Pro Eye and a Vive controller were used with a laptop equipped with an RTX 3080 Ti 16GB GPU and i9-12900HK CPU. The experimental environment was developed in Unity 2020.3.32f1.

2.2 Experimental Conditions

In this preliminary study we used a within-subject design with the following four conditions (Figure 1).

- Blur (BL)
- Overlay (OL)
- Arrows (AR)
- Crosshair (CH)

We propose a BL technique (Figure 2) that applies a sphere around the user's head with the normal vectors inverted, such that the blur material can be visible from the interior. Then a blur mask shader is applied to the material, which maps a rectangular clear cut-out on the sphere's surface that would represent the FOV of the trainer. For this preliminary study the target views were static (not dynamically changing, as expected in a virtual co-embodiment system) and the blur technique was compared with three others inspired by previous works. The AR technique attached several arrows in the user's peripheral, acting as a compass always pointing towards the target POI. The CH technique simply attached a crosshair both in the center of the user's visual field and in the target location, with the aim of properly matching them. The OL technique overlaid the target FOV as a semitransparent copy of the environment within the target visual field.

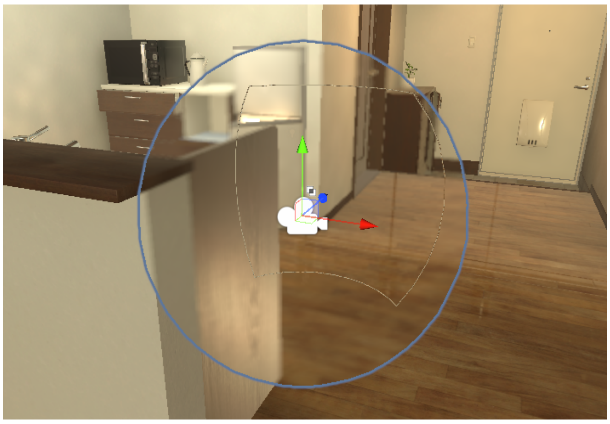

図 2: The proposed Blur technique: a blurred sphere surrounding the user's camera with a rectangular clear cut-out mask.

2.3 Measurements

For subjective measurements we used the following questionnaires after each condition (after 10 trials).

1. IPQ (igroup presence questionnaire) to measure the users' sense of presence in the environment.
2. SSQ (simulation sickness questionnaire) to measure the users' level of sickness.
3. SUS (system usability scale) to measure the usability of each system/technique.
4. NASA-TLX (NASA task load index) to assess the perceived workload.

For each trial the application recorded the following.

1. Duration for view matching the target orientation.
2. Subjective difficulty score on a 7-Likert scale with a Single Ease Question(SEQ) where 1 is extremely easy and 7 is extremely hard.

2.4 Experimental Procedure

The experiment was carried out according to the following procedure.

1. Topic introduction and flow explanation, followed by participant's given consent and a basic questionnaire for demographic data.
2. Tutorial VR session where users familiarize with the VR equipment and the four conditions (BL, OL, AR, CH) with 2 test trials per condition.
3. Participants perform a total of 40 trials, 10 for each of the four conditions (BL, OL, AR, CH). The four questionnaires (SSQ, IPQ, SUS, NASA-TLX) are answered after each condition.
4. Semi-structured interviews for feedback.

表 1: **Perceived trial difficulty with a SEQ and trial duration for all conditions. SEQ scores are on a 7-point Likert scale and duration is in seconds.**

	Blur		Overlay		Crosshair		Arrows	
	M	SD	M	SD	M	SD	M	SD
SEQ	2.57	1.26	4.05	1.69	2.89	1.33	3.13	1.64
Duration	5.76	2.79	11.25	10.15	7.80	4.39	9.58	7.89

5. Participant receives ￥2000 Amazon gift card as a reward.

For the experiment trials the participants were seated on a swivel chair that allowed them to fully rotate around and were equipped with the HMD to look for targets. At the start of the experiment they were immersed in an isolated hub area where instructions were presented. When users felt ready, they pressed the start button, and one of the four conditions was randomly assigned. They were transported in a randomly selected area out of set spawn locations within a virtual apartment where they were tasked to find the randomly placed target view by rotating around and orienting their head to match the entire desired visual field. When they felt confident about their view matching they pressed the trigger to lock their orientation and get transported back to the hub area to answer a SEQ about the difficulty of the trial. When they confirmed their rating they were once again transported back to the apartment where the next trial was loaded. After all ten trials for the conditions were completed the participant was assisted in removing the HMD and guided towards a desk to complete the four Questionnaires.

2.5 Experimental Hypotheses

As BL is the proposed technique, we will focus on comparisons between the BL and other techniques and not between the non BL techniques. It is hypothesized that BL will be faster and with less mental and physical workload than the rest of the techniques. We also hypothesize that BL will maintain a high level of immersion and will be perceived as the easiest and less frustrating with lower scores of CS than OL and AR.

3. Results

The completion time and perceived difficulty are summarized in Table 1. BL is perceived as the easiest, which also translates to faster trial completion times. The questionnaire results are summarized in Table 2. Friedman's test revealed a significant p-value ($p = 0.0087$) for the general presence but not for any IPQ subscales. A post hoc analysis showed us that the difference is due to the difference in general presence between Crosshair and Blur ($p = 0.0088$) as well as Crosshair and Overlay ($p = 0.048$). In Figure 4 we plot the parallel coordinates and use boxplots to show the differences between all pairs of groups. In this respect, the post hoc analysis can be thought of as performing paired wilcox test with correction for multiplicity. Blur ranks above the average (68 on a scale from 0-100) for the SUS score along with CH and AR, but the difference between them was not significant ($p = 0.066$).

表 2: **Results for presence (IPQ) and usability (SUS) questionnaires for all conditions. SUS scores have a range of 0-100, whereas IPQ scores are on a 7-point Likert scale.**

		Blur		Overlay		Crosshair		Arrows	
Quest	Subscale	M	SD	M	SD	M	SD	M	SD
IPQ	General	3.85	1.19	3.91	1.5	4.83	1.03	4.5	1.17
	Realism	2.5	0.91	2.54	0.73	2.98	0.76	2.52	0.82
	Involvement	4.13	1.11	4.31	1.08	4.27	1	4.23	1.3
	Spatial	4.1	0.8	3.87	1	4.37	0.94	3.75	1.36
SUS		74.16	19.22	53.75	23.51	77.7	19.84	68.95	26.42

Figure 3 shows the mean SSQ and NASA-TLX overall scores for each condition. Aside a few reports of eye strain from new to VR participants after conditions OL and CH, no discomforts were reported. In total sickness BL performed better than OL but not AR. The assessment of work load revealed the lowest scores for the BL condition overall and in all subscales. With an average weighted overall score of 34.28 it is ranked on the medium to somewhat high workload.

4. Discussion

The results from the study show that the Blur technique performed comparatively well in tasks where users need to find a static target visual field in an immersive environment. As hypothesised, BL outperformed the other techniques in mental workload and is perceived as the easiest and less frustrating. Further, we found comparable levels of presence, usability and CS for the technique.

Regarding the completion time, the favorable results might be due to the fact that there was less visual feedback information provided that assisted in the accuracy of matching the target view point. Many participants interestingly stated that they didn't feel the urge to try as hard to precisely match the target orientation as in OL or CH, but instead they were satisfied with just making the entire FOV unblurred. Despite BL also providing accuracy assisting information in the form of a thin

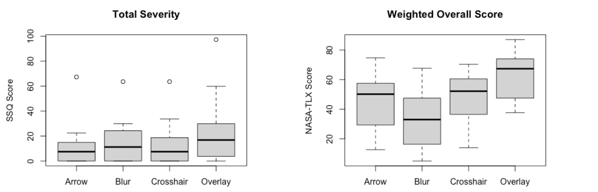

図 3: SSQ total scores (left) and NASA-TLX weighted overall scores (right) for all conditions.

and subtle border around the clear cut-out of the mask shader, it was not mentioned in the interviews probably due its placement in the peripheral view of users when they where looking at the center of the target. The AR condition was the only one that provided initial direction information that negated the need to arbitrary look around to find the target, something that was especially true in the CH condition where the target indicator was hard to find. Some participants utilized the spatial information provided by the OL technique in combination with memory to get a sense of initial direction.

In a virtual co-embodiment scenario where instead of static targets, the target is dynamically changing based on the movement of the counterpart's head or gaze, most participants came to agreed that if the goal is to visually locate a specific object or when precision is crucial, then their preference is shifted towards techniques other than BL. In scenarios where non-obtrusive visual guidance is required for just the general region in the environment (which is the case we want in a virtual co-embodiment system), the BL technique was favored by the majority.

5. Conclusion

In this preliminary study, we propose a visual attention guidance technique using blur, with the aim of being an alternative approach to view sharing in virtual co-embodiment systems. In our experiment, participants used our technique along with three other visual guidance techniques on matching static view targets in a virtual apartment. We evaluated and compared our technique using subjective questionnaires on presence, mental workload, usability and time performance and derived feedback for possible improvements. Results, indicate that our blur technique can be a subtle, quick and low workload visual guidance system with comparable levels of presence and CS. In future work, we will examine whether the good effects of the blur technique translate well in dynamically changing targets, with focus on accuracy and the ability to follow the target while also incorporating eye-tracking technology. Furthermore, we plan to apply an improved version of our technique to a virtual

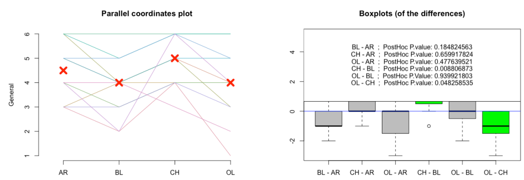

図 4: Visualization of post hoc analisys for friedman's test for General Presence. Parallel coordinates on the left and a boxplot of the differences between all pairs of groups on the right.

co-embodiment system between a trainer and a trainee.

謝辞 This work was supported by JSPS Grant-in-Aid for Scientific Research (A) Grant Number 21H04883 and Grant-in-Aid for Scientific Research (S) Grant Number 19H05661.

参考文献

[1] S. Grogorick and M. Magnor, "Subtle visual attention guidance in vr," in *Real VR–Immersive Digital Reality*, pp. 272–284, Springer, 2020.

[2] S. Grogorick, G. Albuquerque, J.-P. Tauscher, and M. Magnor, "Comparison of unobtrusive visual guidance methods in an immersive dome environment," *ACM Transactions on Applied Perception (TAP)*, vol. 15, no. 4, pp. 1–11, 2018.

[3] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose, and A. Lécuyer, "Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 10, pp. 4023–4038, 2020.

[4] H. Hata, H. Koike, and Y. Sato, "Visual attention guidance using image resolution control," **情報処理学会論文誌ジャーナル** *(Web)*, vol. 56, no. 4, pp. 1152–1161, 2015.

[5] S. Ang and J. Quarles, "Gingervr: An open source repository of cybersickness reduction techniques for unity," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 460–463, IEEE, 2020.

[6] Y.-X. Lin, R. Venkatakrishnan, R. Venkatakrishnan, E. Ebrahimi, W.-C. Lin, and S. V. Babu, "How the presence and size of static peripheral blur affects cybersickness in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 17, no. 4, pp. 1–18, 2020.