Supplementary Material of “GazeRing: Enhancing Hand-Eye Coordination with Pressure Ring in Augmented Reality”

Zhimin Wang[1] Jingyi Sun[1] Mingwei Hu[2] Maohang Rao[1] Weitao Song[2] Feng Lu[1]*

[1] State Key Laboratory of VR Technology and Systems, School of CSE, Beihang University, Beijing, China
[2] School of Optics and Photonics, Beijing Institute of Technology, Beijing, China

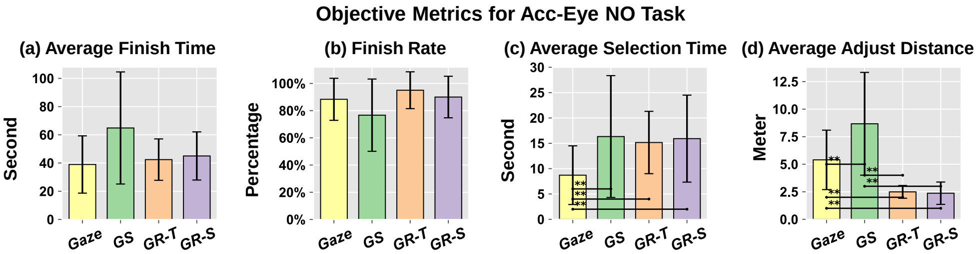

Figure 1: Bar charts of performance of multimodal interactions under different objective metrics for the Acc-Eye NO Task. Error bars indicate the standard error. The statistical significances are labeled with ** ($p < 0.05$).

1 ANOVA ASSUMPTIONS

We conducted repeated-measures ANOVAs for the results in Section 4.7 of the main text. Prior to using ANOVA, normality tests and homogeneity of variance tests should be conducted. Here is a brief summary of the test results:

Normality Tests: we use the Shapiro-Wilk (S-W) test and plot normal distribution histograms for all data used for ANOVA analyses. We then calculated the kurtosis and skewness of the data distributions. The S-W test indicated that the data did not strictly follow a normal distribution (p $<$ 0.05). However, the histograms of the data distributions showed a bell-shaped curve, indicating the presence of normality. Furthermore, the kurtosis values were within the range of -10 to +10, and the skewness values were within the range of -3 to +3. Based on guidelines [1], we conclude that the data is not absolutely normal, but can be accepted as normally distributed.

Homogeneity of Variance Tests: we found that some data did not satisfy the homogeneity of variance assumption. Consequently, the computational tool automatically employed Welch’s ANOVA, which has been shown to be robust in cases of heterogeneity of variance [2].

In summary, the data in this paper can be accepted as normally distributed, and for data that did not meet the homogeneity of variance assumption, Welch’s ANOVA was employed. Therefore, the data analysis in our paper is considered reasonable.

2 RESULTS

In Section 4.7 of the main text, we primarily focused on the most complex scenario: the Insufficient Eye Tracking Heavy Occlusion (Ins-Eye HO) Task. Here, we provide supplementary discussion on the results of the three remaining scenarios: the Accurate Eye Tracking No-Occlusion (Acc-Eye NO) Task, the Accurate Eye Tracking Heavy Occlusion (Acc-Eye HO) Task, and the Insufficient Eye Tracking No-Occlusion (Ins-Eye NO) Task.

Acc-Eye NO Task. The objective metrics for Acc-Eye NO Task are shown in Fig. 1. Repeated measures ANOVAs revealed significant differences among the four interaction techniques in terms of Average Selection Time ($p = 0.019$) and Average Adjust Distance (p<0.001). However, no significant differences were observed for Average Finish Time ($p = 0.217$) or Finish Rate ($p = 0.17$). The **Average Selection Time** for *Gaze* was significantly lower than *GS*, *GR-T*, and *GR-S* ($p = 0.024, 0.01, 0.025$), indicating that *Gaze* can quickly complete selection tasks in the absence of occlusion and with accurate eye tracking. However, the **Average Adjust Distance** showed a contrasting result, with *GR-T* and *GR-S* demonstrating superior performance compared to *Gaze* ($p = 0.002, 0.003$) and *GS* (p < 0.001, < 0.001), suggesting that *GR-T* and *GR-S* exhibit significant advantages in the efficiency and accuracy of moving spheres, even in tasks most favorable to *Gaze*.

Acc-Eye HO Task. The objective metrics for Acc-Eye HO Task are shown in Fig. 2. Repeated measures ANOVAs revealed significant differences among the four interaction techniques in terms of Average Finish Time ($p = 0.031$), Finish Rate ($p = 0.021$), Invalid Selection Count (p<0.001) and Average Adjust Distance ($p = 0.014$). However, no significant difference is observed for Average Selection Time ($p = 0.237$). With the introduction of occluding objects, *Gaze* no longer excelled in the **Average Finish Time**, with both *GR-T* and *GR-S* showing significant advantages over *Gaze* ($p = 0.003, 0.017$). Additionally, *GR-T* was significantly shorter than that of *GS* ($p = 0.039$) and comparable to *GR-S*. *GR-T* and *GR-S* maintained a higher **Finish Rate**, whereas the finish rates for *Gaze* and *GS* were significantly lower than those for *GR-T* ($p = 0.004, 0.034$) and *GR-S* ($p = 0.01, 0.028$), indicating that the presence of occluding objects greatly increased the task’s difficulty. Although the occluding objects did not completely obscure the target objects, the *Gaze* was severely impacted, as evidenced by the higher number of **Invalid Selection Count** compared to *GS*, *GR-T* and *GR-S* (p<0.001, <0.001, <0.001), requiring more attempts to move the occluding objects. However, analysis of the **Average Adjust Distance** revealed that *GR-T* and *GR-S* demonstrated a notable advantage over *GS* ($p = 0.006, 0.008$), resulting in shorter movement distances.

Ins-Eye NO task. The objective metrics for Ins-Eye NO Task are shown in Fig. 3. Repeated measures ANOVAs revealed significant differences among the four interaction techniques on all objective metrics ($p = 0.003$ for the Average Selection Time and p<0.001 for others). Without the influence of occluding objects but with the interference of insufficient eye tracking, the **Average Finish Time** for *Gaze* significantly increased, becoming more than twice that of *GS* (p $<$ 0.001) and more than three times that of *GR-T* (p $<$ 0.001) and *GR-S* (p $<$ 0.001), while *GR-*

---

*Corresponding Author: Feng Lu (lufeng@buaa.edu.cn)

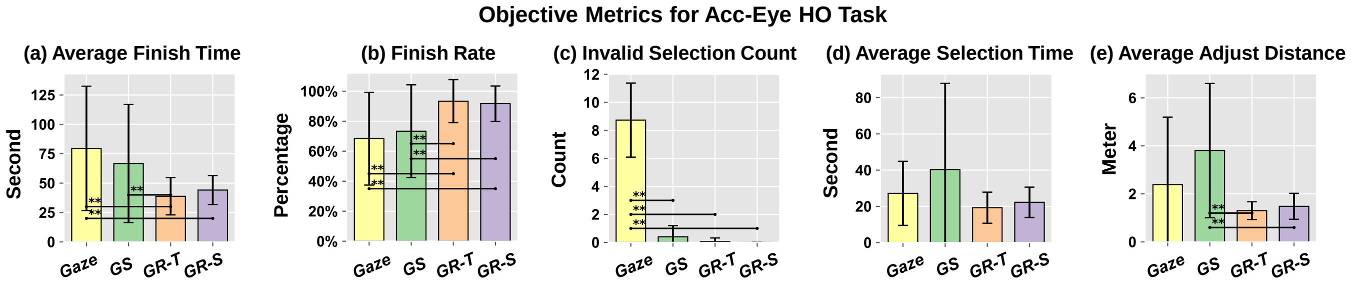

Figure 2: Bar charts of performance of multimodal interactions under different objective metrics for the Acc-Eye HO Task. Error bars indicate the standard error. The statistical significances are labeled with ** ($p < 0.05$).

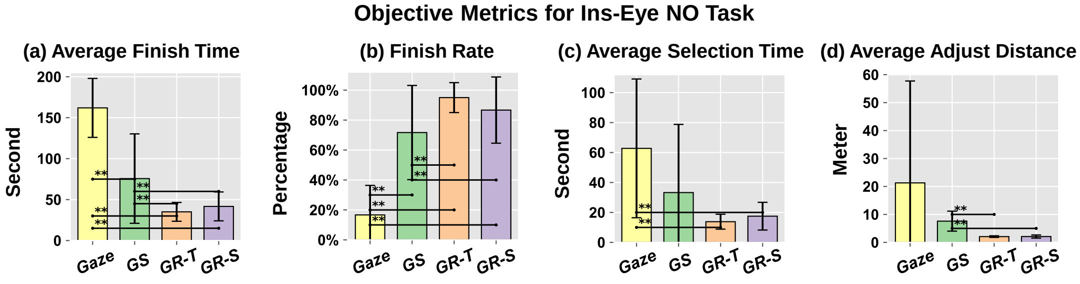

Figure 3: Bar charts of performance of multimodal interactions under different objective metrics for the Ins-Eye NO Task. Error bars indicate the standard error. The statistical significances are labeled with ** ($p < 0.05$).

T and *GR-S* still outperformed *GS* ($p = 0.016, 0.027$). The **Finish Rate** for *Gaze* and *GS* continued to decline, both significantly lower than those for *GR-T* (p < 0.001, = 0.025) and *GR-S* (p < 0.001, = 0.033). In terms of **Average Selection Time**, *GR-T* and *GR-S* performed excellently, being significantly lower than *Gaze* ($p = 0.001, 0.003$) and slightly lower than *GS*. Regarding **Average Adjust Distance**, *GR-T* and *GR-S* were significantly lower than *GS* (p < 0.001, < 0.001), with *Gaze* showing a large variance, indicating excessively long movement distances in some cases. Overall, regardless of the influence of occlusions or the insufficient of eye tracking, *GR-T* and *GR-S* consistently exhibited excellent performance, whereas *Gaze* and *GS* gradually became less suited to these challenging tasks.

3 ACCESSBILITY OF GAZERING

The accessibility of GazeRing to users with disabilities. Our GazeRing demonstrates potential benefits for users with limited hand mobility, such as those with arm fractures, as it supports subtle finger interactions without requiring extensive arm movement. The design of GazeRing also shows promise for users with visual impairments, such as myopia. The study explored object manipulation under the Insufficient Eye Tracking condition (4° error). The interaction strategies designed in this research, including the Refine Gaze Cone and Refine Gaze Beam phases, enable fine-tuning of selections when eye tracking accuracy is insufficient.

Future iterations of GazeRing. The GazeRing technique facilitates private and subtle interactions, potentially enhancing user experience in public settings. As demonstrated in our supplementary video, users can employ GazeRing with less noticeable gestures in environments such as subway carriages. Future iterations may implement skin-inspired flexible sensors attached to the finger surface for more private interactions, with no visible hardware.

REFERENCES

[1] R. B. Kline. *Principles and practice of structural equation modeling*. Guilford publications, pp. 76-77. 2023. 1

[2] K. Moder. Alternatives to f-test in one way anova in case of heterogeneity of variances (a simulation study). *Psychological Test and Assessment Modeling*, 52(4):343, 2010. 1