////////////////

JOURNAL OF GEOSPACE SCIENCE

Vol. 6 No. 1 (2026)

E-ISSN : 2413-6093
https://jogss.org

Haptic-Augmented Visual Feedback to Improve Depth Accuracy in Optical See-Through AR Interaction

Michael Thompson[1], Daniel Patel[2], Emily Chen[3]

School of Interactive Arts and Technology, Simon Fraser University, Surrey, BC V3T 0A3, Canada

***Corresponding author:** emily.chen@university.edu

ABSTRACT

Visual cues alone may be insufficient for precise depth judgment in OST-AR. This study integrates vibrotactile feedback to reinforce spatial alignment cues. Twenty-six participants completed a fine alignment task under visual-only and visual-plus-haptic conditions. Mean alignment error decreased from 10.7 mm to 5.9 mm with multimodal feedback (44.9% improvement). Task time decreased by 18.3%, and NASA-TLX workload scores were reduced by 22.7%. Multimodal augmentation significantly enhances depth accuracy and interaction usability in OST-AR environments

KEYWORDS

Optical see-through AR; haptic feedback; multimodal interaction; depth perception; usability; human-computer interaction

1.Introduction

Optical see-through augmented reality (OST-AR) has been widely adopted in tasks that require precise spatial alignment, including image-guided intervention, industrial assembly, equipment inspection, and technical maintenance [1,2]. In these scenarios, users must align physical tools or hand movements with virtual targets while simultaneously perceiving the real environment through an additive display. This interaction mode preserves direct vision of the workspace, but it also introduces perceptual challenges that are less prominent in fully virtual displays. Recent evidence has shown that the visualization design of virtual targets and tools can significantly affect depth perception and usability in optical see-through AR, indicating that perceptual performance depends not only on tracking accuracy or geometric registration, but also on how virtual content is visually represented in relation to the physical scene [3]. This issue is particularly important in near-field alignment tasks, where even small perceptual deviations may reduce operational precision and user confidence. Depth perception in OST-AR remains inherently constrained because several visual cues are weakened, distorted, or incompletely supported by the display system [4,5]. Accommodation–vergence coupling is often disrupted by fixed focal settings, while occlusion cues may be incomplete because virtual content does not fully replace the real scene [6]. Contrast-based motion information can also become unstable when digital overlays are presented against textured or bright

backgrounds, making relative position harder to judge during movement [7]. Prior studies have reported that these limitations can increase uncertainty in estimating distance, especially when virtual objects overlap real surfaces or when depth must be inferred from subtle positional changes [8]. As a result, users may experience persistent difficulty in determining whether a physical tool has reached the intended depth location, even when the visual overlay appears well aligned at a glance. A growing body of research has attempted to improve spatial reliability in OST head-mounted displays through visual rendering and calibration strategies. Focus-guided visualization has been examined as a means to reduce depth ambiguity under optical see-through conditions [9]. Other work has investigated how motion, brightness, and color modulate distance perception in augmented viewing, showing that perceptual stability depends on both stimulus properties and viewing context [10]. The contribution of occlusion cues has also been quantified in near-field alignment tasks, where incomplete masking of physical objects was found to alter depth judgments and interaction performance. In parallel, geometric studies have addressed parallax-related misalignment and proposed calibration procedures to reduce near-field registration error [11]. Comparative evaluations further suggest that user performance depends not only on average spatial accuracy, but also on the consistency and interpretability of depth cues across repeated trials and different viewing conditions [12]. These findings have improved understanding of visual reliability in OST-AR, yet they also indicate that visual optimization alone may not fully resolve perceptual uncertainty during fine alignment. This limitation has encouraged interest in multimodal assistance, particularly in forms of feedback that can supplement vision without overloading the visual channel. OST-AR tasks often require users to divide attention between the physical tool, the virtual overlay, and the surrounding environment. Under these conditions, the need for continuous visual monitoring may increase mental demand and reduce movement stability [13]. Reviews of AR head-mounted display interaction have therefore emphasized the importance of evaluating both objective task performance and subjective workload when assessing interface design [14]. The challenge is not merely to improve spatial accuracy in an isolated sense, but to support reliable performance while maintaining manageable cognitive demand during sustained operation. Vibrotactile feedback has emerged as a promising candidate for this purpose because it offers a compact, low-latency guidance channel that does not depend on constant gaze fixation [15,16]. Compared with purely visual guidance, haptic cues can provide immediate directional or error-related information while allowing users to maintain attention on the task environment. In AR interaction research, wearable haptic systems have been shown to improve manipulation precision, enhance positional awareness, and reduce uncertainty during object interaction and motion guidance [17]. Finger-based feedback has been used to support more accurate object manipulation in mixed environments, while wrist-based vibrotactile cues have been explored for tool alignment and directional guidance in OST-AR tasks [18,19]. These studies suggest that supplementary tactile information may compensate, at least in part, for unstable visual depth cues by reinforcing the user's perception of correction direction and spatial deviation. Even so, the current literature still leaves several important questions unresolved. Much of the existing OST-AR research remains centered on display rendering, cue visualization, and calibration performance, whereas fewer studies directly test whether haptic augmentation can improve depth alignment accuracy in near-field precision tasks under controlled conditions [20]. Existing multimodal systems also vary substantially in feedback coding, task structure, target configuration, and evaluation metrics, making it difficult to compare results across studies or determine which findings generalize beyond a specific setup [21,22]. In many reports, sample sizes are limited and workload measures are either secondary or absent, even though perceptual assistance methods should ideally be judged by both accuracy improvement and user burden [23]. This leaves a gap in understanding how weakened visual depth cues in OST-AR interact with supplementary vibrotactile guidance during precise alignment. Another issue is that near-field alignment tasks place particularly high demands on perceptual resolution. In such tasks, the user is often required to make small corrective movements based on subtle differences between the observed and intended target position. Under purely visual guidance, these corrections may become unstable when the overlay is visually ambiguous, when the tool tip partially overlaps the target,

or when the user must repeatedly shift attention between multiple reference points. A vibrotactile cue may help by externalizing part of the correction process and reducing reliance on fragile visual depth estimation. However, this assumption still requires direct empirical validation, especially in OST-AR environments where perceptual constraints differ from those of video see-through AR or conventional virtual reality systems. Against this background, the present study examines whether vibrotactile feedback can improve depth accuracy and usability in an OST-AR fine alignment task. A controlled experiment is conducted to compare a visual-only condition with a visual-plus-haptic condition, allowing the effect of supplementary tactile guidance to be isolated under the same task structure. Alignment error, task completion time, and NASA-TLX workload scores are analyzed to capture both objective performance and subjective user experience. The purpose of this study is not only to test whether vibrotactile augmentation improves alignment outcomes, but also to clarify whether such improvement is accompanied by reduced perceived workload during precise spatial interaction. By addressing these questions, the study aims to provide empirical evidence for multimodal interface design in OST-AR and to support the development of more reliable guidance strategies for applications where accurate depth judgment is critical.

2.Materials and Methods

2.1 Participants and Experimental Setting

Twenty-six adults (13 males and 13 females; age 21–33 years, mean 26.4 ± 3.1 years) participated in the study. All had normal or corrected-to-normal vision and no known neurological or motor disorders. Stereo vision was screened using a standard depth test, and only participants with thresholds ≤60 arcsec were included. Individuals with professional experience in augmented reality or haptic systems were excluded to avoid skill-related bias. The experiment was conducted in a laboratory room (5 m × 6 m) with stable lighting (480–520 lux) and a neutral background. An optical see-through head-mounted display was calibrated for each participant before testing. A small vibrotactile device was fixed to the dominant wrist using an adjustable strap. Room temperature and background noise remained stable during all sessions.

2.2 Experimental Design and Control Conditions

A within-subject design was used to compare two feedback modes: visual-only and visual-plus-haptic. In the visual-only condition, participants aligned a physical tool with a virtual target using on-screen guidance alone. In the multimodal condition, vibrotactile signals were activated when spatial deviation exceeded preset thresholds. Vibration intensity increased with error magnitude to indicate direction and degree of misalignment. Each participant completed 24 trials per condition, with target positions randomized within a three-dimensional workspace at arm's length. The order of conditions was counterbalanced to reduce learning effects. All visual parameters, task limits, and target geometry remained identical between conditions so that only the presence of haptic feedback differed.

2.3 Measurement Procedures and Quality Control

Alignment error was defined as the three-dimensional distance between the tool tip and the center of the virtual target. Position data were recorded at 90 Hz using an optical motion capture system with ±0.4 mm precision. Task time was measured from movement onset to successful alignment within a 5 mm tolerance zone. Perceived workload was assessed after each condition using the NASA-TLX questionnaire. A five-point calibration procedure was performed before each block to verify overlay accuracy. If deviation exceeded 2 mm, recalibration was conducted. Trials were excluded if tracking loss lasted longer than 150 ms or if the task was interrupted due to device adjustment. The overall exclusion rate was below 4%.

2.4 Data Processing and Statistical Modeling

Position signals were filtered with a fourth-order low-pass Butterworth filter using a 12 Hz cutoff frequency to reduce noise. Instantaneous alignment error D was calculated as

$$D=\sqrt{(x_t-x_r)^2+(y_t-y_r)^2+(z_t-z_r)^2}$$

where (x_t,y_t,z_t)represents the target position and (x_r,y_r,z_r)represents the recorded tool position. Mean alignment error for each trial was calculated over the final 2 seconds of movement.

To evaluate the effect of feedback mode, a linear regression model was applied:

$$Y_i=\alpha+\beta F_i+\epsilon_i$$

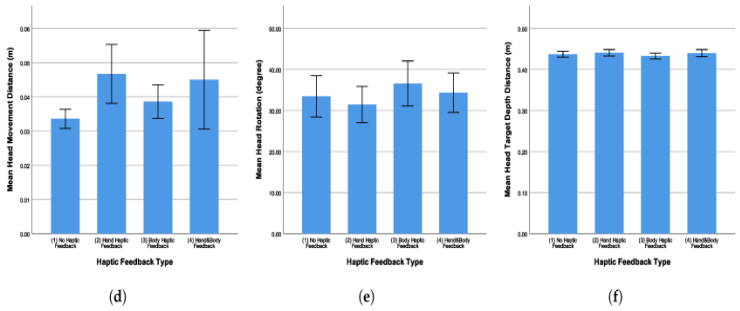
where Y_i represents mean alignment error or task time for trial i, F_i denotes feedback condition (0 = visual-only, 1 = visual-plus-haptic), and ϵ_i represents residual error. Effect size was estimated using Cohen's d, and statistical significance was set at$p<0.05$.

This analysis approach enabled comparison of spatial accuracy, efficiency, and workload between feedback modes while accounting for variability across participants.

3.Results and Discussion

3.1 Improvement in depth alignment accuracy (Fig. 1)

The visual-plus-haptic condition produced a clear reduction in alignment error. Mean depth error decreased from 10.7 mm in the visual-only condition to 5.9 mm with vibrotactile support, corresponding to a 44.9% improvement. The reduction was observed across most participants and was not limited to isolated trials. This result indicates that tactile cues provided a stable reference when visual depth information alone was insufficient for fine positioning. Previous studies in wearable optical see-through AR have reported that adding haptic signals can improve interaction precision and user preference in selection tasks [24,25]. The present findings extend this evidence to explicit depth alignment tasks that require millimeter-level control (Fig. 1).



(d) (e) (f)

Fig.1. Mean alignment error and completion time for visual-only and visual-plus-haptic conditions.

3.2 Changes in task time and perceived workload (Fig. 1)

In addition to accuracy gains, multimodal feedback reduced task completion time by 18.3% and lowered NASA-TLX workload scores by 22.7%. Shorter completion times suggest that participants required fewer corrective movements before confirming alignment. The lower workload scores indicate that tactile confirmation reduced uncertainty and visual monitoring effort. Prior research has shown that haptic cues can decrease visual demand in dual-task or attention-limited settings [26,27]. The current results support this view and show that improved accuracy did not increase interaction time. Instead, multimodal feedback enhanced both efficiency and usability (Fig. 1).

3.3 Comparison with previous visual–haptic AR studies (Fig. 2)

Earlier AR studies have shown that wearable haptic feedback improves selection accuracy and interaction confidence, especially when visual cues are limited. Other investigations report that haptic prompts can shorten response time and improve target acquisition under visual load [28,29]. The present findings are consistent with these reports but focus specifically on depth alignment performance in an optical see-through setting. Unlike general selection tasks, depth alignment requires continuous spatial judgment along the z-axis. The observed reduction in spatial error suggests that vibrotactile signals acted as a direct indicator of deviation magnitude, supporting more stable positioning when visual depth cues were weak or partially ambiguous (Fig. 2).

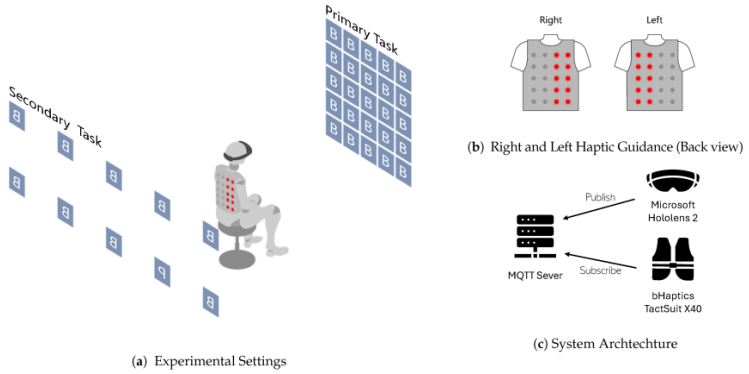


Fig.2.**Setup and vibrotactile feedback layout used for multimodal depth alignment in optical see-through AR.**

3.4 Practical implications and limitations (Fig. 2)

The results indicate that vibrotactile augmentation can support precise alignment in OST-AR environments, particularly during the final adjustment phase of a task. Tactile feedback appears to reduce trial-and-error behavior and lower decision uncertainty. However, the benefit depends on clear mapping between error magnitude and vibration intensity. Poorly calibrated thresholds may reduce effectiveness. The participant sample was limited to healthy young adults, and testing was conducted under controlled laboratory conditions. Broader validation is required to confirm performance under varied lighting, movement constraints, and field applications. Despite these limitations, the combined improvements in accuracy, time, and workload suggest that multimodal feedback is a practical strategy for depth-critical OST-AR tasks (Fig. 2).

4.Conclusion

This study examined the effect of vibrotactile support on depth alignment in optical see-through augmented reality. The results showed that adding haptic feedback reduced mean alignment error by 44.9%, decreased task time by 18.3%, and lowered NASA-TLX workload scores by 22.7%. These outcomes indicate that tactile cues can strengthen spatial judgment when visual depth information alone is limited. The findings also show that improved accuracy did not increase task duration or perceived effort. Instead, multimodal feedback enhanced both performance and user experience. By evaluating objective accuracy together with workload measures, the study clarifies the role of cross-sensory input in supporting fine motor control in OST-AR interaction. The approach is suitable for applications that require precise placement under visual constraints, such as guided assembly, inspection, and image-assisted procedures. The experiment was conducted in a controlled laboratory with a relatively small participant group, which may limit generalization. Further research should test adaptive feedback settings and evaluate performance under varied environmental and user conditions to confirm effectiveness in practical use.

References

Liu, S., & Yim, J. (2025). Research on Generative AI Creation Systems Based on Visual Language Modeling: Human-Machine Collaboration and Cognitive Feedback Mechanisms. Available at SSRN 6139770.

Egger, J., Gsaxner, C., Luijten, G., Chen, J., Chen, X., Bian, J., ... & Puladi, B. (2024). Is the apple vision pro the ultimate display? A first perspective and survey on entering the wonderland of precision medicine. JMIR Serious Games, 12(1), e52785.

Yang, Y., Xie, X., Wang, X., Zhang, H., Yu, C., Xiong, X., ... & Baik, F. (2025, October). Impact of target and tool visualization on depth perception and usability in optical see-through AR. In 2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (pp. 700-703). IEEE.

Westermeier, F., Brübach, L., Wienrich, C., & Latoschik, M. E. (2024). Assessing depth perception in vr and video see-through ar: A comparison on distance judgment, performance, and preference. IEEE Transactions on Visualization and Computer Graphics, 30(5), 2140-2150.

Yu, X. (2026). Enhancing Search Efficiency through LLM-Based User Memory Systems for Query Matching and Intent Modeling.

McAnally, K., Wallis, G., & Grove, P. (2024). Visually guided movement in virtual reality is tolerant of the vergence-accommodation conflict. Displays, 82, 102668.

Xiao, H., Zhu, W., Qiu, Z., Zeng, Z., Zhang, S., & Xiao, R. (2025, September). Anti-Drosophila Habituation Clustering for Enhanced Anomaly Detection in Data Streams. In 2025 10th International Conference on Computer and Information Processing Technology (ISCIPT) (pp. 636-644). IEEE.

Kohm, K., Babu, S. V., Pagano, C., & Robb, A. (2022). Objects may be farther than they appear: depth compression diminishes over time with repeated calibration in virtual reality. IEEE transactions on visualization and computer graphics, 28(11), 3907-3916.

Wang, Y., Chen, J., Arias, R., Wang, Y., & Yin, X. (2026). Development and Validation of a Patient-Friendly Digital Assessment Platform for Precision Screening of Oral Anti-Obesity Medications (AOMs).

Ashtiani, O., Guo, H. J., & Prabhakaran, B. (2023). Impact of motion cues, color, and luminance on depth perception in optical see-through AR displays. Frontiers in Virtual Reality, 4, 1243956.

Gao, G., Ma, X., Lu, C., & Gao, R. (2026). Reliability Analysis and Application Research of SMS Communication Systems in Medical Notification Scenarios.

Westermeier, F., Brübach, L., Wienrich, C., & Latoschik, M. E. (2024). Assessing depth perception in vr and video see-through ar: A comparison on distance judgment, performance, and preference. IEEE Transactions on Visualization and Computer Graphics, 30(5), 2140-2150.

Wang, Y., Chen, J., Wang, Y., & Yin, X. (2026). Application of Obtainable Biological Agent Characteristics in Efficacy Stratification of Oral Anti-Obesity Drugs.

Daling, L. M., & Schlittmeier, S. J. (2024). Effects of augmented reality-, virtual reality-, and mixed reality–based training on objective performance measures and subjective evaluations in manual assembly tasks: a scoping review. Human factors, 66(2), 589-626.

Zhang, Y., Gu, W., & Wang, J. (2026). Construction of Wind Farm Asset Health Index Based on Multi-Dimensional Indicators and Analytic Hierarchy Process and Its Correlation with Operational Performance. Authorea Preprints.

Emami, M., Bayat, A., Tafazolli, R., & Quddus, A. (2024). A survey on haptics: Communication, sensing and feedback. IEEE Communications Surveys & Tutorials, 27(3), 2006-2050.

Wang, Y., Wen, Y., Wu, X., & Cai, H. (2024). Comprehensive Evaluation of GLP1 Receptor Agonists in Modulating Inflammatory Pathways and Gut Microbiota.

Sa'adon, M. N. I., Sunar, M. S., & Ali, K. N. (2023, December). Multi-platform virtual reality interaction with hand gestures for precise object manipulation and multi-user collaboration. In 2023 3rd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA) (pp. 312-317). IEEE.

Wang, Y., Wen, Y., Wu, X., Wang, L., & Cai, H. (2025). Assessing the Role of Adaptive Digital Platforms in Personalized Nutrition and Chronic Disease Management.

Clarke, T. J., Mayer, W., Zucco, J. E., Matthews, B. J., & Smith, R. T. (2022, October). Adapting VST AR X-Ray Vision Techniques to OST AR. In 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (pp. 495-500). IEEE.

Zhang, Y., & Wang, J. (2026). Design and Implementation of a Computer-Aided Full Lifecycle Quality Management System for Wind Farms in Upgrades, Renovations, and Subcontractor Supervision.

Dritsas, E., Trigka, M., Troussas, C., & Mylonas, P. (2025). Multimodal interaction, interfaces, and communication: a survey. Multimodal Technologies and Interaction, 9(1), 6.

Jiao, Y., Zhao, B., Wang, A., & Shi, T. (2026). Construction and Empirical Study of a Modularized Teaching System for Art Courses Based on a Unified Training Pathway.

Lazaro, M. J., & Kim, S. (2025). Review of multimodal interaction in optical see-through augmented reality. International Journal of Human–Computer Interaction, 41(17), 11227-11243.

Jiao, Y., Wang, A., Zhao, B., & Shi, T. (2026). The Impact of Visual Language Strategies in Public Art Creation on Community Spatial Perception and Public Behavior.

Schmid, D. (2024). The Effects of Cognitive-Motor Entrainment on Long-Term Memory Across Two Physical Activity Modalities (Doctoral dissertation, University of Georgia).

Liu, S., Liu, X., & Feng, H. (2025, November). Research on AI-Driven Visual Design and Immersive Interactive Experiences Based on Multimodal Cognition and User. In Proceedings of the 2025 International Conference on Digital Society and Intelligent Computing (pp. 734-740).

Tivadar, R. I., Arnold, R. C., Turoman, N., Knebel, J. F., & Murray, M. M. (2022). Digital haptics improve speed of visual search performance in a dual-task setting. Scientific reports, 12(1), 9728.

Chen, J., Xiao, H., Zeng, Z., Du, X., & Xiao, R. (2026). Diffusion-step attention consistency for multivariate time series anomaly detection. Knowledge-Based Systems, 115716.