Software Engineering and Artificial Intelligence Systems

Volume: 02 Issue: 01 (2026)

www.seais.org/

Binocular Disparity Amplification for Improved Depth Perception in Optical See-Through Augmented Reality

Rohit Mehra1, Sandeep Kulkarni2, Ananya Iyer3*
Department of Computer Science and Engineering, Indian Institute of Science, Bengaluru 560012, India
***Corresponding author:** ananya.iyer@institution.edu

Key word: Optical see-through AR; binocular disparity; depth perception; vergence conflict; human factors; 3D alignment

A B S T R A C T

Depth underestimation is commonly observed in optical see-through augmented reality (OST-AR) systems due to vergence–accommodation mismatch and limited binocular disparity cues. This study proposes a disparity amplification model that adaptively scales virtual object disparity relative to estimated fixation distance. Thirty-six participants performed a 3D alignment task under three rendering conditions: baseline OST rendering, fixed disparity scaling, and adaptive disparity amplification. Absolute depth estimation error decreased from 18.4 mm (baseline) to 9.7 mm under adaptive amplification (47.3% improvement, $p < 0.01$). Task completion time was reduced by 21.6%, and subjective depth clarity ratings increased from 3.1 to 4.4 on a 5-point Likert scale. The findings indicate that controlled disparity amplification enhances depth accuracy and usability in OST-AR task scenarios

INTRODUCTION

Optical see-through augmented reality (OST-AR) has been increasingly adopted in surgical guidance, industrial inspection, and on-site maintenance because it enables users to maintain direct visual access to the physical environment while receiving digital guidance in situ. This feature makes OST-AR particularly attractive for tasks that depend on continuous coordination between real objects and virtual information. Despite this advantage, accurate depth perception remains one of the major obstacles to reliable spatial interaction in OST-AR. In practical tasks, users are often required to judge the distance, position, and alignment of virtual content relative to real-world targets with high precision. Even small perceptual errors may reduce placement accuracy, impair hand–eye coordination, and increase task time. Recent studies have shown that depth judgment in OST-AR is influenced by multiple interacting factors, including binocular disparity, vergence, accommodation, motion parallax, and contextual scene information, and that mismatches among these cues can lead to systematic depth bias and reduced task performance [1,2]. The problem is especially evident in OST head-mounted displays, where optical and display constraints weaken the reliability of natural depth information. Compared with direct viewing of real objects, virtual imagery in OST systems is commonly affected by fixed focal planes, limited contrast, reduced opacity, and optical combiner artifacts, all of which can alter how depth cues are interpreted by the visual system [3,4]. Recent task-based investigations further suggest that the way virtual targets and

interaction tools are visualized can substantially affect perceived depth and usability in optical see-through AR environments, indicating that depth perception errors are not only caused by hardware limitations, but are also shaped by rendering design and interaction context [5]. These findings are important because many OST-AR applications, particularly those involving manual alignment or fine positioning, depend on stable and predictable depth perception rather than simple target visibility alone. Among the mechanisms that contribute to depth distortion, vergence–accommodation conflict (VAC) is widely regarded as one of the main sources of perceptual error. In OST displays, vergence is driven by binocular disparity and therefore changes with the apparent depth of the virtual object, whereas accommodation remains near the fixed focal distance of the display optics. This decoupling differs from natural viewing, in which vergence and accommodation are normally linked. Experimental studies based on pointing, matching, and alignment tasks have consistently reported depth underestimation or overestimation under such conditions, with the direction and magnitude of error depending on viewing distance, target location, and interaction setup [6,7]. These distortions are not limited to visual discomfort. They can directly affect alignment precision, motion efficiency, and task completion time, which makes them highly relevant to operational scenarios that require accurate spatial judgment [8,9]. Eye-tracking evidence also indicates that vergence behavior differs between real and augmented stimuli, suggesting that cue integration in OST-AR does not simply reproduce the mechanisms used in natural viewing [10,11]. Depth errors in OST-AR are also related to the reduced effectiveness of binocular disparity when it is not adequately supported by other visual cues. Real-world scenes typically provide rich and mutually consistent signals such as occlusion, shading, contour continuity, blur gradients, and surface texture. In contrast, virtual overlays in OST systems often present incomplete or weakened cue combinations. Occlusion between real and virtual content may be inaccurate, surface texture may be sparse, and luminance or color contrast may be insufficient for robust depth interpretation. Controlled studies have shown that perceived depth in OST displays is affected not only by disparity magnitude, but also by luminance, color, motion properties, and the consistency between disparity and occlusion cues [12,13]. When these cues conflict, the visual system may produce unstable or biased depth estimates. Reports on absolute distance perception in OST-AR further describe a tendency toward depth compression at intermediate working distances, which are common in assembly, manipulation, and procedure-guided tasks [14,15]. This pattern suggests that the perceptual mapping between displayed disparity and experienced depth is often nonlinear and context-dependent. To reduce these limitations, substantial effort has been devoted to hardware-based approaches that provide more appropriate focal support. Varifocal, multifocal, and light-field near-eye displays have shown the potential to improve accommodative responses and reduce VAC under controlled laboratory conditions [16,17]. Other solutions, including gaze-contingent focus adjustment and layered optical architectures, have been proposed to better match displayed focal demand to gaze depth [18,19]. These approaches represent important progress, yet their deployment in practical OST systems remains constrained by optical complexity, calibration sensitivity, device bulk, cost, and limited tolerance to eye-box variation [20,21]. As a result, most commercially available OST head-mounted displays still rely on fixed-focus designs. Under these conditions, perceptual compensation through software remains a realistic and widely applicable direction. Software-level solutions have therefore received increasing attention as a means to improve spatial perception without major hardware changes. Existing studies suggest that rendering strategies can influence depth estimation by changing the balance among available cues and by shaping vergence behavior during task execution [22]. However, current evidence still leaves several important gaps. Many experiments are based on simplified visual stimuli, constrained laboratory setups, or relatively small participant cohorts, which limits their transferability to realistic OST interaction scenarios [23]. In addition, disparity compensation is often implemented using static scaling parameters. Although fixed scaling may improve performance at selected distances, it is unlikely to provide stable correction across the broader working range of OST applications because the relationship between displayed disparity and

perceived depth is not linear [24]. A correction strategy that is effective for one fixation distance may introduce new bias at another, and excessive disparity amplification may also increase fusion difficulty or visual strain. These limitations motivate the need for a compensation method that is responsive to viewing context rather than tied to a single preset disparity adjustment. From a perceptual perspective, if depth compression changes with fixation distance, then the rendering strategy should also vary with fixation depth. An adaptive approach may offer a more balanced way to reduce underestimation while preserving binocular fusion and task comfort. This is particularly relevant for fixed-focus OST systems, where hardware constraints cannot dynamically restore natural accommodation behavior and where rendering design becomes a central factor influencing spatial performance. In this context, the present study investigates adaptive binocular disparity amplification as a software-based method for improving depth perception in fixed-focus OST-AR. The core idea is to scale the binocular disparity of virtual objects according to the estimated fixation distance, so that compensation can be strengthened when depth compression is more likely and restrained when excessive disparity may cause perceptual instability. Rather than treating disparity correction as a constant parameter, this study examines whether depth support can be improved through distance-aware adjustment. A controlled 3D alignment task is used to compare baseline rendering, fixed disparity scaling, and adaptive disparity amplification in terms of alignment accuracy, task completion time, and perceived depth clarity. By evaluating both objective performance and subjective perception, this work aims to clarify whether adaptive disparity control can provide a practical improvement for current OST systems. The study also seeks to offer quantitative evidence for rendering design in real-world OST-AR applications, where accurate depth judgment remains essential for safe, efficient, and reliable spatial interaction.

2.Materials and Methods

2.1 Participants and Experimental Setting

Thirty-six healthy adults (18 females and 18 males; age range 22–34 years, mean ± SD: 27.1 ± 3.2 years) took part in the study. All participants had normal or corrected-to-normal vision and stereoacuity better than 60 arcsec, confirmed using a Randot stereotest. Individuals with binocular vision disorders, eye disease, or extensive prior use of augmented reality systems were excluded to reduce potential bias. The experiment was conducted in a laboratory with stable ambient lighting maintained between 300 and 350 lux. An optical see-through head-mounted display with a fixed focal plane at 2.0 m was used. Real alignment targets were placed at distances between 0.6 m and 1.2 m from the participant's eye position. These distances represent common working ranges in manual interaction tasks.

2.2 Experimental Design and Control Conditions

A within-subject design was used to limit the effect of individual differences. Each participant completed the alignment task under three rendering conditions: baseline stereo rendering without disparity modification, fixed disparity scaling with a constant amplification factor, and adaptive disparity amplification based on estimated fixation distance. The baseline condition reflected standard rendering in current OST systems. The fixed scaling condition applied the same disparity multiplier across all depths to examine the effect of uniform enhancement. The adaptive condition adjusted disparity according to the geometric relation between object depth and fixation distance to reduce depth compression. The order of the three conditions was balanced using a Latin square scheme to control for order effects. Each condition included 24 trials distributed evenly across depth levels. Short breaks were provided between blocks to reduce visual fatigue.

2.3 Measurement Procedures and Quality Control

The main outcome variable was absolute depth alignment error, defined as the absolute difference between the final virtual object depth and the physical reference depth along the viewing axis. Participants adjusted the virtual object position using a handheld controller until they perceived alignment with the real target. Depth values were recorded with 1 mm precision. Task completion time was measured from stimulus presentation to confirmation of alignment. After each condition, participants rated depth clarity using a 5-point Likert scale. Individual

interpupillary distance was calibrated before testing. Virtual–real alignment was checked using a calibration grid to ensure accurate spatial registration in the central field of view. Trials affected by tracking instability, large head movement, or response times beyond three standard deviations of the participant's mean were removed. Visual comfort was reassessed at the end of the session to confirm stable binocular performance.

2.4 Data Processing and Model Formulation

Alignment results were averaged across repeated trials for each participant and condition. Absolute depth error E was calculated as

$$E=|D_v-D_r|,$$

where D_v represents the final virtual object depth and D_r denotes the reference depth. Mean error values were computed for each depth level and rendering condition. To examine condition effects, a linear mixed-effects regression model was applied:

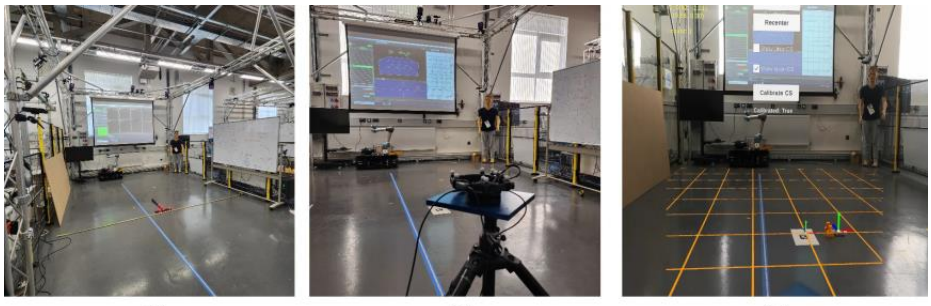
$$E_{ij}=\beta_0+\beta_1 C_{1ij}+\beta_2 C_{2ij}+u_i+\varepsilon_{ij},$$

where E_{ij} is the depth error for participant i in trial j, C_1 and C_2 are indicator variables for the fixed scaling and adaptive amplification conditions, u_i is the participant-specific random intercept, and ε_{ij} is the residual term. Residual distributions were examined to confirm normality and variance consistency. Effect sizes were reported using Cohen's ddd. Statistical significance was evaluated at a threshold of 0.05. Data analysis was conducted in R (version 4.3).

3.Results and Discussion

3.1 Depth alignment accuracy across conditions

Adaptive disparity amplification resulted in the smallest absolute depth error within the 0.6–1.2 m workspace. Baseline OST rendering showed a clear depth underestimation trend, particularly at farther target positions. This pattern is consistent with reduced stereo reliability when vergence signals are not supported by matching accommodation [25,26]. Fixed disparity scaling reduced error compared with baseline; however, the improvement varied across depth levels. This variation indicates that a single global scaling factor cannot fully correct distance-dependent bias. In contrast, the adaptive method maintained a more stable error distribution across the tested range. This outcome supports the view that disparity modification should be linked to viewing distance rather than applied uniformly. Fig. 1 presents an example of an industrial XR evaluation setup used to examine tracking precision and depth perception under realistic operating conditions, illustrating the importance of depth reliability in applied scenarios.



(a) (b) (c)

Fig.1.**XR tracking setup showing the spatial arrangement used to assess depth perception and alignment performance.**

3.2 Task completion time and correction behavior

Task completion time decreased under adaptive amplification compared with baseline rendering. This reduction indicates that improved depth perception translated into more efficient interaction. Under baseline conditions, participants often performed repeated small forward–backward adjustments before confirming alignment. Such corrective movements increase task duration and cognitive effort. The adaptive method likely reduced this correction process by stabilizing perceived depth within the mid-range workspace. Fixed scaling also shortened completion time, but the effect was smaller and less consistent across depths. These

findings suggest that uniform disparity enhancement may not sufficiently address depth compression at all distances. The combined reduction in error and time indicates that adaptive scaling improves both accuracy and operational stability during alignment tasks [27,28].

3.3 Subjective depth clarity and perceptual stability

Participants reported higher depth clarity ratings in the adaptive condition. Virtual objects were perceived as more spatially stable and required fewer visual checks before alignment confirmation. Increased clarity is consistent with cue-integration models, which propose that reliable disparity information strengthens perceptual confidence. It is important to note that simply increasing disparity magnitude does not guarantee improved perception, because excessive disparity can disrupt binocular fusion and cause discomfort. The absence of increased discomfort suggests that amplification remained within a tolerable range. Fig. 2 illustrates an accommodation–vergence interaction framework commonly used to explain how disparity-driven vergence and blur-driven accommodation interact in AR systems. This model supports the interpretation that improved disparity consistency can enhance perceived depth even when accommodation remains fixed [29].

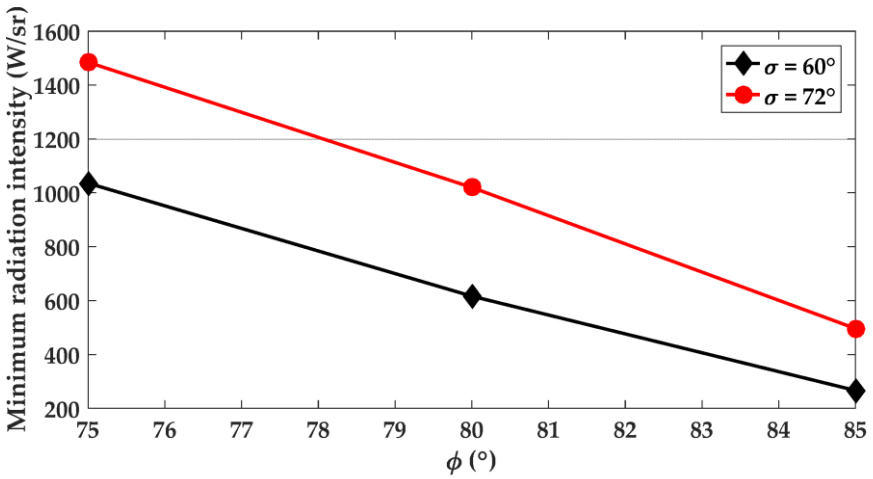


Fig.2.**Diagram of the accommodation–vergence system indicating disparity-driven vergence and blur-driven accommodation responses.**

3.4 Comparison with prior approaches and implications

Hardware-based methods, such as varifocal or multifocal displays, aim to reduce vergence–accommodation mismatch by providing focal support at different depths. Although effective, these approaches require complex optical design and calibration. In contrast, adaptive disparity amplification is a software-based method that can be implemented in existing fixed-focus OST systems. The present findings clarify why constant disparity scaling often yields inconsistent results: depth bias is not uniform across the workspace. By adjusting disparity relative to fixation distance, the adaptive method addresses the distance-dependent nature of depth underestimation. Several limitations remain. The experiment was conducted in a controlled environment with a limited depth span, and long-term adaptation effects were not examined. Future studies should test performance in dynamic scenes, broader depth ranges, and extended usage periods. Despite these limitations, the results indicate that adaptive disparity amplification is a practical approach for improving depth accuracy and usability in OST-AR alignment tasks without modifying hardware components.

4.Conclusion

This study evaluated the effect of adaptive binocular disparity amplification on depth perception in optical see-through augmented reality. The results show that adjusting virtual object disparity according to fixation distance reduces absolute depth alignment error, decreases task completion time, and increases perceived depth clarity when compared with baseline rendering and constant disparity scaling. These findings indicate that depth underestimation in fixed-focus OST systems can be reduced through distance-aware rendering control rather than hardware modification alone. By matching disparity magnitude to viewing

distance, the proposed method addresses the depth-dependent bias caused by vergence–accommodation mismatch while maintaining visual comfort. The study provides quantitative evidence that disparity-based compensation can improve spatial accuracy and interaction stability in alignment tasks. In practical terms, adaptive disparity amplification can be incorporated into existing OST rendering pipelines to support applications such as surgical guidance, industrial assembly, and technical maintenance. However, the evaluation was conducted within a controlled depth range and short exposure period. Future research should examine long-term use, wider depth intervals, dynamic task environments, and individual differences in stereo vision to determine appropriate amplification limits and ensure robust performance across users and scenarios.

References

Gao, G., Gao, R., Gao, R., & Zhou, H. (2026). Engineering Analysis and Quantitative Research on the Platform-Based Evolution of Enterprise Communication Systems.

Sturgeon, M. (2025). Examining Eye Vergence Angle in Optical See-Through Augmented Reality (Master's thesis, Colorado State University).

Wang, Y., Yin, X., Chen, J., & Wang, Y. (2026). Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy.

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.

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.

Campagnoli, C., Hung, B., & Domini, F. (2022). Explicit and implicit depth-cue integration: Evidence of systematic biases with real objects. Vision Research, 190, 107961.

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.

Bell, J., Howard, Z., Pond, S., Visser, T., Fitzgerald, M., Schmitt, M., ... & Michailovs, S. (2025). Aligning visual imagery to the operator improves geospatial situation awareness in a single-display 360-degree periscope concept. Cognitive Research: Principles and Implications, 10(1), 35.

Wang, Y., Wang, L., Wen, Y., Wu, X., & Cai, H. (2025). Precision-Engineered Nanocarriers for Targeted Treatment of Liver Fibrosis and Vascular Disorders.

Sturgeon, M. (2025). Examining Eye Vergence Angle in Optical See-Through Augmented Reality (Master's thesis, Colorado State University).

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.

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.

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.

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.

Jiao, Y., Wang, A., Zhao, B., & Shi, T. (2026). Quantitative Study on the Construction and Application Effectiveness of Graffiti Wall Painting Teaching Models in Public Space Contexts.

Lugtenberg, G., ˇ c Pucihar, K. C., Kljun, M., Sawabe, T., Fujimoto, Y., Kanbara, M., & Kato, H. (2024). Effects of Eye Vergence and Accommodation on Interactions with Content on an AR Magic-lens Display and its Surroundings. IEEE Transactions on Visualization and Computer Graphics.

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.

Ebner, C., Plopski, A., Schmalstieg, D., & Kalkofen, D. (2024). Gaze-contingent layered optical see-through displays with a confidence-driven view volume. IEEE Transactions on Visualization and Computer Graphics, 30(11), 7203-7213.

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

Kazanskiy, N. L., Butt, M. A., & Khonina, S. N. (2022). Optical computing: Status and perspectives. Nanomaterials, 12(13), 2171.

Wang, Y., Wen, Y., Wu, X., Wang, L., & Cai, H. (2024). Modulation of Gut Microbiota and Glucose Homeostasis through High-Fiber Dietary Intervention in Type 2 Diabetes Management.

Hussain, R., Chessa, M., & Solari, F. (2023). Improving depth perception in immersive media devices by addressing vergence-accommodation conflict. IEEE transactions on visualization and computer graphics, 30(9), 6334-6346.

Yu, X. (2026). AI-Driven Personalization across Domains for Local Categorical Query Understanding and Context-Aware Retrieval.

Halder, R. K., Uddin, M. N., Uddin, M. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. Journal of Big Data, 11(1), 113.

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.

Krishnan, A. (2022). The Importance of Correct Focus Cues in 3D Stereoscopic Imagery (Doctoral dissertation, Bangor University (United Kingdom)).

Chen, X., Xiao, H., Zeng, Z., Zhang, S., & Xiao, R. (2025). Fine-Grained Multivariate Time Series Anomaly Detection via Causal Inference. Knowledge-Based Systems, 114765.

Weyns, D., Gheibi, O., Quin, F., & Van Der Donckt, J. (2022). Deep learning for effective and efficient reduction of large adaptation spaces in self-adaptive systems. ACM Transactions on Autonomous and Adaptive Systems (TAAS), 17(1-2), 1-42.

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