Algorithms and Intelligence

Vol 2 Issue 1 (2026)

www.algointel.org

Influence of Virtual Target Size Scaling on Depth Estimation Bias in Optical See-Through Augmented Reality

Luca Bianchi1, Matteo Conti2, Giulia Romano3*

Department of Computer Science and Engineering, University of Bologna, 40126 Bologna, Italy

***Corresponding author:** giulia.romano@university.it*

Abstract

Perceived depth in OST-AR may be influenced by target size cues, leading to systematic bias. This study quantifies the relationship between virtual target scaling and perceived depth distortion. Twenty-eight participants estimated distances to virtual targets rendered at 0.5×, 1×, and 1.5× physical scale equivalents across distances from 0.3 m to 1.2 m. Linear regression analysis revealed a consistent compression bias at smaller scales, with mean underestimation of 12.6%. Size-matched rendering reduced bias to 4.3%, while exaggerated scaling introduced overestimation errors. Results demonstrate that accurate scale calibration is essential for reliable depth perception in OST-AR interaction tasks.

Keywords

Depth bias; optical see-through AR; size scaling; perceptual calibration; spatial perception; mixed reality

INTRODUCTION

Optical see-through augmented reality (OST-AR) enables users to observe the physical environment directly while virtual content is superimposed within the same visual field. Because it preserves direct sight of real objects and allows digital information to be integrated into ongoing manual tasks, OST-AR has attracted increasing interest in surgical navigation, industrial assembly, maintenance support, and other scenarios that require accurate spatial interaction [1]. Despite these advantages, stable and accurate depth perception remains a major limitation. Recent studies have shown that users often misjudge the distance of virtual content in head-mounted OST displays, and this distortion becomes especially evident during near-field and mid-range interaction, where precise alignment is often required [2]. Related work has also reported that the way virtual targets and tools are visualized can significantly influence perceived depth and task usability, indicating that rendering design is not merely a cosmetic issue but a perceptual factor that can directly affect operational accuracy [3]. These findings suggest that depth estimation errors in OST-AR are not caused by a single mechanism, but emerge from the combined influence of display optics, visual cue conflict, and rendering configuration.

The perceptual difficulty is closely related to the fact that OST-AR does not reproduce natural viewing conditions in a fully consistent manner. Many studies have linked depth underestimation in head-mounted AR systems to vergence–accommodation mismatch, insufficient focal support, imperfect registration, and conflicts between virtual and real depth cues [4,5]. Under these conditions, the observer may receive binocular disparity information that is only partially consistent with the physical environment, while accommodation remains constrained by the optical characteristics of the display. This mismatch can weaken confidence in metric distance judgments and promote systematic compression of perceived depth. Although binocular disparity and focal conflict have been widely discussed in the AR perception literature, the contribution of target scale has received comparatively less direct attention in OST-specific settings [6,7]. This omission is important, because any rendering parameter that changes the apparent size of a virtual object may also alter the balance among available depth cues and thereby shift the observer's spatial judgment [8]. Human depth perception depends on the integration of multiple sources of information, including binocular disparity, motion parallax, shading, texture gradients, occlusion, and familiar size [9]. In natural vision, retinal image size provides a meaningful cue when the observer has prior knowledge, either explicit or implicit, about the likely physical dimensions of an objectb [10,11]. If an object appears smaller or larger than expected, the visual system may reinterpret its distance accordingly. In mixed and augmented environments, however, virtual objects do not possess fixed physical dimensions in the same way as real objects. Their apparent scale is determined by rendering settings, display geometry, and scene composition rather than by a stable material presence in the world [12]. As a result, the size cue associated with a virtual target can be manipulated independently of physical distance, creating a condition in which scale may either support or distort depth interpretation [13,14]. Recent studies in immersive and mixed environments indicate that rendered scale can influence spatial judgments, particularly when other cues are weak, noisy, or inconsistent [15,16]. Under such circumstances, observers may rely more strongly on size-related information, which can introduce systematic bias in absolute distance estimation. This issue is especially relevant in OST-AR because virtual content is perceived together with visible real-world references. Unlike fully immersive virtual reality, where the entire scene is computer generated, OST displays present virtual objects against a real background that remains continuously available to the user. This configuration may reduce some forms of spatial ambiguity, yet it can also produce new perceptual inconsistencies. Real objects in the scene can serve as anchors for scale and distance, but they may also compete with virtual information when contrast, luminance, edge sharpness, or occlusion relationships are not well matched [17,18]. Optical combiners and display waveguides can further modify the visual appearance of virtual content by altering brightness, transparency, contour clarity, and apparent solidity. These factors may affect how object boundaries are interpreted and how strongly size information contributes to perceived depth. Consequently, findings obtained in immersive VR cannot be directly generalized to OST-AR, because the observer in OST conditions is simultaneously exposed to physical references and display-mediated virtual cues that do not always obey the same perceptual rules. Evidence from immersive virtual environments suggests that smaller rendered objects are often perceived as closer than intended, whereas larger rendered objects may appear farther away [19,20]. Even so, the transferability of this pattern to OST-AR remains uncertain. In a see-through configuration, the real scene remains visible at all times, and the perceptual system may weight environmental references differently from the way it does in a fully synthetic scene. The effect of target scale may therefore be attenuated, amplified, or qualitatively altered depending on the degree of cue consistency between the virtual content and the surrounding real environment. In addition, OST tasks often involve practical workspaces with constrained distances and action-oriented judgments rather than free exploration of a virtual scene. In such contexts, small errors in

perceived distance can directly affect reaching accuracy, alignment quality, procedural safety, and task efficiency. For this reason, the role of virtual object scale in OST depth estimation deserves more direct and systematic examination than it has received so far. Existing studies on mixed-reality perceptual calibration have explored disparity correction, cue reweighting, and other compensation strategies intended to improve spatial accuracy [21,22]. These studies have advanced understanding of how users adapt to display-induced distortions, but they do not fully resolve the question of how target scale itself influences absolute distance judgments in OST systems. Much of the available literature is still limited by relatively small participant samples, narrow distance ranges, or comparisons among only a few discrete display conditions [23,24]. In many cases, the outcome measure focuses on relative depth discrimination rather than absolute metric estimation. That distinction matters in application domains such as surgery, assembly, and maintenance, where users must often place a tool, align an overlay, or verify a position at a specific physical distance rather than merely decide which of two targets appears nearer. The interaction between physical viewing distance and virtual scale is therefore not only a theoretical perceptual question, but also a practical design issue for OST interfaces intended to support precision tasks.

A clearer characterization of this relationship is needed to determine whether scale mismatch contributes to the depth compression commonly observed in OST-AR and to identify rendering conditions that reduce spatial bias. The present study addresses this need by systematically examining perceived distance under reduced, matched, and enlarged virtual target scales across multiple physical viewing distances in an OST environment. Absolute distance estimation is used as the primary outcome so that the results directly reflect metric spatial judgment rather than ordinal discrimination. Regression-based analysis is further employed to quantify how perceived distance changes with target scale and to determine whether the resulting bias follows a consistent trend across the tested workspace. By clarifying the role of virtual size calibration in OST depth perception, this study aims to provide quantitative evidence that can support perceptually informed rendering design and improve spatial reliability in mixed-reality applications that depend on accurate hand–eye coordination and near-field alignment.

2.Materials and Methods

2.1 Participants and Experimental Setting

Twenty-eight adults (14 females and 14 males; age range 21–33 years, mean ± SD: 26.4 ± 3.1 years) participated in the study. All had normal or corrected-to-normal visual acuity and stereoacuity better than 60 arcsec, confirmed using a standard stereo test. Individuals with binocular vision disorders, uncorrected refractive errors, or known eye disease were excluded. The experiment was carried out in a laboratory with stable ambient lighting between 280 and 320 lux. An optical see-through head-mounted display with a fixed focal distance of 2.0 m was used. Virtual targets were presented against a neutral real background without strong depth cues to limit contextual influence. Physical reference distances were set at 0.3, 0.6, 0.9, and 1.2 m from the participant's eye position, covering common near and mid-range interaction distances.

2.2 Experimental Design and Control Conditions

A repeated-measures design was applied to reduce the effect of individual differences. Each participant completed trials under three size conditions: reduced scale (0.5× physical equivalent), matched scale (1×), and enlarged scale (1.5×). The matched condition served as the baseline because it preserved geometric consistency between rendered and intended size. The reduced and enlarged conditions were included to test how deviations from correct scaling affect perceived distance. Target distances were randomized within each block. The order of size conditions was counterbalanced across participants. Each condition included 20 trials at each distance level to ensure balanced data for statistical analysis. Short breaks were provided between blocks to reduce visual fatigue.

2.3 Measurement Procedures and Quality Control

Participants verbally reported the perceived distance of each virtual target in centimeters. No physical measuring tool or visual reference scale was provided during testing to avoid response bias. Reported values were recorded digitally and paired with the actual rendered distance. Before data collection, interpupillary distance was calibrated individually to ensure accurate stereo presentation. Spatial alignment was checked using a calibration grid displayed in the central field of view. Trials with tracking errors or response times exceeding three standard deviations from the participant's mean were excluded. A brief practice session was conducted before formal testing to familiarize participants with the procedure without emphasizing any specific size condition.

2.4 Data Processing and Model Formulation

For each trial, relative depth bias was calculated as

$$\text{Bias}=\frac{D_p-D_a}{D_a}\times 100\%,$$

where $D_p$represents perceived distance and $D_a$represents actual rendered distance. Mean bias values were computed for each participant, distance, and size condition. To examine the effect of scaling on perceived depth, a linear regression model was used:

$$D_p=\alpha+\beta_1 D_a+\beta_2 S+\varepsilon,$$

where S denotes the size condition (coded as −1 for 0.5×, 0 for 1×, and 1 for 1.5×), α is the intercept, β_1 reflects sensitivity to actual distance, β_2 represents the effect of scaling, and ε is the residual term. Residual plots were examined to check distribution and variance consistency. Statistical significance was determined at $\alpha = 0.05$. Data analysis was performed using R (version 4.3).

3.Results and Discussion

3.1 Size scaling produced systematic shifts in distance estimates

Virtual target size scaling produced consistent changes in reported distance across the 0.3–1.2 m range. The 0.5× condition led to clear depth compression, with a mean underestimation of 12.6%. Size-matched rendering (1×) reduced this bias to 4.3%, indicating that correct geometric scale supports more accurate metric judgment in OST-AR. In contrast, the 1.5× condition shifted responses toward overestimation, and the effect became stronger at longer distances. These results show that scale manipulation changes the direction and magnitude of depth error rather than simply increasing response variability. Similar sensitivity to object-related visual factors has been observed in controlled VR perception studies, where changes in object representation influence spatial interpretation [25,26]. Fig. 1. VR experimental setup used for assessing object perception under controlled visuomotor interaction conditions.

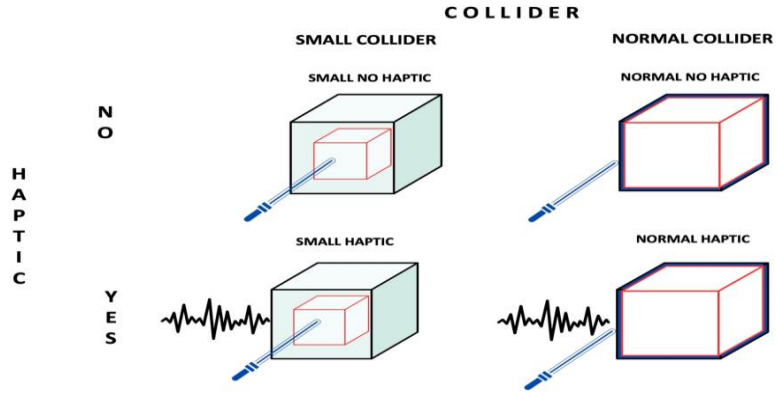


Fig.1.**VR setup showing controlled visuomotor conditions used to evaluate object perception.**

3.2 Regression analysis revealed changes in depth-response gain

Perceived distance was strongly correlated with rendered distance in all scale conditions; however, regression parameters differed. Under 0.5× scaling, the slope of the fitted relationship was lower, indicating compression across the full range of distances. The 1× condition produced a slope closest to unity and showed minimal intercept shift, consistent with reduced systematic bias. The 1.5× condition increased the slope and shifted the response upward, which explains the observed overestimation at greater depths. These findings indicate that scale scaling alters the gain of distance mapping rather than introducing a constant offset [27,28]. Therefore, a single correction value would not remove bias across all distances in the workspace.

3.3 Interpretation under OST display constraints

The pattern of results is consistent with cue weighting in optical see-through viewing. In OST displays, accommodation remains fixed at the display focal setting, while other depth cues may be weak or partially inconsistent. Under these conditions, retinal size can contribute more strongly to the final distance estimate. A smaller-than-expected target may reinforce compression already present in the system, whereas an enlarged target may counteract compression but can also produce overestimation [29]. In addition, optical geometry and parallax-related misalignment influence the stability of the perceived virtual–real relationship. Fig. 2 illustrates how selecting a display focus plane near the working distance can reduce parallax-related registration error, which is closely related to stable spatial perception in near-field tasks.

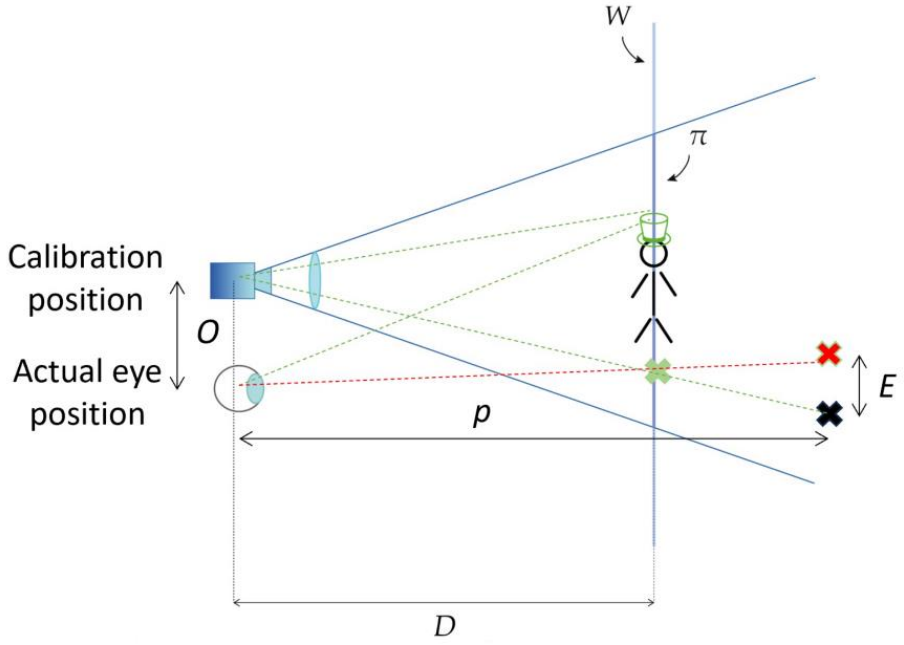


Fig.2.**Optical layout demonstrating how focus-plane placement reduces parallax-related misalignment in OST head-mounted displays.**

3.4 Comparison with related studies and design implications

Previous research in immersive and mixed environments has shown that object scale influences perceived space. The present findings confirm that the same effect appears in OST-AR during absolute distance estimation. In OST systems, real background cues and optical limits further shape how size information is interpreted. The results indicate that correct geometric scaling should be considered a default requirement for applications that rely on accurate metric placement. Any intentional size adjustment for readability or emphasis should be tested against depth bias across the relevant working distances. The study was conducted with simple targets and a controlled background, and the exposure period was short. More complex scenes, familiar objects, and extended use may change cue weighting and adaptation patterns. Future research should

examine varied target shapes, richer environments, and user-specific calibration limits to improve depth accuracy while maintaining stable perception.

4.Conclusion

This study investigated the effect of virtual target size scaling on depth estimation bias in optical see-through augmented reality. The results indicate that reduced scale produces consistent depth compression, matched scale reduces systematic error, and enlarged scale shifts perception toward overestimation, especially at longer viewing distances. Regression analysis shows that size scaling changes the gain of the distance–response relationship rather than adding a constant offset. These findings demonstrate that object size is a critical factor in metric depth judgment in OST-AR. From a scientific perspective, the data confirm that monocular size cues interact with optical constraints and fixed focal settings to influence spatial perception. From an application perspective, accurate geometric scaling should be considered a basic requirement for tasks that depend on precise distance interpretation. If size adjustments are introduced to improve visibility or emphasis, their impact on depth bias should be tested across the intended working range. The study was conducted in a controlled setting with simple targets and a limited depth interval, and long-term adaptation was not examined. Future research should assess performance in more complex scenes, across wider distance ranges, and during extended use to determine stable calibration strategies for practical deployment.

References

Zhang, Y., Gu, W., & Wang, J. (2025). Research on First Article Inspection (FAI)-Driven Quality Assurance Methods for Wind Turbine Installation and Operation & Maintenance and Their Effect on Reliability Improvement. Available at SSRN 6094206.

Kelly, J. W. (2022). Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics. IEEE transactions on visualization and computer graphics, 29(12), 4978-4989.

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.

Lee, S., & Hua, H. (2024). Effects of focal distance on near-field depth perception and accommodative response in a VariFocal optical see-through augmented reality display. IEEE Transactions on Visualization and Computer Graphics.

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.

Yoo, S. A., Lee, S., & Joo, S. J. (2023). Monocular cues are superior to binocular cues for size perception when they are in conflict in virtual reality. Cortex, 166, 80-90.

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.

Dorado, J. (2022). A framework to characterize and influence spatial perception in heterogeneous virtual reality displays (Doctoral dissertation, HESAM Université; Universidad de los Andes (Bogotá, Colombia; 1948-....)).

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

Bracci, S., & Op de Beeck, H. P. (2023). Understanding human object vision: a picture is worth a thousand representations. Annual review of psychology, 74(1), 113-135.

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

Andrews, S. (2024). The depth object exploring non-naturalistic spatial composition in virtual reality (Doctoral dissertation, RMIT 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.

Au, D., Allison, R. S., & Wilcox, L. M. (2026). Depth Cue Conflicts Between Real and Virtual Objects Disrupt Visually-Guided Reaching in Augmented Reality. IEEE Transactions on Visualization and Computer Graphics.

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.

Salehi, F., Pariafsai, F., & Dixit, M. K. (2024, June). The impact of different levels of spatial cues on size perception: a spatial perception study of altered conditions. In International Conference on Human-Computer Interaction (pp. 74-85). Cham: Springer Nature Switzerland.

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).

Franz, R. L., Junuzovic, S., & Mott, M. (2024). A virtual reality scene taxonomy: Identifying and designing accessible scene-viewing techniques. ACM Transactions on Computer-Human Interaction, 31(2), 1-44.

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.

Rzepka, A. M., Hussey, K. J., Maltz, M. V., Babin, K., Wilcox, L. M., & Culham, J. C. (2023). Familiar size affects perception differently in virtual reality and the real world. Philosophical Transactions of the Royal Society B: Biological Sciences, 378(1869).

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

Jiménez-Navarro, D., Groth, C., Peng, X., Pina, J., Sun, Q., Chakravarthula, P., ... & Serrano, A. (2026). Overdriving Visual Depth Perception via Sound Modulation in VR. IEEE Transactions on Visualization and Computer Graphics.

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.

Hoogerbrugge, A. J., Hooge, I. T., Hessels, R. S., & Strauch, C. (2025). When is enough enough? Empirical guidelines to determine participant sample size for scene viewing studies. Behavior Research Methods, 57(9), 241.

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

Unger, K., Kacin, M., & Abdel Rahman, R. (2025). Evolving object concepts in the adult brain: An electrophysiological investigation. Cognitive, Affective, & Behavioral Neuroscience, 25(4), 1040-1062.

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.

Lachmuth, S., Capblancq, T., Prakash, A., Keller, S. R., & Fitzpatrick, M. C. (2024). Novel genomic offset metrics integrate local adaptation into habitat suitability forecasts and inform assisted migration. Ecological Monographs, 94(1), e1593.

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.