This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

An Anatomically Contextualized Mixed Reality Testbed for Evaluating Neurovascular Depth Visualization

Zoé Jones[1], Simon Drouin[2], and Marta Kersten-Oertel[1]

[1] Concordia University, Montréal, Canada
za_jones@live.concordia.ca
[2] École de technologie supérieure, Montréal, Canada

Abstract. Depth-enhancement techniques can improve perception of isolated vascular renderings, but their benefit in anatomically contextualized mixed reality (MR) remains unclear. We present an MR visualization that embeds patient-specific cerebral vasculature within surrounding brain anatomy and evaluate three established depth-enhancement techniques: Phong shading, aerial perspective (Fog), and pseudo-chromadepth. In a within-subject study with 12 participants, users completed vessel depth-selection and surface-placement tasks while objective performance and subjective workload were recorded. No technique significantly improved selection accuracy, placement accuracy, or task completion time, although pseudo-chromadepth was preferred by participants. These results suggest that anatomical context combined with stereoscopic viewing and natural motion parallax may reduce the contribution of additional depth cues. While our exploratory design cannot isolate these factors, the findings highlight the importance of evaluating surgical visualization techniques in perceptually realistic MR environments rather than isolated vascular renderings.

Keywords: Mixed Reality · Medical Imaging · Computer Graphics

1 Introduction

Neurovascular procedures such as aneurysm clipping, arteriovenous malformation (AVM) resection, and the treatment of arteriovenous fistulae (AVFs) require surgeons to reason about the complex three-dimensional relationships between cerebral vessels and the surrounding brain anatomy [14]. Unlike many other image-guided procedures, vessels are rarely viewed in isolation. Instead, they are partially occluded by surrounding tissue and visualized through narrow surgical corridors, making accurate depth perception important for surgical planning and intraoperative navigation [10].

To improve the perception of vascular anatomy, numerous visualization techniques have been proposed over the past two decades. Classical approaches such as aerial perspective (fog), chromadepth colour mapping, illustrative rendering, and enhanced surface shading have consistently been shown to improve depth

judgements of vascular datasets displayed on conventional desktop systems [15, 9, 3]. Building on these approaches, Drouin et al. [5] demonstrated that interaction-driven visualization, in which depth cues dynamically adapt relative to a tracked instrument, significantly improved local depth perception while preserving users' global understanding of angiographic volumes. More recently, these visualization techniques have been investigated in immersive virtual reality (VR), where stereoscopic viewing and motion parallax provide strong natural depth cues [8]. Several studies have reported that once these natural cues are available, differences between visualization techniques become substantially smaller or disappear entirely, suggesting a ceiling effect in immersive environments [7, 17, 12].

An important limitation of many previous works is that neurovascular perception has been studied largely outside its anatomical context. Evaluations typically present isolated vascular trees or angiographic volumes against simple backgrounds, deliberately removing the surrounding brain anatomy to isolate visualization effects [9, 7, 17]. While appropriate for understanding individual depth cues, these experimental paradigms differ substantially from how surgeons perceive anatomy during interventions, where vessels are interpreted together with surrounding tissue while continuously exploring the scene through natural head movement.

Recent mixed reality (MR) head-mounted displays (HMDs) employing video passthrough provide stereoscopic viewing, natural motion parallax, and registration of virtual anatomy within the user's physical workspace, making them a promising platform for surgical visualization. At the same time, video passthrough introduces its own perceptual challenges, including reduced image quality, imperfect depth perception, limited occlusion handling, and competition between virtual anatomy and the real environment [1, 13, 4]. Although MR has received increasing attention for IGS and neurovascular visualization [14, 11], it remains unclear whether visualization techniques developed for isolated vascular rendering continue to provide measurable benefit once vessels are viewed within a richer anatomical context.

In this work, we introduce an anatomically contextualized MR visualization platform that embeds patient-specific cerebral vasculature within surrounding brain anatomy rather than presenting the vasculature in isolation. Although not intended to reproduce the exact intraoperative view, the platform more closely reflects the spatial reasoning required during neurovascular procedures while maintaining experimental control over visualization conditions, providing a controlled yet perceptually richer testbed for evaluating neurovascular visualization techniques in MR. Using this platform, we compare three established depth-enhancement techniques, Phong shading, aerial perspective (Fog), and pseudo-chromadepth, during a depth-selection and placement task. Objective performance is evaluated using selection accuracy, depth and lateral placement error, task completion time, and head movement, complemented by subjective workload and user preference measures.

2 User Study

2.1 MR Visualization Environment

A standalone mixed reality (MR) application was developed in Unity for the Meta Quest 3. The application displays a patient-specific cerebral vascular model embedded within surrounding anatomical brain tissue, providing anatomical context while maintaining full experimental control over the visualization. Unlike previous studies that present isolated vascular structures, participants viewed the vasculature within the surrounding anatomy while freely exploring the scene using natural head movement.

The cerebral vasculature was reconstructed from anonymized time-of-flight magnetic resonance angiography (TOF-MRA) data and rigidly registered to an anatomical brain model based on the ICBM-152 template. Because the vasculature is patient-specific while the surrounding brain is a population template, the environment provides a simplified, template-based anatomical context rather than a fully patient-specific or intraoperative representation. Both models were rendered at anatomical scale and rigidly anchored to a fixed location in the user's physical workspace. Participants interacted with the visualization using the Quest controller, represented as a virtual probe. A localized clipping window centred on the probe tip dynamically revealed the embedded vasculature while preserving the surrounding anatomy (Fig. 1).

2.2 Task

Participants completed a depth perception task under each visualization condition. Three candidate target points were displayed at different depths within the vascular tree, and participants identified the point located at the intermediate depth using the virtual probe. They then placed a virtual fiducial on the outer brain surface directly above the selected vessel while avoiding excessive penetration into the model.

Each visualization condition consisted of five depth-selection trials followed by five surface-placement trials. A sixth trial set was administered but excluded from analysis, yielding 180 trials per task across the 12 participants and three conditions ($12 \times 3 \times 5$). Participants were free to naturally move their head and body throughout the experiment, allowing unrestricted use of stereoscopic viewing and motion parallax while exploring the anatomy.

2.3 Depth Visualization Techniques

Three commonly studied depth-enhancement visualization techniques were evaluated in the study (see Fig. 1):

- **Phong shading:** Standard diffuse and specular surface shading served as the baseline condition.

- **Fog:** Vessel colour and opacity gradually attenuated with increasing distance from the probe tip, providing a proximity-based depth cue through aerial perspective.
- **Pseudo-chromadepth:** Vessel surfaces were coloured using a red-to-blue colour gradient according to their distance from the probe tip, adapting the chromadepth concept to the anatomically contextualized visualization.

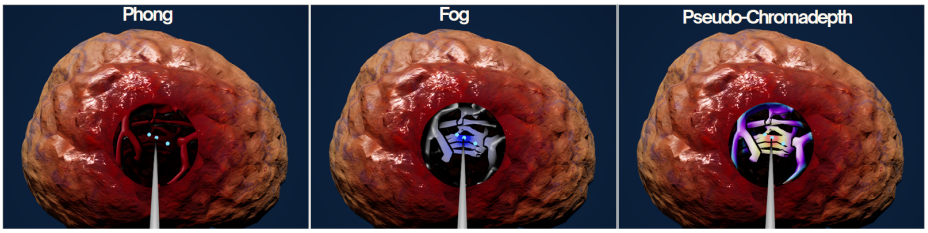


Fig. 1. The three depth-enhancement visualizations (Phong, Fog, and pseudo-chromadepth) with candidate depth points in cyan.

For both Fog and pseudo-chromadepth, depth was encoded relative to the probe tip, allowing the visualization to update continuously as participants explored the anatomy. All other rendering parameters remained identical across conditions.

2.4 Study Design

The study (approved by the Concordia REB) employed a within-subject design in which each participant completed the task under all three visualization conditions. The order of the visualization conditions was counterbalanced across participants to reduce ordering effects. Prior to data collection, participants completed a brief familiarization session to become comfortable with the MR environment and interaction technique.

2.5 Outcome Measures

Objective measures included target selection accuracy, selection time, surface-placement time, fiducial placement error, and head movement throughout each trial. Placement error was decomposed into components parallel and perpendicular to the viewing direction to separately quantify depth and lateral accuracy. Following each visualization condition, participants completed the NASA Task Load Index (NASA-TLX) questionnaire [6] and rated the perceived usefulness and overall preference of each visualization technique. Each NASA-TLX subscale was rated on a raw, unweighted 7-point ordinal scale; we report the per-subscale raw ratings rather than the weighted 0–100 score of the original instrument, as no pairwise weighting procedure was administered.

3 Results

3.1 Participants

Participants were recruited from a non-clinical population because the task evaluates fundamental depth perception, and previous work reported similar depth cue rankings between novices and experts [9]. Twelve participants (9 female, 3 male) completed the study, while two participants withdrew due to motion sickness and were excluded from the analysis. Five participants were graduate students, one was a medical professional, and the remaining participants represented a range of professions. Seven participants reported using VR less than once per year, and eight reported little or no familiarity with neuroanatomy.

3.2 Objective Performance

Table 1 summarizes the objective performance measures across the three visualization techniques. No significant differences were observed in selection accuracy, unsigned or signed depth error, lateral error, or task completion time. Likewise, no pairwise comparison survived Holm correction, indicating that none of the evaluated visualization techniques measurably improved objective depth perception within the anatomically contextualized MR environment.

Table 1. Objective outcomes (per-participant means, $n = 12$, 180 trials). Errors are reported in cm and time in seconds. Depth and lateral errors include only correct target selections. No significant differences were observed.

Outcome	Phong	Fog	Chromadepth	Friedman
Selection accuracy	0.867	0.850	0.900	$\chi^2 = 1.53$, $p = 0.466$
Depth error, unsigned	0.315	0.368	0.322	$\chi^2 = 0.50$, $p = 0.779$
Lateral error	0.204	0.228	0.261	$\chi^2 = 2.00$, $p = 0.368$
Depth error, signed	−0.069	−0.167	−0.036	$\chi^2 = 1.17$, $p = 0.558$
Time to place (s)	6.85	8.98	7.71	$\chi^2 = 1.17$, $p = 0.558$
Time to select (s)	21.57	20.13	26.57	$\chi^2 = 2.67$, $p = 0.264$

A null result is only informative if the task itself requires participants to make meaningful depth judgements. To verify this, we decomposed each placement error into a view-aligned orthonormal coordinate frame and compared the RMS error along the viewing direction with the two lateral directions. This analysis requires no distributional assumptions.

Error along the viewing axis was 0.429 cm compared with a per-axis lateral RMS error of 0.193 cm, corresponding to a ratio of 2.37 (median 2.41). The effect was consistent across all participants: every participant exhibited greater depth than lateral error (Wilcoxon signed-rank $W = 0$, $p = 0.00049$). These findings indicate that depth judgement was approximately 2.4× more difficult

than lateral positioning, consistent with previous 3D pointing studies [16, 2]. The task therefore contained a substantial depth-specific component that could have benefited from improved depth visualization. Signed depth error was slightly negative across all visualization techniques (grand mean -0.088 cm), consistent with the depth compression commonly reported for HMDs, although the bias was not significantly different from zero ($t = -1.45$, $p = 0.174$). Time usage was fairly consistent across participants, who took considerably longer to select the correct point (mean 22.75 s) than to place the fiducial marker (mean 7.84 s). This suggests that participants often struggled to judge the correct depth initially; once the depth of the vasculature became apparent, placing the fiducial marker was straightforward.

3.3 Head Movement and Viewing Behaviour

As the availability of motion parallax depends on how much participants actually moved, we logged the head pose throughout every trial. Participants made extensive use of natural viewing behaviour rather than adopting a single fixed viewpoint: the median head-position range during a session was 2.08× the full spatial extent of the vascular model, meaning that participants translated their viewpoint across a volume roughly twice the size of the anatomy under inspection. This exploration was sustained over time, with individual selection trials lasting up to 235 s while participants repositioned to disambiguate depth. Because participants freely selected and continuously changed their viewpoints, the effective depth separation between the three candidate points varied from 0.2 to 18.9 cm across trials (median 4.0 cm). The magnitude of this movement indicates that participants continuously exploited stereoscopic viewing and motion parallax while interpreting the vasculature. This behavioural evidence supports our interpretation that strong natural depth cues were available in every condition, offering a plausible mechanism for why the additional artificial depth cues conferred no measurable objective benefit.

3.4 Subjective Evaluation

Although objective performance did not differ between visualization techniques, subjective responses showed a clear preference for pseudo-chromadepth. Of the ten participants who selected a single preferred visualization, seven chose pseudo-chromadepth and three chose Phong, while none selected Fog (binomial test against 1/3 chance, $p = 0.020$). Including the two participants who selected multiple techniques, pseudo-chromadepth was chosen by nine of the twelve participants. This result should be interpreted cautiously. The response options were modified during data collection, requiring one visualization condition to be merged across two labels, and participants were permitted to select multiple preferred techniques. We therefore consider this finding suggestive rather than confirmatory.

NASA-TLX responses showed a similar trend (Table 2). Phong received the highest (worst) ratings on all six workload subscales, with Effort approaching

Table 2. NASA-TLX subscale scores (per-participant means, $n = 12$; higher scores indicate greater workload). Phong received the highest workload ratings on all six subscales, although no comparison reached statistical significance.

Subscale	Phong	Fog	Chromadepth	Friedman
Mental Demand	4.33	3.50	3.42	$\chi^2 = 1.37$, $p = 0.505$
Physical Demand	3.75	3.33	3.58	$\chi^2 = 2.70$, $p = 0.260$
Temporal Demand	2.25	1.92	1.50	$\chi^2 = 3.56$, $p = 0.168$
Performance	3.83	3.42	3.50	$\chi^2 = 1.15$, $p = 0.562$
Effort	4.17	3.50	3.42	$\chi^2 = 5.52$, $p = 0.063$
Frustration	3.83	2.75	2.75	$\chi^2 = 3.71$, $p = 0.156$

significance ($p = 0.063$). Although none of the workload measures reached statistical significance, the consistent direction across all subscales suggests that participants perceived pseudo-chromadepth and Fog as requiring slightly less effort than the baseline visualization.

4 Discussion

The primary finding of this study is that established depth-enhancement visualization techniques did not measurably improve objective depth perception when cerebral vasculature was viewed within an anatomically contextualized MR environment. Previous work has consistently demonstrated benefits of chromadepth, aerial perspective, and interaction-driven visualization for isolated vascular renderings viewed on conventional displays [15, 9, 3, 5]. In contrast, none of the evaluated visualization techniques significantly affected selection accuracy, placement accuracy, or task completion time in our study. Rather than contradicting previous findings, we believe these results demonstrate that the effectiveness of depth-enhancement techniques depends strongly on the perceptual context in which they are evaluated.

A key difference between our work and previous studies is that participants viewed cerebral vessels embedded within surrounding brain anatomy rather than as isolated vascular trees or angiographic volumes. In addition, participants were free to naturally explore the scene using stereoscopic viewing and unrestricted head movement, allowing continuous motion parallax while interpreting vessels relative to neighbouring anatomical structures. Together, these cues likely reduced reliance on artificial depth enhancements. This interpretation agrees with previous immersive VR studies showing that the benefits of depth-enhancement techniques diminish once binocular disparity and motion parallax become available [7, 17, 12]. Our findings suggest that this ceiling effect extends beyond immersive VR to anatomically contextualized MR visualization, where anatomical context itself becomes an additional source of spatial information.

Beyond the findings for the three visualization techniques evaluated here, the proposed MR platform provides a controlled evaluation environment for studying

future neurovascular visualization methods under conditions that more closely resemble surgical perception than conventional desktop experiments. Such a testbed may help bridge the gap between visualization techniques developed in isolation and their eventual application within computer-assisted surgical systems. Importantly, this does not imply that depth visualization techniques are no longer useful. Rather, it suggests that visualization methods developed using simplified experimental paradigms should be re-evaluated in environments that more closely resemble how surgeons actually perceive anatomy. Much of the vascular visualization literature intentionally removes surrounding anatomy to isolate perceptual effects. While this approach has produced valuable insights into individual visualization techniques, our results indicate that conclusions drawn from isolated vascular renderings may not directly translate to clinically relevant MR visualization systems, where surgeons simultaneously exploit anatomical context, stereopsis, and motion parallax.

Interestingly, subjective responses diverged from objective performance. Participants consistently preferred pseudo-chromadepth, and NASA-TLX scores showed a similar trend toward reduced perceived workload despite no measurable improvement in task performance. This partially contrasts with interaction-driven visualization work, where chromadepth both improved performance and was preferred by users [5], again suggesting that the perceptual context substantially influences the effectiveness of depth-enhancement techniques. Our findings indicate that visualization techniques may still improve confidence or perceived ease of interpretation even when objective performance is unchanged. For surgical visualization systems, this distinction is important: techniques that improve user experience may still be valuable, but they should not necessarily be interpreted as improving perceptual accuracy.

Several limitations should be considered when interpreting these findings. First, we evaluated three representative depth-enhancement visualization techniques adapted for interactive MR exploration. Other visualization strategies, including alternative implementations of chromadepth, illustrative rendering, or interaction-driven techniques, may exhibit different behaviour in anatomically contextualized MR environments. Consequently, our conclusions apply specifically to proximity-based depth encodings and do not exclude potential benefits of alternative visualization strategies in MR. Second, participants were allowed unrestricted viewpoint changes and unlimited exploration time. Although this reflects natural interaction with MR systems, it also maximized the availability of motion parallax and may therefore represent a best-case scenario for natural depth perception. Third, because anatomical context, stereoscopic viewing, and motion parallax were present simultaneously in every condition, our design cannot attribute the absence of a visualization effect to any individual factor. Isolating their contributions would require a factorial design that varies anatomical context and viewing mode independently; we therefore present the roles of context, stereopsis, and parallax as plausible explanations rather than demonstrated causes. Finally, this exploratory study included a relatively small sample of participants, and we did not conduct an a priori power or equivalence anal-

ysis. The non-significant results should accordingly be read as an absence of a detectable difference in this sample rather than as evidence that the techniques are equivalent. While the depth-specific error analysis demonstrated that the task itself was sufficiently sensitive to reveal performance differences, future studies involving clinicians, more realistic surgical tasks, and varying levels of anatomical context will be necessary to determine under which conditions depth-enhancement techniques provide measurable benefit.

5 Conclusion

We presented an anatomically contextualized mixed reality platform for evaluating neurovascular visualization techniques and used it to compare three established depth-enhancement methods. Although pseudo-chromadepth was consistently preferred by participants, none of the evaluated techniques significantly improved objective task performance. These findings suggest that anatomical context, stereoscopic viewing, and natural motion parallax may reduce the contribution of additional artificial depth cues compared with previous studies using isolated vascular renderings, though our exploratory design cannot separate their individual contributions. Beyond these findings, the proposed platform provides a controlled testbed for evaluating future MR visualization techniques within a perceptually realistic neurovascular context.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Adams, H., Stefanucci, J., Creem-Regehr, S., Bodenheimer, B.: Depth Perception in Augmented Reality: The Effects of Display, Shadow, and Position. pp. 792–801 (Mar 2022). https://doi.org/10.1109/VR51125.2022.00101
2. Batmaz, A.U., Mughrabi, M.H., Barrera Machuca, M.D., Stuerzlinger, W.: Effect of Stereo Deficiencies on Virtual Distal Pointing. In: Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (VRST). pp. 1–8. ACM (2022). https://doi.org/10.1145/3562939.3565621
3. Behrendt, B., Berg, P., Preim, B., Saalfeld, S.: Combining Pseudo Chroma Depth Enhancement and Parameter Mapping for Vascular Surface Models. In: Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM). The Eurographics Association (2017). https://doi.org/10.2312/vcbm.20171250
4. Do, T.D., LaViola, J.J., McMahan, R.P.: The Effects of Object Shape, Fidelity, Color, and Luminance on Depth Perception in Handheld Mobile Augmented Reality. In: 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 64–72 (Nov 2020). https://doi.org/10.1109/ISMAR50242.2020.00026
5. Drouin, S., Di Giovanni, D.A., Kersten-Oertel, M., Collins, D.L.: Interaction-driven enhancement of depth perception in angiographic volumes. IEEE Transactions on Visualization and Computer Graphics **25**(1), 927–937 (2019). https://doi.org/10.1109/TVCG.2018.2865245

6. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, Advances in Psychology, vol. 52, pp. 139–183. North-Holland (1988)
7. Heinrich, F., Apilla, V., Lawonn, K., Hansen, C., Preim, B., Meuschke, M.: Estimating depth information of vascular models: A comparative user study between a virtual reality and a desktop application. Computers & Graphics **98**, 210–217 (2021). https://doi.org/10.1016/j.cag.2021.05.014
8. Hombeck, J., Meuschke, M., Zyla, L., Heuser, A.J., Toader, J., Popp, F., Bruns, C., Hansen, C., Datta, R., Lawonn, K.: Evaluating Perceptional Tasks for Medicine: A Comparative User Study Between a Virtual Reality and a Desktop Application (Apr 2022). https://doi.org/10.1109/VR51125.2022.00071
9. Kersten-Oertel, M., Chen, S.J.S., Collins, D.L.: An Evaluation of Depth Enhancing Perceptual Cues for Vascular Volume Visualization in Neurosurgery. IEEE Transactions on Visualization and Computer Graphics **20**(3), 391–403 (Mar 2014). https://doi.org/10.1109/TVCG.2013.240
10. Kersten-Oertel, M., Gerard, I., Drouin, S., Mok, K., Sirhan, D., Sinclair, D., Collins, L.: Augmented reality in neurovascular surgery: feasibility and first uses in the operating room. International journal of computer assisted radiology and surgery **10** (Feb 2015). https://doi.org/10.1007/s11548-015-1163-8
11. Kersten-Oertel, M., Gerard, I.J., Drouin, S., Mok, K., Sirhan, D., Sinclair, D.S., Collins, D.L.: Augmented Reality for Specific Neurovascular Surgical Tasks. In: Linte, C.A., Yaniv, Z., Fallavollita, P. (eds.) Augmented Environments for Computer-Assisted Interventions. pp. 92–103. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-24601-7_10
12. Lesar, Ž., Bohak, C., Marolt, M.: Evaluation of depth perception in crowded volumes. Computers & Graphics **120**, 103918 (May 2024). https://doi.org/10.1016/j.cag.2024.103918
13. Macedo, M.C.F., Apolinario, A.L.: Occlusion Handling in Augmented Reality: Past, Present and Future. IEEE Transactions on Visualization and Computer Graphics **29**(2), 1590–1609 (Feb 2023). https://doi.org/10.1109/TVCG.2021.3117866
14. Preim, B., Baer, A., Cunningham, D., Isenberg, T., Ropinski, T.: A Survey of Perceptually Motivated 3D Visualization of Medical Image Data. Computer Graphics Forum **35**(3), 501–525 (Jun 2016). https://doi.org/10.1111/cgf.12927, https://onlinelibrary.wiley.com/doi/10.1111/cgf.12927
15. Ropinski, T., Steinicke, F., Hinrichs, K.: Visually Supporting Depth Perception in Angiography Imaging. In: Smart Graphics (SG 2006), Lecture Notes in Computer Science, vol. 4073, pp. 93–104. Springer (2006). https://doi.org/10.1007/11795018_9
16. Teather, R.J., Stuerzlinger, W.: Pointing at 3D targets in a stereo head-tracked virtual environment. In: 2011 IEEE Symposium on 3D User Interfaces (3DUI). pp. 87–94. IEEE (2011). https://doi.org/10.1109/3DUI.2011.5759222
17. Titov, A., Kersten-Oertel, M., Drouin, S.: The Effect of Interactive Cues on the Perception of Angiographic Volumes in Virtual Reality. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization **10**, 1–9 (Nov 2021). https://doi.org/10.1080/21681163.2021.1999332