Technical Challenges in Building Cross Reality Applications for Analyzing 3D Medical Images

Brody Wells* Nanjia Wang† Frank Maurer‡

University of Calgary

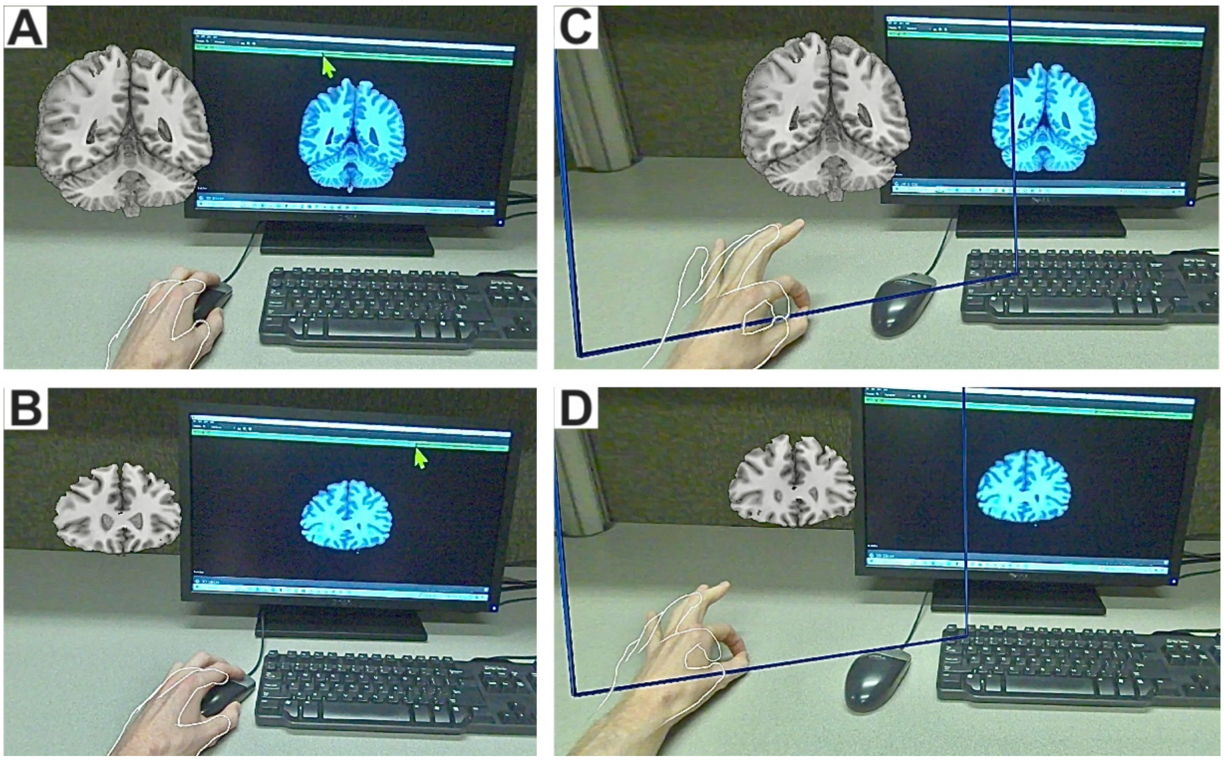

Figure 1: Series of images showing bi-directional cross reality (CR) interactions between our mixed reality (MR) system and the 3DSlicer application for analyzing medical images. Image A: The user is using the mouse (cursor is highlighted in green) on the desktop running 3DSlicer software. Image B: The user changed the cross-section view with the mouse, which is updated on both the desktop and MR systems. Image C: The user performs a ”pinch” gesture with hand interactions on the MR system. Image D: The user has changed the cross-section view with a ”pinch” gesture in MR, which is updated on both the desktop and MR systems.

Abstract

Current extended reality (XR) systems lack seamless interaction with established desktop applications for medical image analysis, resulting in inefficiencies. This gap highlights the need for XR design guidelines and cross reality (CR) frameworks to integrate these systems. Designing an effective XR system and implementing a CR framework would enable users to leverage XR's immersive capabilities alongside the precision and advanced tools of desktop software without the need to switch between systems. Addressing these challenges could streamline medical image analysis and inform the development of future CR applications.

Index Terms: Serious extended reality, Cross reality framework, Mixed reality, 3D visualization, Medical imaging, Human computer interaction.

*e-mail:brody.wells@ucalgary.ca
†e-mail:nanjia.wang1@ucalgary.ca
‡e-mail:frank.maurer@ucalgary.ca

1 Introduction

The integration of mixed reality (MR) with traditional medical desktop software, using cross reality (CR) technologies, offers the potential to improve workflow in medical image analysis. Effectively combining the intuitive and immersive capabilities of MR with the precision and established utility of traditional desktop tools, may allow medical professionals to perform more efficient analysis of medical imaging, in forms such as Computed Tomography (CT) or Magnetic Resonance Image (MRI) scans. This integration would preserve the benefits of traditional software while leveraging the advanced visualization and interaction possibilities of MR.

1.1 Why Integrate with Cross Reality

Extended reality (XR), which includes augmented reality (AR) and virtual reality (VR), along with MR, is being adopted by medical professionals, with most current interest focused on AR tools for surgery and medical training [4]. However, there has been less research and development focused on the analysis and evaluation of

2771-1110/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR-Adjunct64951.2024.00040

medical images, despite the significant potential and emerging interest in this area. Previously, most XR systems have been developed to operate on a single point of the Milgram reality-virtuality continuum (RVC) [3], primarily utilizing XR input modalities such as controllers or hand interactions. CR technologies would allow users to seamlessly transition between multiple points along the RVC, giving them access to more input modalities, such as a mouse and keyboard.

Traditional desktop software for analyzing medical images has been developed and refined over decades, with the assumption that users would interact using a mouse and keyboard. XR systems that abandon these input modalities may lose many of the benefits of traditional systems, requiring users to physically switch between systems to leverage the strengths of each. With CR, users can utilize both systems simultaneously, gaining the advantages of each without needing to switch back and forth, thus avoiding the inconveniences such as frequently donning and doffing a head-mounted display (HMD). This seamless integration would allow for a more cohesive and uninterrupted workflow, improving the overall efficiency of medical image analysis.

1.2 Research Questions and Challenges

Summarizing the challenges, we have formed the following research questions:

1. **Synchronization Frameworks**: How can established patterns such as Model-View-Controller (MVC) and publish-subscribe be utilized to effectively synchronize data between MR systems and traditional desktop software? What frameworks and protocols can be adopted or adapted to ensure seamless data exchange and consistency in single-user and multi-user collaborative environments?

2. **Interaction Modalities**: What are suitable interaction modalities for different tasks within this integrated system? When should hand interactions be prioritized over mouse and keyboard inputs, and vice versa?

3. **User Experience and Affordance**: How can we design an intuitive user interface that caters to non-technical medical users, facilitating a smooth transition between XR and traditional desktop tools?

2 Related Work

To enhance the effectiveness of surgery, the medical field has been developing VR and AR solutions for many decades, with the first use of AR in spinal surgery occurring in 1997 [4].

The application of XR technologies in medical image analysis and evaluations is less developed, but an emerging field. Mustafa et al. reported a significant increase in the number of radiologists who believed VR would play a significant role in radiology after testing a VR reading room developed by SieVRt, Luxsonic Technologies [11]. The study also found that radiologists felt the VR system provided adequate image resolution and contrast, underscoring VR's potential as a diagnostic tool for medical image analysis. Additionally, VR prototypes have recently been explored in neurology for clinical imaging evaluations, including RealityFlow [2] and SurgicalTheater systems [12].

Guerroudji et al. proposed an AR visualization and interaction system to assist clinicians in diagnosing breast cancer by visualizing tumor masses. Their method involved segmenting tumor masses using MATLAB [1], creating 3D models with 3DSlicer[2], and then utilizing their AR system to translate, rotate, and scale these visualizations [8]. This system suggested a linear workflow, where the clinician uses traditional desktop software for segmentation and 3D model creation, and then transitions to the AR system for visualization and interaction. Roessel et al. found immersive technologies can be useful for pre-operative planning and the actual operation, as MR can expand the vision and information available to the surgeon [14]. However they claim that the creation of interactable 3D organ with complex detail for immersive analysis remains a challenge.

Huang et al. proposed a CR interactive visualization system that bridges VR with a desktop display for cardiac surgery planning and education [9]. They identified the need for combining 2D desktop-based visualization with 3D VR visualization, as 3D heart models offer cardiologists more depth cues and intuitive manipulation. Meanwhile, 2D visualization applications are well-established and provide sufficient information for decision-making during the pre-surgery planning workflow. Aigner et al. explored a CR approach for cardiac evaluations by transitioning medical images from 2D displays into MR [1]. This approach allowed users to analyze the images in a 3D space by walking around and viewing them from different angles. Their pilot study concluded that traditional desktop software should not be replaced but rather integrated with XR solutions. Reinschluessel et al. proposed integrating MR HMDs with desktop visualization as a complementary interface in order to achieve optimal and time-efficient surgery planning to increase patient safety [13].

The integration of XR tools with traditional medical imaging software offers numerous benefits, yet there is currently no framework for such CR applications. Various frameworks have been developed for similar applications. For instance, WebXR supports accessing XR-capable devices, such as HMDs and smartphones, through a web browser [10]. Building upon WebXR, frameworks like ATON enhance Cultural Heritage experiences [7]. To facilitate synchronizing collaboration among users, ATON uses an API to subscribe clients to network events. Other frameworks have facilitated the development of diverse web-based CR applications, such as WiXaRd for multi-party collaboration in 3D spaces [16], and the collaborative interaction space proposed by Seo et al. [15], which uses deepstreaml.io for real-time synchronization between browsers via a publish-subscribe pattern [6]. For stationary single-user applications, Cool et al. proposed a framework which would expand the domain of interaction methods, such as enabling the mouse to track beyond the bounds of the desktop screen into XR space [5]. Although their proposed framework is not developed with medical image analysis in mind, it's novel approaches may have benefits in such use cases.

3 Design Challenges

3.1 Synchronization of XR and Traditional Desktop Software

- **Framework for Synchronization**: One of the primary challenges is developing an effective framework for synchronizing data between XR systems and traditional desktop software. Established patterns like Model-View-Controller (MVC) and publish-subscribe can be leveraged for this purpose. This also can involve evaluating methods such as web sockets combined with REST APIs to facilitate data exchange. A critical decision is whether each system should operate with its own client-server setup or utilize a central server to manage synchronization. The complexity increases in multi-user collaborative systems, where real-time synchronization across multiple devices is essential.

- **Existing Solutions and Frameworks**: Reviewing existing CR systems that manage synchronization can provide valuable insights into building generic frameworks to simplify

[1]https://www.mathworks.com/products/matlab.html?s_tid=hp_products_matlab
[2]https://www.slicer.org/

the integration of XR systems with traditional medical imaging software. Open-source medical imaging software like 3DSlicer, which uses a RESTful API running on a Python web server, can be enhanced for bi-directional CR interactions. By adding listeners to input actions, changes in state could be detected, and updates can be sent to an XR client system, whether via a publish-subscribe pattern or some other means. This can be achieved through a peer-to-peer (P2P) method for single-user applications, however implementing a central server would provide a framework to allow multiple devices and users can communicate. Discrepancies in how different systems structure their data could cause issues without a framework to properly handle communications. For instance, applications built with Unity use quaternions to measure rotations in virtual space, whereas 3DSlicer uses Euler angles.

3.2 Interaction Challenges

- **Smooth Interaction**: One significant technical challenge is running two systems simultaneously while ensuring smooth synchronization of model states, view states, and interactions. Determining when hand interactions are most effective and which tasks benefit from them is crucial. Conversely, identifying tasks where mouse and keyboard interactions are more suitable is equally important. Exploring novel interaction methods that combine hand gestures in XR with traditional input devices can lead to more intuitive and efficient workflows.

- **Affordance of UX/UI**: Enhancing the user interface to be intuitive for non-technical medical users is essential. The design should minimize the inconvenience of switching between XR controllers and traditional input devices. Ensuring a user-friendly interface that allows seamless interaction with both XR and traditional tools.

4 Traditional Desktop Software Challenges

4.1 2D vs. 3D Interfaces

Traditional medical imaging software conveys 3D data using 2D interfaces, which can complicate the conceptualization of 3D structures. While repositioning orthogonal cross-sections is quick and efficient, adjusting these cross-sections along oblique angles remains challenging and time-consuming. The integration of XR may help address such issues by providing an immersive 3D environment where data can be viewed in its original dimensional form and allowing users to manipulate the data using additional input modalities such as hand interactions or controllers.

4.2 Established Utility

Traditional desktop software has been developed over decades, providing proven accuracy and visualization quality. Medical professionals are highly trained and experienced in using these tools, making them an indispensable part of the medical imaging workflow. The challenge lies in augmenting these established tools with XR technologies without compromising their reliability and precision.

5 Methodology

To address the research questions outlined in this position paper we will focus on using qualitative research methodologies. For exploring how data can be effectively synchronized between XR systems and traditional desktop software, we will use knowledge obtained from related works regarding current CR frameworks to inform the design of our medical imaging CR framework. To determine suitable interaction modalities for different tasks as well as designing an effective user experience with high affordance, we propose performing heuristic evaluations as well as conducting demonstration sessions with medical experts. The heuristic evaluations are relatively quick and low cost, not requiring end-users to conduct, which can help speed up the prototyping process, as well as ensuring the user interface is being designed according to best practices. The demonstration sessions with medical experts would be audio or video recorded and participants would be asked to think aloud while using the system and its varying input modalities, gathering feedback on usability issues, cognitive or decision making processes, and to capture immediate reactions to gauge user satisfaction or frustration levels. This method can be paired with more structured interview questions relating to the user experience, intuitiveness of the system, and use case suggestions of how and when certain XR input modalities could be used to enhance their workflow. We will cross-validate the feedback results, comparing the demo session feedback with responses to interview questions to find consistent patterns and remove bias.

6 Prototyping

Part of our investigation involves the development of a prototype CR system. This prototype is being created through an iterative design process and serves as a case study to define core functionalities, test required features, and run experiments. The current stage of the prototype is controllerless and only uses hand tracking, with two main hand gestures used for XR interactions. A "grab" gesture is used for moving and rotating the 3D brain, allowing the user to reposition MR models to best use the space in their environment, as well as rotating the models to easily view from any perspective (see Figure 2). A "pinch" gesture is used to reposition the cross-sections. When repositioning cross-sections, the user can choose to use hand gestures in MR, or to use the mouse and keyboard. The views in the XR system and on the desktop software are updated simultaneously regardless of the input modality used to reposition a cross-section (see Figure 1).

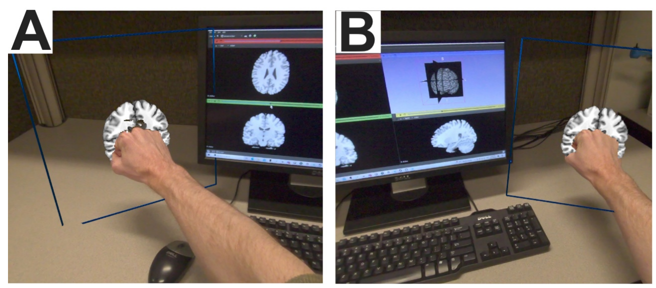

Figure 2: Series of images showing the grab interaction in MR. Image A: The user performs the grab gesture on the 3D brain model using their right hand. Image B: While continuously performing a grab gesture with their right hand, the user has relocated the brain from the left side of the computer display to the right side.

The system integrates 3DSlicer [3] with a modified version of its WebServer module to enable bi-directional CR synchronization, ensuring image views are consistent between desktop and XR systems. The XR system is being developed using Unity (2022.3 LTS)[4], utilizing the standard 3D core, the UltraLeap [5] package for hand tracking, which uses the inside-out tracking from the head-mounted display (HMD). We use the Varjo XR3 HMD [6], which has

[3] https://www.slicer.org/
[4] https://unity.com/
[5] https://www.ultraleap.com/tracking/
[6] https://varjo.com/products/varjo-xr-3/

relatively high resolution 12-megapixel video pass-through cameras. To convert NIfTI formatted medical images into 3D renderings in Unity we are using the VolumeViewerPro package from LISCINTEC [7] (deprecated from Unity asset store).

7 Discussion and Future Work

This paper outlines the potential and challenges of integrating MR with traditional desktop medical imaging software using CR. The integration aims to merge the immersive capabilities of MR with the precision and established utility of traditional tools, allowing both modalities to be used simultaneously. This will involve the investigation and development of synchronization frameworks, interaction modalities, and user interface designs. An extensive literature review will need to be conducted into the best practices of framework development, along with investigations of case studies.

Prototyping will involve an iterative design process to test and refine core functionalities, ensuring effective bi-directional synchronization between desktop and XR systems and informing the design of a CR framework. The insights gained from this prototyping phase will guide the development of a robust CR system that has potential to enhance the capabilities of medical professionals.

Future work will focus on investigating and developing synchronization frameworks, interaction modalities, and user interface designs Current literature will help inform the design of our CR framework. Heuristic evaluations and structured demo and feedback sessions will identify ideal input modalities for different use cases as well as develop an effective user experience and interface.

By systematically addressing these design challenges and leveraging advanced XR technologies, we aim to pave the way for future developments in CR applications within the medical field. This research holds significant promise for advancing medical imaging workflows.

Acknowledgments

This research was supported by a grant from Alberta Innovates and an NSERC Discovery Grant.

References

[1] D. Aigner, N. Wang, D. Kielmayer, J. Steiner, J. Hochpöchler, C. Heinzl, D. Roth, F. Maurer, and C. Anthes. Cardiac visualisation along the rv-continuum - a high-fidelity pilot study. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 675–680. IEEE, 2023. 2

[2] Z. Aminolroaya. *Streamlining the epilepsy pre-surgical evaluation workflow with virtual reality*. PhD thesis, University of Calgary, Calgary, Apr. 2023. 2

[3] J. Auda, U. Gruenefeld, S. Faltaous, S. Mayer, and S. Schneegass. A scoping survey on cross-reality systems. *ACM Computing Surveys*, 56(4):1–38, Apr. 2024. doi: 10.1145/3616536 2

[4] E. Barcali, E. Iadanza, L. Manetti, P. Francia, C. Nardi, and L. Bocchi. Augmented reality in surgery: A scoping review. *Applied Sciences*, 12(14):6890, Jan. 2022. Number: 14 Publisher: Multidisciplinary Digital Publishing Institute. doi: 10.3390/app12146890 1, 2

[5] R. Cools, M. Gottsacker, A. Simeone, G. Bruder, G. Welch, and S. Feiner. Towards a desktop-ar prototyping framework: Prototyping cross-reality between desktops and augmented reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 175–182, Oct. 2022. ISSN: 2771-1110. doi: 10.1109/ISMAR-Adjunct57072.2022.00040 2

[6] DeepStream. Deepstream — an open-source real-time stream processing engine. https://deepstream.io/. Accessed: July. 2024. 2

[7] B. Fanini, D. Ferdani, E. Demetrescu, S. Berto, and E. d'Annibale. ATON: An open-source framework for creating immersive, collaborative and liquid web-apps for cultural heritage. *Applied Sciences*, 11(22):11062, Nov. 2021. doi: 10.3390/app112211062 2

[8] M. A. Guerroudji, K. Amara, and N. Zenati. Augmented reality aid in diagnostic assistance for breast cancer detection. *Multimedia Tools and Applications*, Mar. 2024. doi: 10.1007/s11042-024-18979-2 2

[9] J. Huang, J. D. Plasencia, D. M. Bardo, N. C. Huber, E. G. Ellsworth, S. D. Zangwill, and C. Bryan. Phoenix virtual heart: A hybrid vr-desktop visualization system for cardiac surgery planning and education. In *2021 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 36–40, 2021. doi: 10.1109/VAHC53616.2021.00012 2

[10] B. Jones, M. Goregaokar, and R. Cabanier. Webxr device api. https://www.w3.org/TR/2024/CRD-webxr-20240701/, 2024. W3C Candidate Recommendation Draft, 1 Jul. 2024. 2

[11] A. R. Mustafa, F. Moloudi, E. Balasalle, M. Lang, and R. N. Uppot. Virtual reading room for diagnostic radiology. *Current Problems in Diagnostic Radiology*, p. S0363018823002001, Dec. 2023. doi: 10.1067/j.cpradiol.2023.12.002 2

[12] T. N. Phan, K. J. Prakash, R.-J. S. Elliott, A. Pasupuleti, W. D. Gaillard, R. F. Keating, and C. O. Oluigbo. Virtual reality–based 3-dimensional localization of stereotactic eeg (seeg) depth electrodes and related brain anatomy in pediatric epilepsy surgery. *Child's nervous system*, 38(3):537–546, 2022. 2

[13] A. V. Reinschluessel and J. Zagermann. Exploring hybrid user interfaces for surgery planning. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 208–210, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00048 2

[14] J. Roessel, M. Knoell, J. Hofmann, and R. Buettner. A systematic literature review of practical virtual and augmented reality solutions in surgery. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 489–498, 2020. doi: 10.1109/COMPSAC48688.2020.0-204 2

[15] D. Seo, B. Yoo, and H. Ko. Webizing collaborative interaction space for cross reality with various human interface devices. In *Proceedings of the 23rd International ACM Conference on 3D Web Technology*, pp. 1–8. ACM, Poznań Poland, June 2018. doi: 10.1145/3208806.3208808 2

[16] L. Stacchio, G. Vallasciani, G. Augello, S. Carrador, P. Cascarano, and G. Marfia. WiXaRd: Towards a holistic distributed platform for multi-party and cross-reality webxr experiences. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 264–272, Mar. 2024. doi: 10.1109/VRW62533.2024.00052 2

[7] https://assetstore.unity.com/packages/vfx/shaders/fullscreen-camera-effects/volume-viewer-pro-83185