3D Selection through Hand Tracking in XR for Point Clouds Using Signed Distance Fields

Luca Fluri, Arzu Çöltekin

FHNW University of Applied Sciences and Arts Northwestern Switzerland, Institute of Interactive Technologies, Windisch, Switzerland

Corresponding author: luca.fluri@fhnw.ch

Keywords: XR, Hand-tracking, Data Selection, Point Clouds, Signed Distance Fields

Introduction & Related Work

Three-dimensional data visualization in extended reality (XR) has become crucial dsdsacross medicine, astronomy, and geography (Fonnet & Prié 2021). While 3D visualizations offer unique spatial insights, selecting specific data subsets remains challenging despite being essential for exploratory analysis, volume measurement, and data annotation.

Current 3D selection solutions face limitations: traditional methods remain confined to 2D interfaces lacking immersive feedback, requiring mental translation between 2D projections and 3D relationships (Çöltekin et al. 2016). Even existing 3D approaches are restricted to simple geometric primitives (spheres, boxes) and fail to optimize for large datasets, creating computational bottlenecks unsuitable for real-time XR applications.

Existing XR selection techniques fall into ray-casting and volumetric categories but employ basic shapes or require specialized hardware, limiting expressiveness (Baloup et al. 2019). Early tools were limited to single-point selection, while recent hand-tracking approaches remain constrained by tracking quality limitations (Lubos et al. 2014). Notable techniques include Slice-n-Swipe (knife metaphor for slicing point clouds) (Bacim et al. 2014) and Tangible Brush (6DOF input with touch gestures) (Sereno et al. 2016). Automatic selection using machine learning requires extensive preprocessing, making it unsuitable for real-time exploratory tasks (Chen et al. 2020).

Despite advances in XR interaction and GPU processing, the intersection of hand-tracking, volumetric selection, and real-time point cloud processing remains largely unexplored. Hand-based selection aligns with natural human behaviors, offering spatially intuitive data interaction (Napier 2021). Compared to classical approaches, SDFs (Signed Distance Fields) enable compact geometry encoding supporting smooth interpolation and anti-aliased rendering across collision detection, mesh generation, and 3D rendering applications (Chermain et al. 2021). Our work combines algorithmic precision with intuitive hand-based interaction. We present a novel GPU-based SDF approach that bridges traditional 2D selection methods and XR paradigms, while maintaining real-time performance for large-scale point cloud datasets.

Our Approach

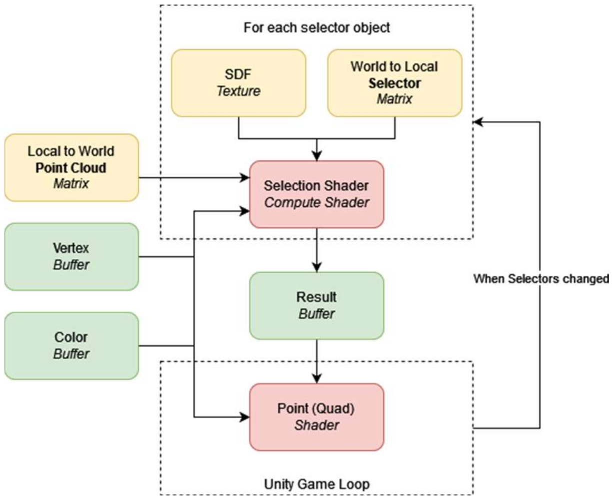

Figure 1 - Overview of our selection handling concept. Yellow represents transformation data, red represents shaders, and green represents GPU data buffers.

To address the limitations of existing 3D selection methods, we developed a Unity-based XR application that combines hand-tracking with GPU-accelerated SDF-based selection algorithms. Our approach leverages SDFs to represent arbitrary 3D selection volumes as compact 3D textures, which are processed entirely on the GPU using compute shaders for parallel point-in-volume testing. The system maintains all point cloud data (positions, colors, selection states) in GPU memory to minimize CPU-GPU transfer bottlenecks, while utilizing Unity's indirect instanced rendering pipeline to visualize millions of points in real-time (Neuman-Donihue et al. 2023) (see Figure 1 for the overall architecture). Hand-tracking serves as the primary input modality, enabling natural gesture-based interaction with volumetric selection tools (Buckingham 2021). This GPU-resident architecture supports efficient processing of large point clouds while providing the flexibility to implement diverse selection techniques ranging from geometric primitives to complex brushing-based interactions, all operating at real-time frame rates suitable for immersive XR environments.

Selection Techniques

We implemented four distinct selection techniques to demonstrate the versatility of our SDF-based framework, categorized into two interaction paradigms: direct geometry-based (Figures 2A and B) and brushing-based approaches (Figures 2C and D).

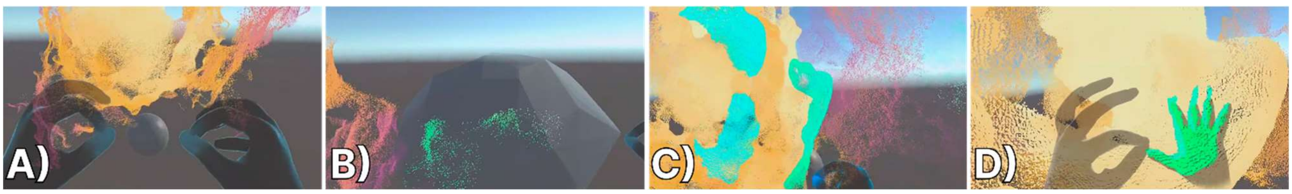

Figure 2 - Our four selection techniques. A) Pinch-Drag to spawn shapes, B) pinch-drag to draw convex hull, C) spherical brushing and D) using hands model directly as brush.

Shapes: Users spawn predefined geometric primitives (sphere, box, cylinder) through a "pinch and pull" gesture, where the distance between hands determines the scale during pinching and the position is automatically centered between the hands (Benko & Feiner, 2007). Once instantiated, shapes can be moved, rotated, and scaled through continued hand manipulation, allowing precise volumetric selection through familiar geometric forms.

Convex Hull: A novel mid-air drawing approach enabling users to sketch convex selection volumes using single-handed pinching gestures. Collected pinch positions are smoothed via Laplacian filtering before generating a 3D convex hull mesh, which is then converted to an SDF volume. This technique bridges the gap between rigid geometric primitives and freeform selection, albeit with restrictions to convex shapes.

Brush Sphere: Implements dynamic sphere-based selection where users spawn and manipulate a spherical brush volume. Users can sweep the sphere through the point cloud, with real-time scaling and positioning control, enabling intuitive "painting" of selections similar to traditional 2D brush tools but extended to 3D space (Akers et al. 2004).

Brush Hands: A novel technique utilizing live hand mesh geometry as selection volumes. Real-time SDF generation from tracked hand geometry enables direct use of natural hand shapes for selection. To prevent unintended selections (the "Midas touch" problem) (Conte 2020), the technique requires activation through pinching with the opposite hand, providing intentional control over when hands function as active selectors.

All techniques are supported by a hand-attached palm menu for mode switching and parameter adjustment, accessed through poking interactions. A bimanual transformation system uses a "double-fist" gesture implementing the handlebar metaphor (Song et al. 2012), where the midpoint between fists controls translation, hand distance controls scaling, and hand orientation defines rotation of the entire point cloud and selection volumes.

Discussion & Conclusion

Our evaluation demonstrates that the SDF-based selection approach successfully addresses key limitations of traditional 3D selection methods while maintaining real-time performance for practical XR applications. The GPU-accelerated architecture achieves linear scalability, supporting interactive selection of point clouds up to 5 million points at acceptable frame rates (>60 FPS), with rendering rather than selection computation serving as the primary bottleneck.

The integration of real-time SDF generation with hand tracking represents a significant technical achievement, enabling the arbitrary selection of volumes that extend beyond simple geometric primitives. Our work demonstrates that efficient, real-time selection of near-arbitrary point cloud subsets is achievable through careful GPU optimization and SDF-based volume representation. However, current limitations include restrictions to convex shapes in the Convex Hull technique, performance degradation beyond 10 million points, and dependence on variable hand-tracking accuracy across devices.

Future enhancements should focus on hierarchical data structures for larger datasets (Schütz et al. 2020), supporting concave selection volumes, implementing Boolean operations between selections, and integrating AI-assisted selection methods. Performance optimizations, including frustum culling and out-of-core streaming, would extend applicability to massive datasets.

This work establishes a robust foundation for immersive scientific visualization, demonstrating that intuitive, precise, and efficient 3D selection is feasible in XR environments. The combination of SDF processing with natural hand-tracking interactions pushes the boundaries of spatial data exploration, opening pathways for next-generation visualization tools that integrate human intuition with computational precision. Our findings suggest significant potential for broader adoption of XR-based data analysis tools across scientific and industrial applications requiring complex 3D data manipulation.

References

Akers, D., Sherbondy, A., Mackenzie, R., Dougherty, R. & Wandell, B. 2004. Exploration of the brain's white matter pathways with dynamic queries. In IEEE Visualization 2004, 377–384.

Bacim, F., Nabiyouni, M. & Bowman, D.A. 2014. Slice-n-Swipe: A freehand gesture user interface for 3D point cloud annotation. In 2014 IEEE Symposium on 3D User Interfaces (3DUI), 185–186.

Baloup, M., Pietrzak, T. & Casiez, G. 2019. RayCursor: A 3D Pointing Facilitation Technique based on Raycasting. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–12.

Benko, H. & Feiner, S. 2007. Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection. In 2007 IEEE Symposium on 3D User Interfaces.

Buckingham, G. 2021. Hand Tracking for Immersive Virtual Reality: Opportunities and Challenges. Frontiers in Virtual Reality 2.

Chen, Z., Zeng, W., Yang, Z., Yu, L., Fu, C.-W. & Qu, H. 2020. LassoNet: Deep Lasso-Selection of 3D Point Clouds. IEEE Transactions on Visualization and Computer Graphics 26(1), 195–204.

Chermain, X., Lucas, S., Sauvage, B., Dischler, J.-M. & Dachsbacher, C. 2021. Real-Time Geometric Glint Anti-Aliasing with Normal Map Filtering. Proc. ACM Comput. Graph. Interact. Tech. 4(1), 1:1–1:16.

Çöltekin, A., Lokka, I. & Zahner, M. 2016. On the usability and usefulness of 3d (geo) visualizations–a focus on virtual reality environments. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 41, 387–392.

Conte, T. 2020. The Midas Touch effect: the most unknown phenomenon in UX design.

Fonnet, A. & Prié, Y. 2021. Survey of Immersive Analytics. IEEE Transactions on Visualization and Computer Graphics 27(3), 2101–2122.

Lubos, P., Beimler, R., Lammers, M. & Steinicke, F. 2014. Touching the Cloud: Bimanual annotation of immersive point clouds. In 2014 IEEE Symposium on 3D User Interfaces (3DUI), 191–192.

Napier, J. 2021. Hands. Princeton: Princeton University Press.

Neuman-Donihue, E., Jarvis, M. & Zhu, Y. 2023. FastPoints: A State-of-the-Art Point Cloud Renderer for Unity. arXiv preprint arXiv:2302.05002.

Schütz, M., Ohrhallinger, S. & Wimmer, M. 2020. Fast Out-of-Core Octree Generation for Massive Point Clouds. Computer Graphics Forum 39(7), 1–13.

Sereno, M., Ammi, M., Isenberg, T. & Besançon, L. 2016. Tangible Brush: Performing 3D Selection with Portable and Position-aware Devices.

Song, P., Goh, W.B., Hutama, W., Fu, C.-W. & Liu, X. 2012. A handle bar metaphor for virtual object manipulation with mid-air interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1297–1306.