InstantXR: Instant XR Environment on the Web Using Hybrid Rendering of Cloud-based NeRF with 3D Assets

Moonsik Park[1], Byonghyun Yoo[2,3], Jee Young Moon[2], Ji Hyun Seo[2,4]

[1]ESTsoft Corporation, [2]Korea Institute of Science and Technology, [3]University of Science and Technology, [4]Korea University

Introduction

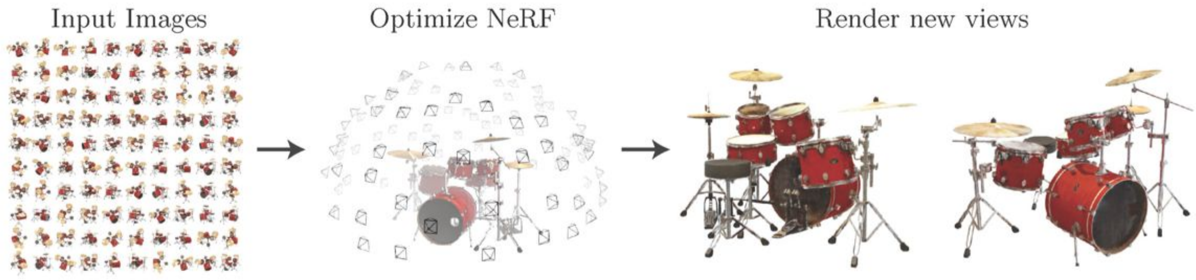
Moonsik Park

- Software Engineer, ESTsoft Corporation (2022. 07. - Current)
- Research Intern, Korea Institute of Science and Technology (2022. 03. - 2022. 06.)
- Undergraduate Student, The Catholic University of Korea (2020. 03. - Current)

Research Interest

- Neural Rendering technologies and their applications

Neural Radiance Fields - Idea



- Synthesizing novel views of complex scenes by optimizing an underlying continuous volumetric scene function using a sparse set of input views
- Surprisingly photorealistic output
- Very slow and compute-intensive
    - Several days to optimize NeRF, several hours to render a video with novel views

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I. Springer-Verlag, Berlin, Heidelberg, 405–421. https://doi.org/10.1007/978-3-030-58452-8_24

Neural Radiance Fields - Advances

- NVLabs “Instant Neural Graphics Primitives”
- Sped up NeRF significantly
- Optimizing
  - Simple scene under 5 seconds
  - Complicated scene under 5 minutes
- Rendering (depending on the resolution)
  - 50 ~ 100 ms

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant neural graphics primitives with a multiresolution hash encoding. ACM Trans. Graph. 41, 4, Article 102 (July 2022), 15 pages. https://doi.org/10.1145/3528223.3530127

Neural Radiance Fields - Applications

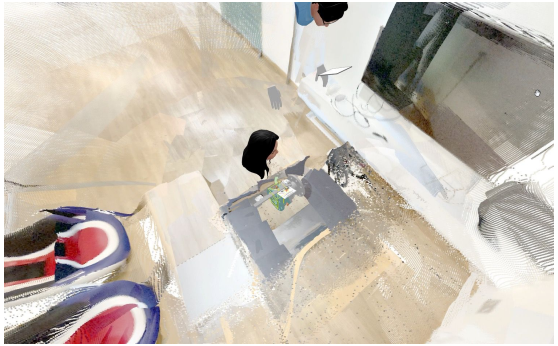
Using NeRF to map large environments

Using NeRF depth estimation to navigate robots

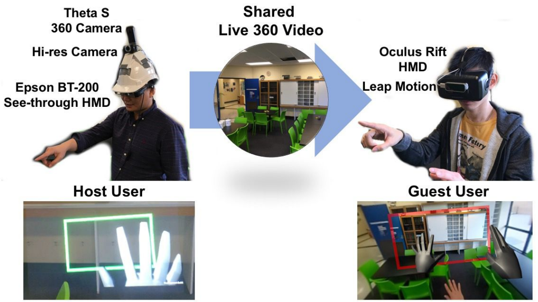
1 Matthew Tancik, Vincent Casser, Xinchen Yan, Sabeek Pradhan, Ben Mildenhall, Pratul P. Srinivasan, Jonathan T. Barron, and Henrik Kretzschmar. 2022. Block-NeRF: Scalable Large Scene Neural View Synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 8248–8258.
2 M. Adamkiewicz et al., "Vision-Only Robot Navigation in a Neural Radiance World," in IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 4606-4613, April 2022, doi: 10.1109/LRA.2022.3150497.

Sharing Environments in Extended Reality (XR)

3D Reconstructed Model[1]



Point Cloud Streaming[2]



3D Camera Streaming[3]

1 Lee, G. A., Teo, T., Kim, S., & Billinghurst, M. (2017). Mixed Reality Collaboration through Sharing a Live Panorama. SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, 1–4. https://doi.org/10.1145/3132787.3139203
2 Yongjae Lee, Byounghyun Yoo, and Soo-Hong Lee. 2021. Sharing Ambient Objects Using Real-time Point Cloud Streaming in Web-based XR Remote Collaboration. In The 26th International Conference on 3D Web Technology (Web3D '21). Association for Computing Machinery, New York, NY, USA, Article 4, 1–9. https://doi.org/10.1145/3485444.3487642
3 Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G., & Billinghurst, M. (2017). Exploring Enhancements for Remote Mixed Reality Collaboration. SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, 1–5. https://doi.org/10.1145/3132787.3139200

Using NeRF to Share Environments

NeRF's exceptional ability

- Excellent novel view synthesis of complex scenes with photorealistic quality
- No preprocessing necessary other than taking a video of the environment
- Very small output (under 50 megabytes)

Roadblocks exist

- Rendering the novel view is compute-intensive
- Modifying the already optimized scene is hard
- Can't make modifications to the "rendering pipeline"

InstantXR

Cloud Rendering

- Scalable architecture with a render server and multiple cloud renderers
- Adding objects to NeRF scenes using depth harmonization

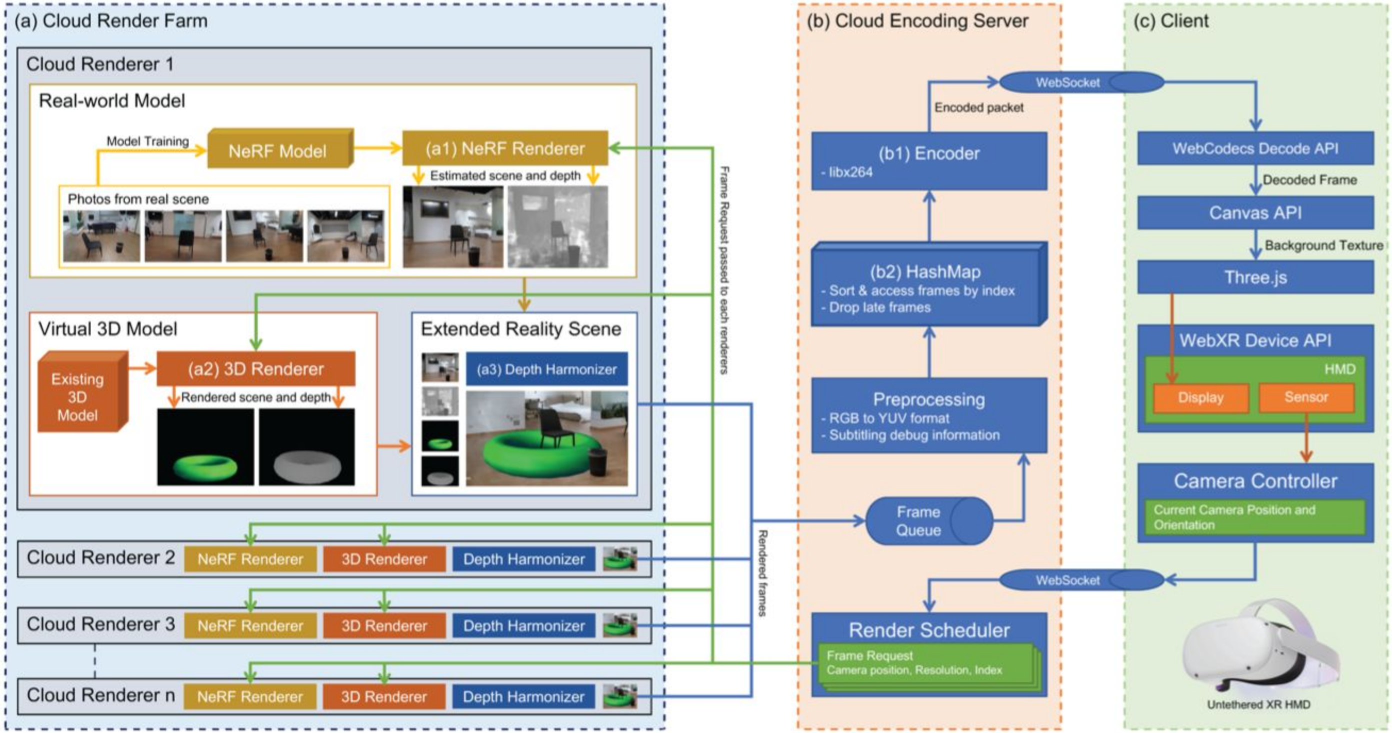
Transport

- Novel low latency streaming method
- Low latency head position sharing

Client

- No requirements to the client other than a web standards compliant browser

InstantXR Architecture



Real-time Cloud-based NeRF with InstantXR

Designing a "distributed remote rendering system"

- Lowering Motion-to-Photon latency using a cloud render farm
- 'Cloud Encoding Server' paired with multiple 'Cloud Renderer's
- Scheduling required
  - Dropping frames that take too long to render
  - Distributing render requests based on rendering speed, latency, and difficulty to render
  - Order of render request and artifacts should be honored
- XR devices have stereoscopic displays: two video streams!
- Low level operations and optimizations required
  - Renderer and server written in C++

Streaming InstantXR

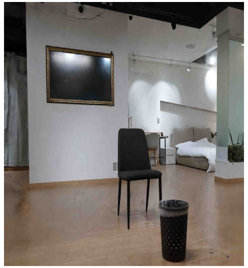
User’s head position

- Periodically sends the position every 20 ms
- Using web standards WebXR and WebSockets API

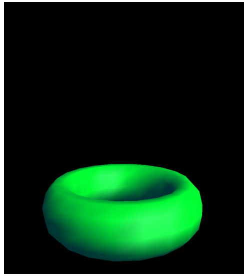
Novel streaming strategy to reduce latency

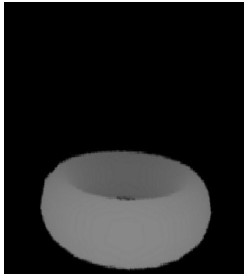
- Sending the user raw compressed packets
- Does not have the notion of “timestamps” (show frames as soon as they arrive)
- No audio/video synchronization
- Using web standards WebCodecs and WebSockets API

Harmonizing 3D assets with InstantXR scenes



Scene and estimated depth from NeRF



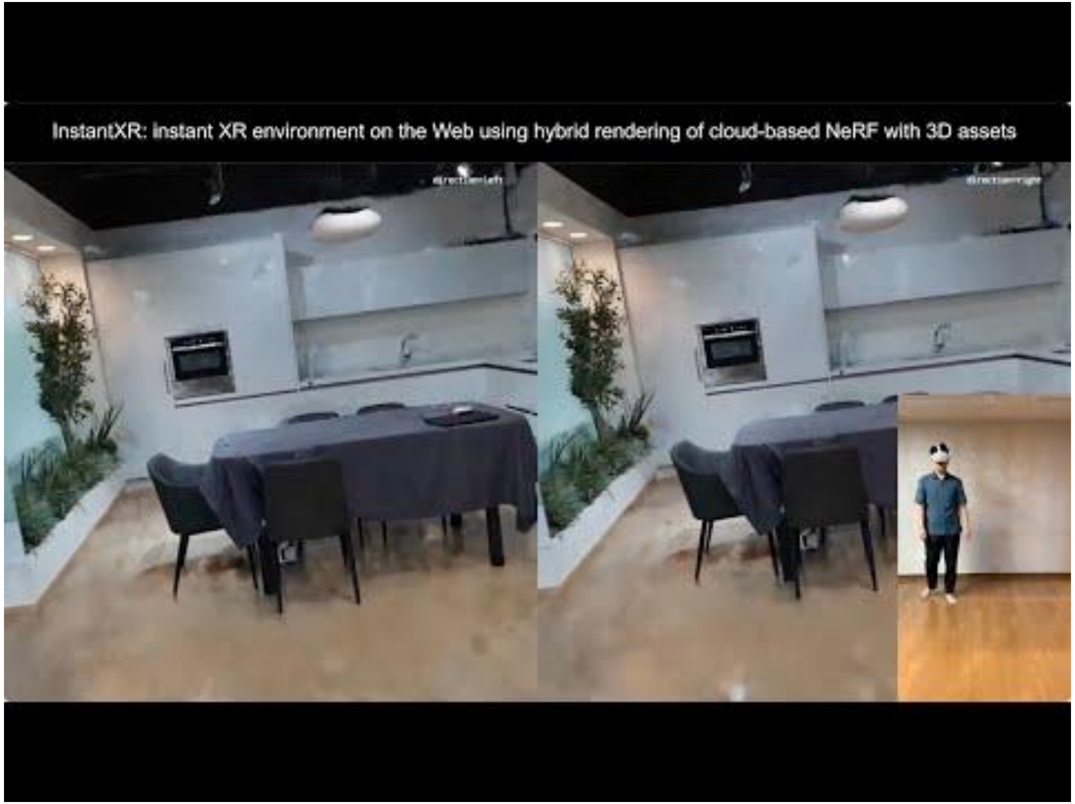


Torus rendered using OpenGL

Depth Testing

Harmonized Result

InstantXR Demo



Conclusion

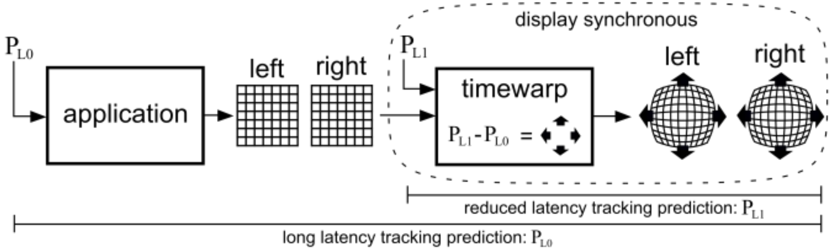
Contribution

- Our method provides a mirror world of an existing space as an immersive XR environment without the cumbersome modeling process
- One of the first attempts to apply nerf to real world use cases

Opinion

- NeRF has a potential to be a new data storage format for scenes
- Research in handling NeRF scenes should be done
- Research on delivery of volume rendered artifacts should be done

Future Work



3D Rendering pipeline with asynchronous time warp[1]

- User experience comparison with other methods with;
  - 3D reconstruction model transfer
  - point cloud streaming
  - 3D camera streaming
- Lowering latency even more
- Improving user experience
  - Judder Control
  - Noise reduction of head position
- Method like asynchronous time warp, for volume rendering

1 J. M. P. van Waveren. 2016. The asynchronous time warp for virtual reality on consumer hardware. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16). Association for Computing Machinery, New York, NY, USA, 37–46. https://doi.org/10.1145/2993369.2993375

Thank you!

Visit https://moonsikpark.github.io/instantxr/ for demo video and code release