



Depth Augmented Stereo Panorama for Cinematic VR

J. Thatte {jayantt@Stanford.edu}, J.-B. Boin, H. Lakshman, G. Wetzstein, B. Girod
Department of Electrical Engineering, Stanford University

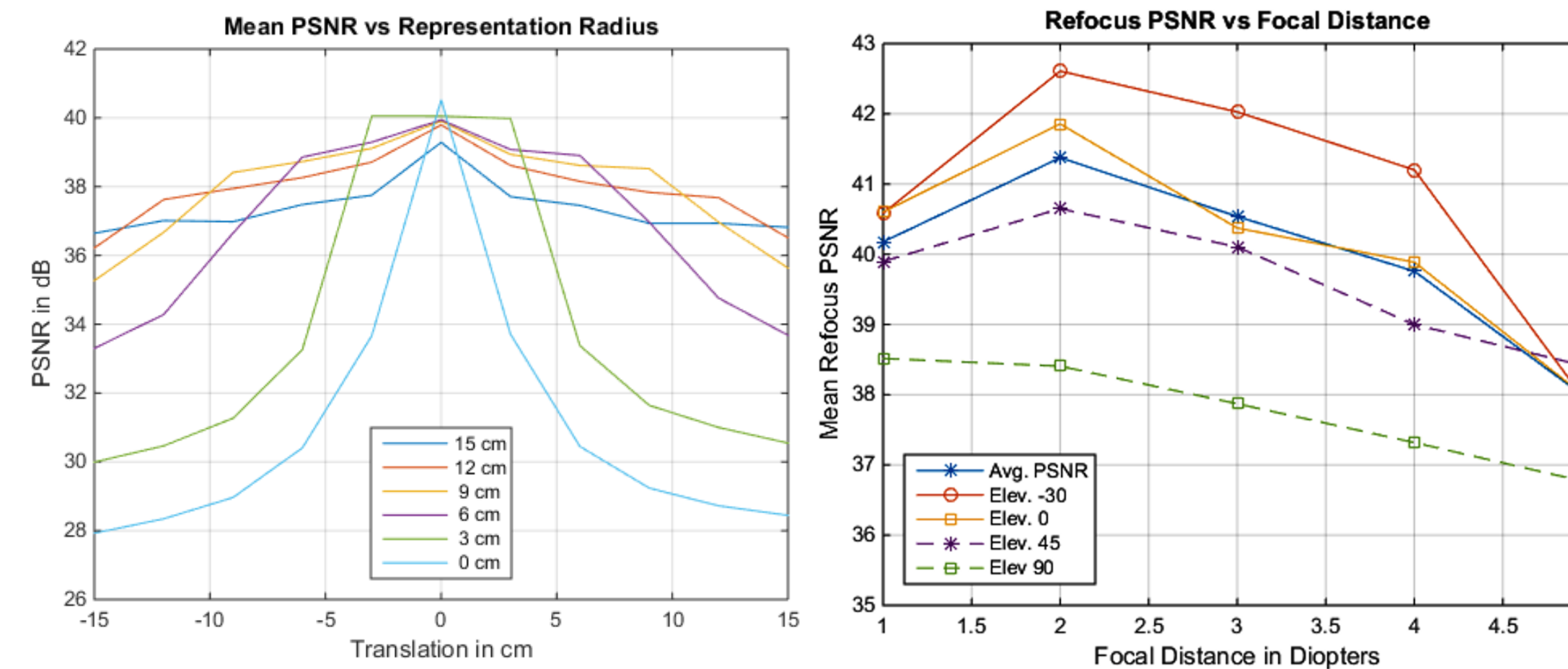
SCIEN

The Stanford Center for
Image Systems Engineering

Abstract

- Current **cinematic VR** systems do not address head-motion parallax or focus cues, both or which are required for viewing comfort and for depth perception in head-mounted displays
- We propose a new stereo content representation called **Depth Augmented Stereo Panorama** (DASP), which has been developed keeping in mind the different stages of VR pipeline from capture to rendering
- Our proposed representation can produce accurate **head-motion parallax** and **focus cues** within a predefined operating range

Quantitative Evaluation



Mean PSNR as a function of head translation for varying representation radius (left), mean PSNR of refocused images using synthesized lightfields (right)

Representation Radius	15 cm	12 cm	9 cm	6 cm	3 cm	0 cm
Mean PSNR (dB)	37.4	37.9	37.9	36.9	33.8	30.9
Std. Dev. (dB)	2.3	2.8	3.2	3.9	5.9	5.8

Mean and std. dev. PSNR over a range of head translations, as a function of representation radius

Proposed Representation

Format	Data Capture	Post-production	Coding & Streaming	Head motion parallax
3D models	hard	hard	medium	simple
Omnistereo [1]	medium	medium	simple	hard
Concentric mosaics [2]	medium	hard	medium	simple
Mono Panorama + Depth	simple	simple	simple	hard
Proposed: Omnistereo + Depth	medium	medium	simple	simple

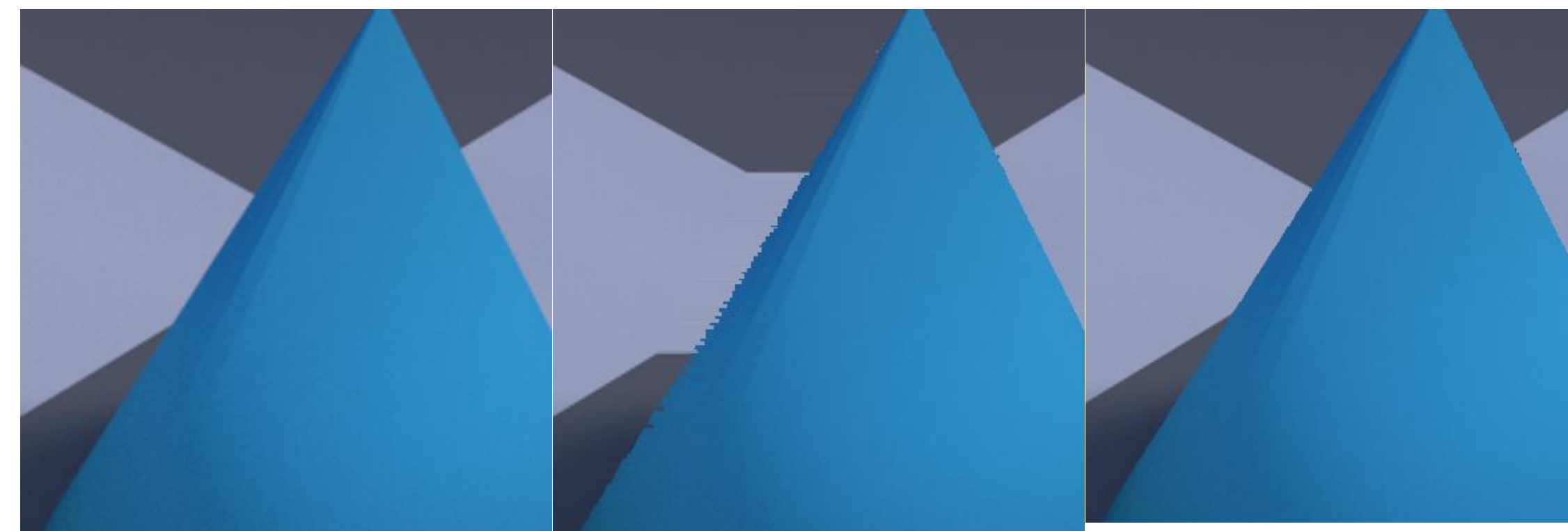
Representation:

- The proposed representation consists of stereo texture panoramas, with a large viewing radius; each panorama is augmented with the corresponding depth
- We extend the traditional omnistereo to use a shrinking viewing radius (as a function of the elevation angle) and a spherical imaging surface.

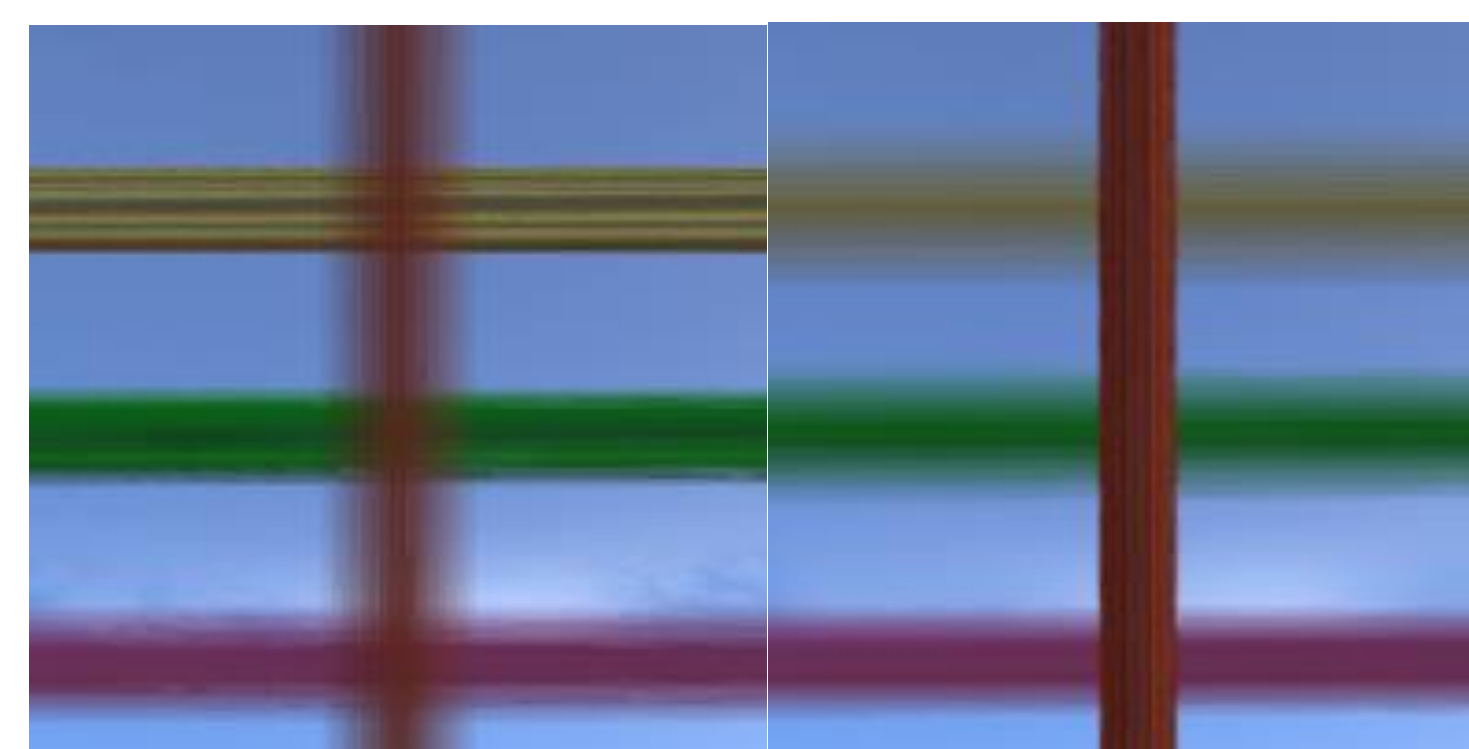
Advantages:

- Unlike traditional omnistereo, we are able to represent the entire scene, including zenith and *nadir* points
- Having two extreme views for each scene point helps drastically reduce the need for hole-filling; suitable for fast and easy rendering
- Adds on top of the existing 3DTV workflow

Occlusions and Blur



Left to right: Ground truth, synthesized image using monoscopic texture and depth and using proposed method

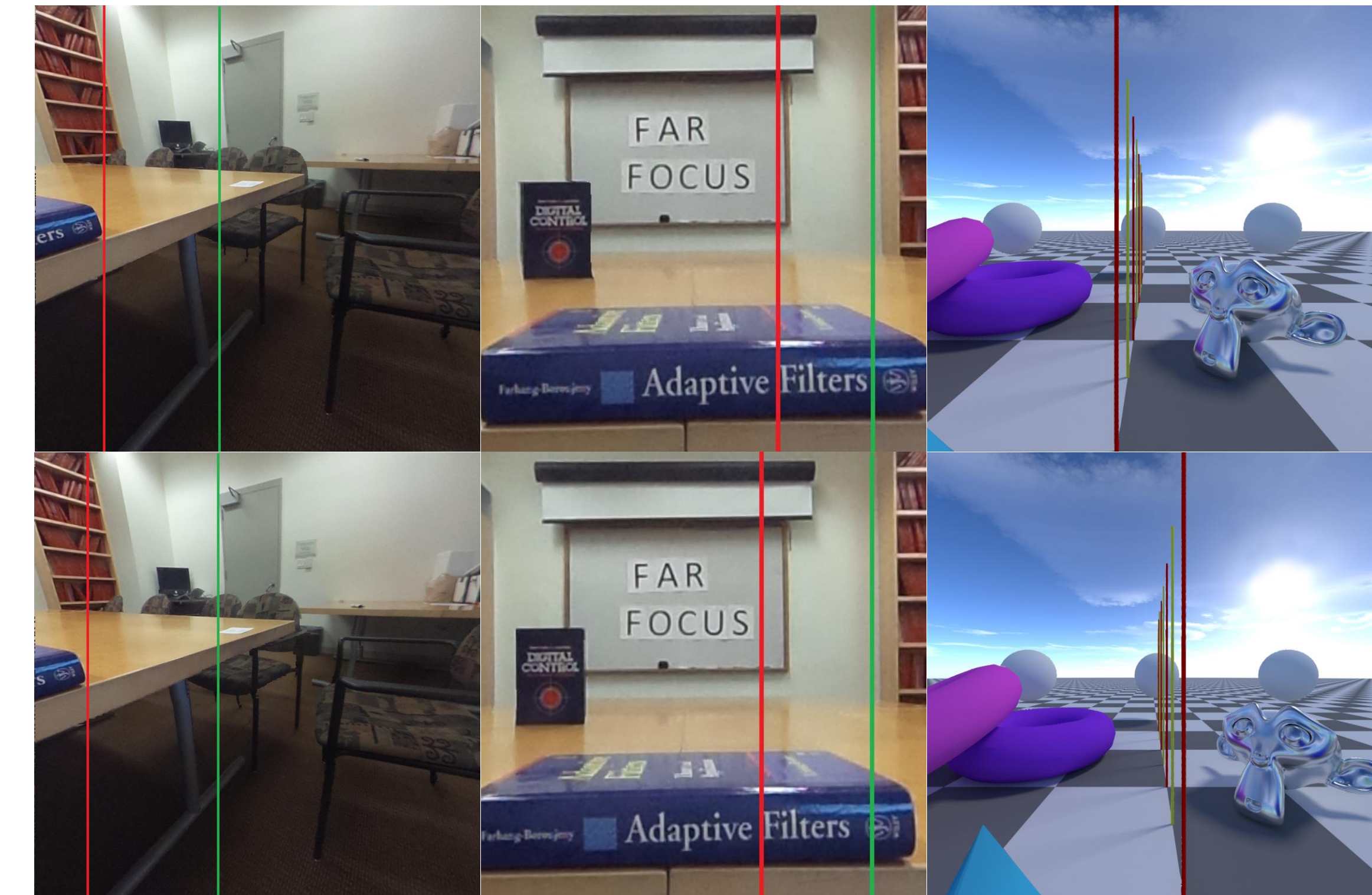


DASP: Closer objects blur over farther objects but not the other way around, as expected.

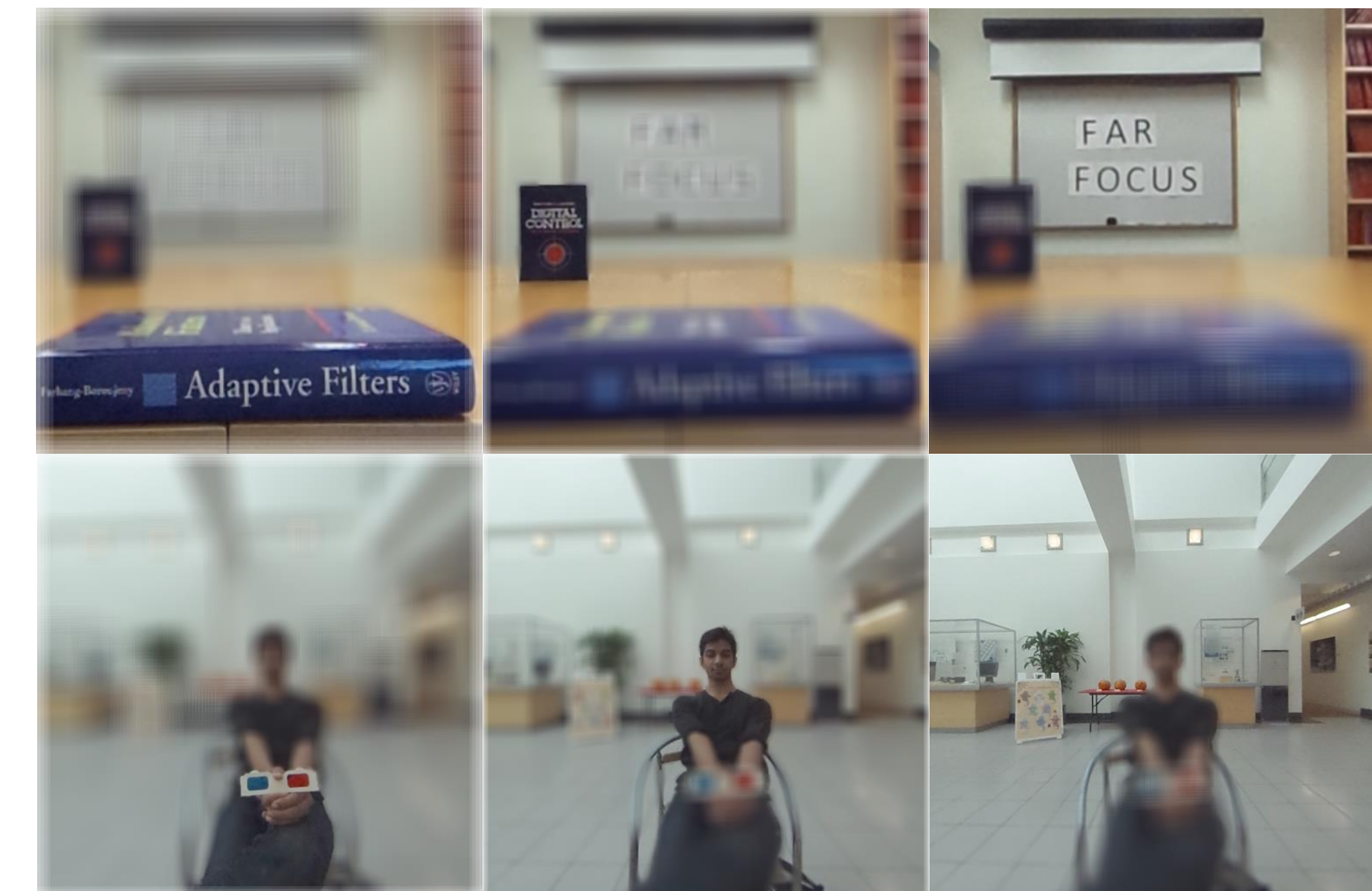
References

- [1] Peleg, et. al., "Omnistereo: Panoramic stereo imaging," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001.
- [2] Shum et. al., "Rendering with concentric mosaics," 26th annual conference on Computer graphics and interactive technique, 1999.

Synthesized Results



Head-Motion Parallax: Examples of synthesized parallax for left-right head translation of 8, 4 and 15 cm, respectively



Refocus using Synthesized Lightfield: (left to right) Near, mid and far focus produced using DASP

Conclusion

- DASP provides a framework to support head motion parallax and focus cues in cinematic VR
- The data overhead is significantly smaller than previously proposed techniques [2]
- DASP requires minimal hole-filling and is able to produce high quality output, as compared to ground truth obtained from synthetic 3D models