View Synthesis in Casually Captured Scenes Using a Cylindrical Neural Radiance Field With Exposure Compensation

Wesley Khademi Jonathan Ventura
California Polytechnic State Universitiy

SIGGRAPH 2021

PROBLEM

Casual capture is when a person tries to capture imagery of an entire scene by spinning a smartphone in a circle. Our goal is to learn to render high quality novel views of a casually captured 360-degree scene.

RELATED WORK

- Neural Radiance Fields (NeRF) [Mildenhall et al. 2020] introduced an implicit representation for learning a scene, producing impressive high quality photorealistic novel views. However, NeRF struggles to faithfully recover appearance and geometry of outward facing unbounded scenes.
- NeRF++ [Zhang et al. 2020] introduced an inverted sphere parameterization enabling NeRF to learn unbounded scenes. However, NeRF++'s novel views suffer from artifacting in casually captured scenes.

MOTIVATION

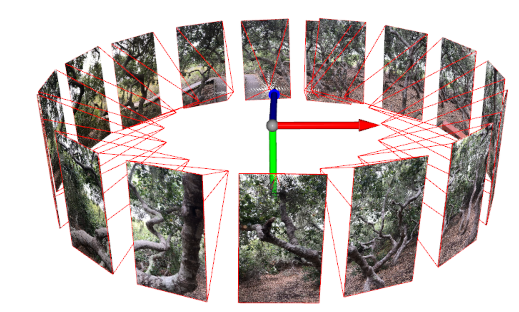
- Extend NeRF to learn to accurately represent the appearance and geometry of 360-degree outward facing scenes.
- Learn to compensate for exposure differences across training images that may arise from casual capture to reduce artifacts and inconsistent exposure in novel views.

OUR APPROACH

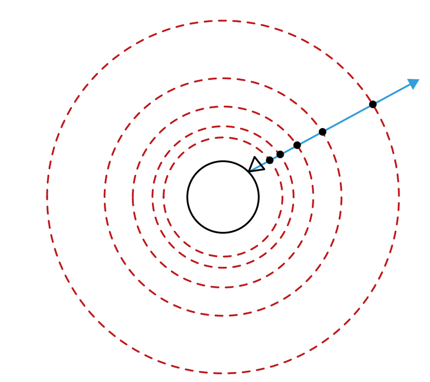
- We introduce a cylindrical parameterization that represents a projected point on the unit cylinder and an inverse radius bounded in [0,1] to recover which cylinder the sample point lies on in space.
- By bounding our input, we overcome NeRF's inability to encode a large range of scene coordinates, allowing our method to accurately reconstruct outward facing unbounded scenes.
- We also propose an exposure compensation technique that accounts for exposure differences in training images.
- By introducing a learnable brightness parameter for each training image, our model learns small exposure adjustments that account for mismatch in exposure across images.

VIEW SYNTHESIS PIPELINE

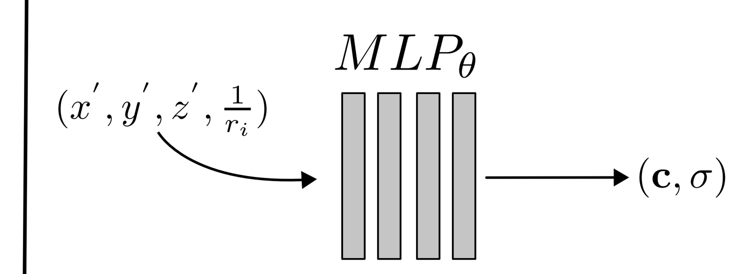
Step 1: Spin in circle and capture photos of a scene.



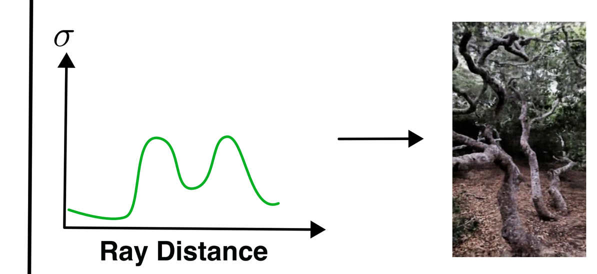
Step 2: Perform Structure from Motion (SfM) to obtain camera poses.



Step 3: For each camera ray, sample points along the ray that intersect cylinders of certain radii.



Step 4: Query MLP at each sampled point along the ray to obtain its color and density.

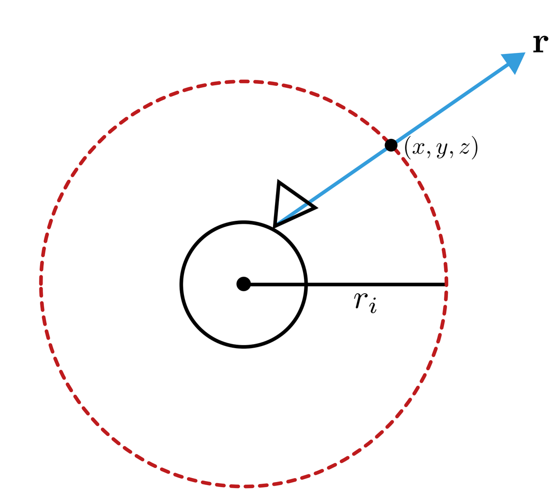


Step 5: Perform volume rendering.

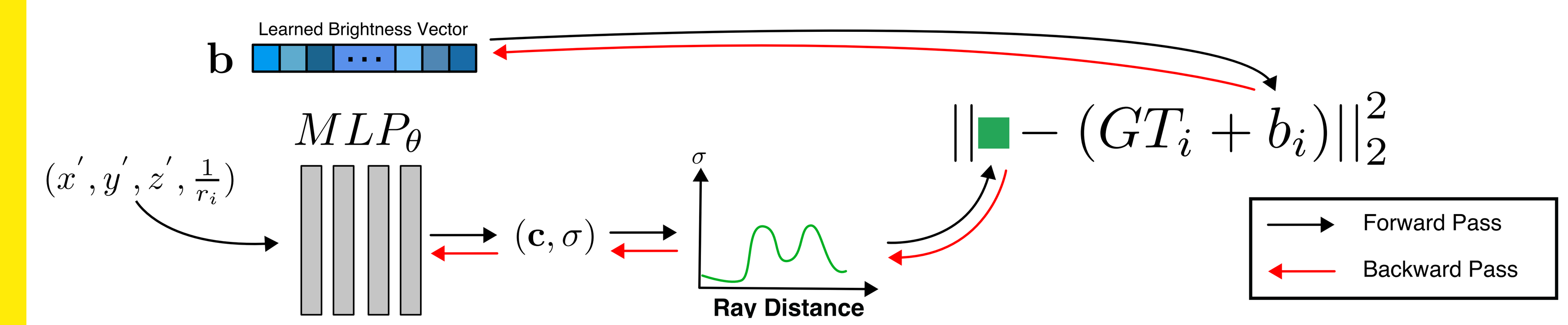
CYLINDRICAL PARAMETERIZATION

Sample points to query the MLP at by:

1) Uniformly sampling $\frac{1}{r_i} \in [0, 1]$ where r_i is the radius of a cylinder
2) Computing the t value where ray $\mathbf{r} = \mathbf{o} + t\mathbf{d}$ intersects a cylinder via constraint:
$$(o_x + t_i d_x)^2 + (o_z + t_i d_z)^2 = r_i^2$$
3) Computing sample point $(x, y, z) = \mathbf{r}(t_i)$
4) Projecting point onto unit cylinder to obtain (x', y', z')
5) Reparameterizing the projected point as the 4D coordinate $(x', y', z', \frac{1}{r_i})$



EXPOSURE COMPENSATION

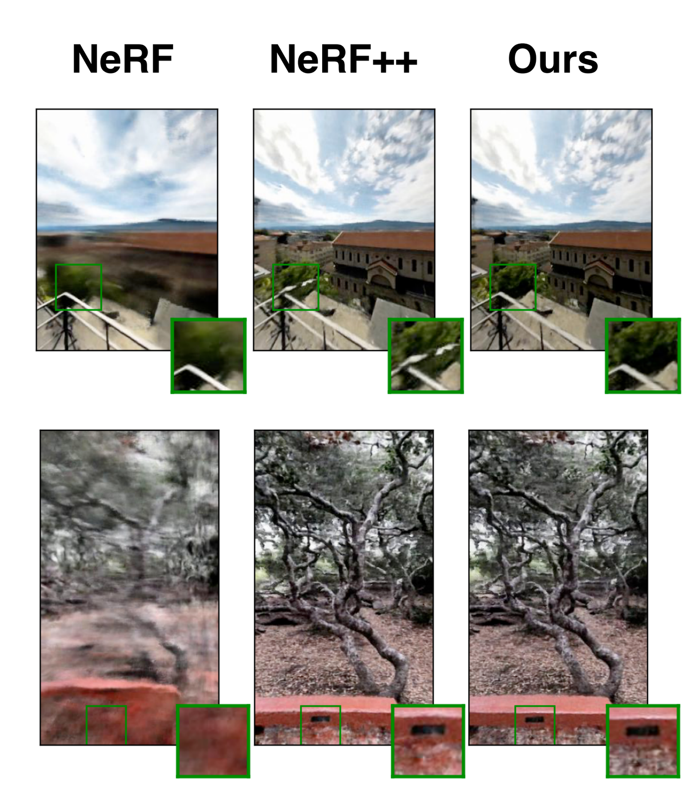


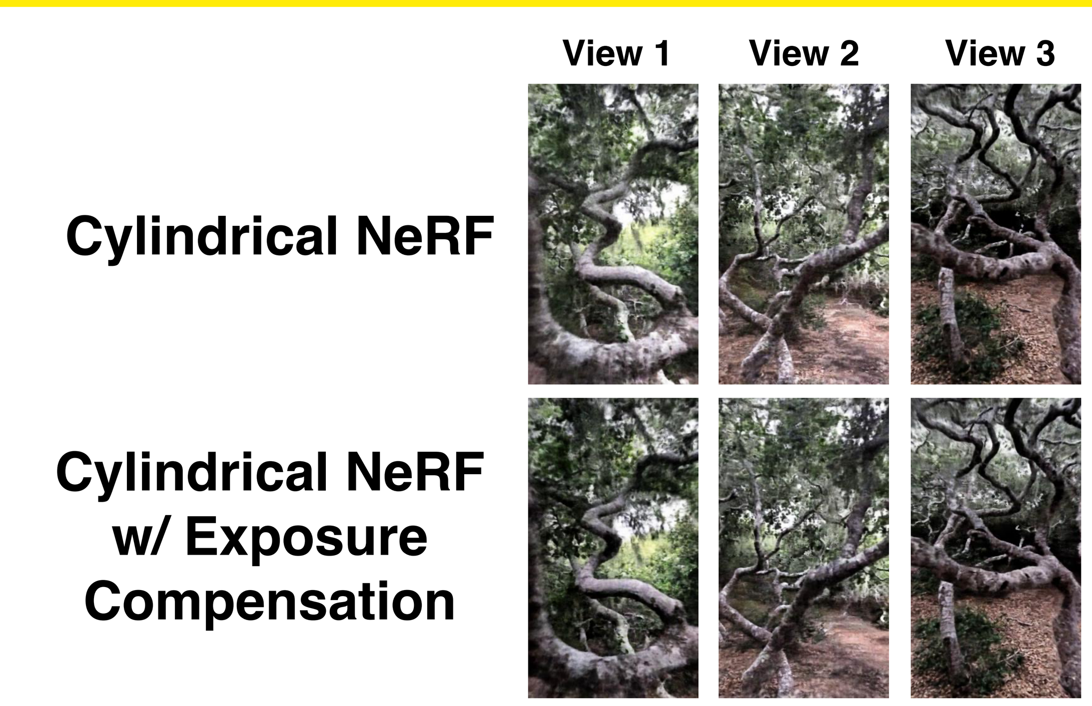
Training Time:
To have each image's exposure match the first training image's exposure, we fix $b_0 = 0$ and jointly optimize the brightness parameters $\{b_i\}_{i=1}^N$ and the weights of the MLP.

Test Time:
Remove the learned brightness vector and run Cylindrical NeRF.

RESULTS

Our method is capable of generating high quality novel views on large unbounded scenes that NeRF struggles to learn. Compared to NeRF++, our method reduces the amount of small floating artifacts that appear within novel views.





Our exposure compensation technique helps maintain consistent exposure across different views of a scene.

REFERENCES

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 *European Conference on Computer Vision*. Springer, 405–421.

Kai Zhang, Gernot Riegler, Noah Snavely, and Vladlen Koltun. 2020. NeRF++: Analyzing and Improving Neural Radiance Fields. arXiv:2010.07492 [cs.CV]