Block-NeRF: Scalable Large Scene Neural View Synthesis

Matthew Tancik, Vincent Casser, Xinchen Yan, Sabeek Pradhan,
Ben P. Mildenhall, Pratul Srinivasan, Jonathan T. Barron, Henrik Kretzschmar

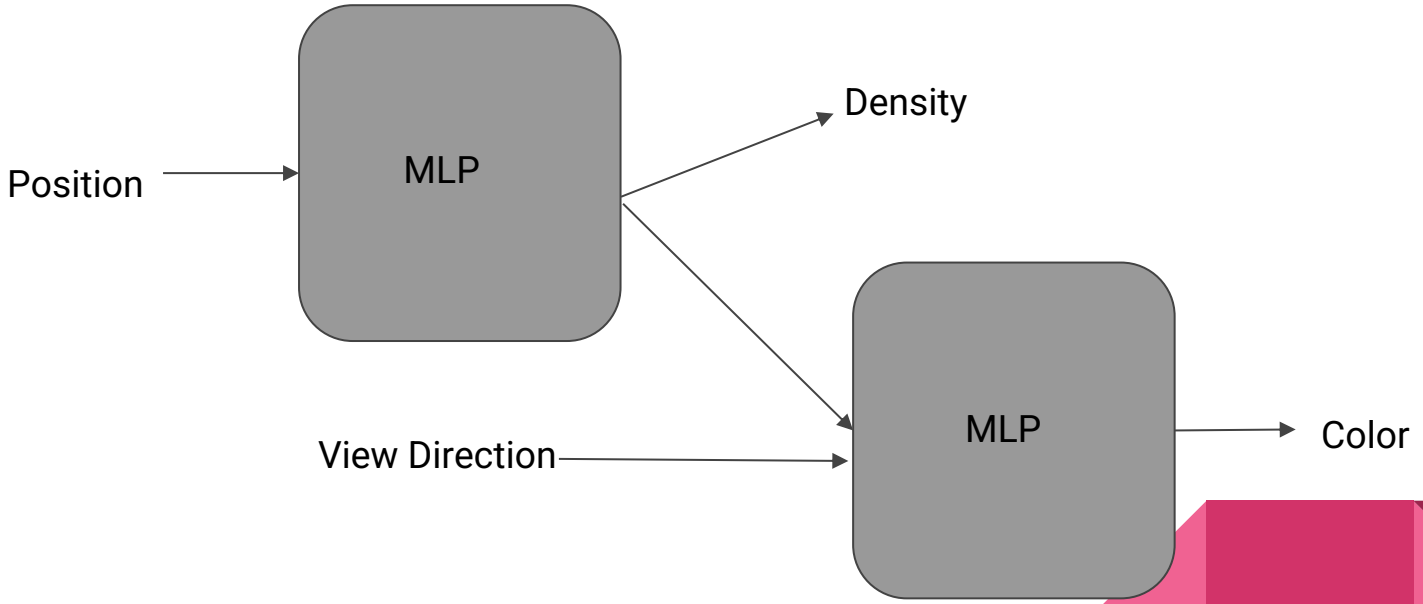
What is Block-NeRF

- Represented large-scale environments
- Individually trained NeRFs
- Aligned the predictions

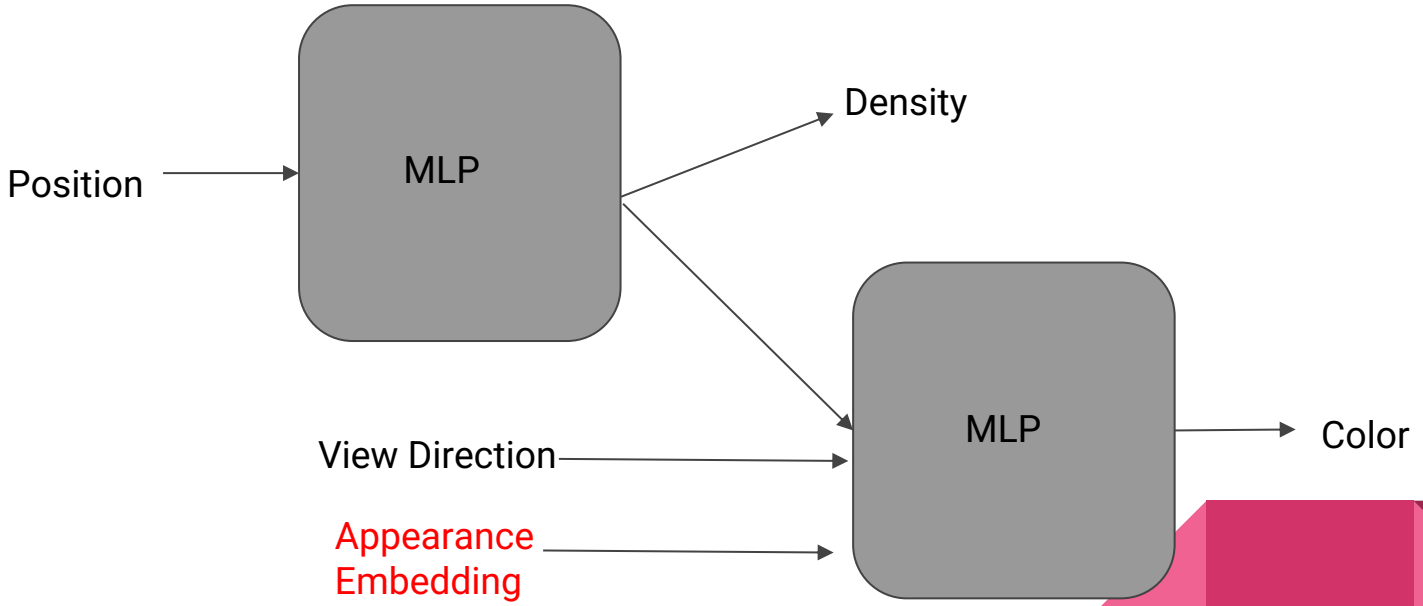
Why Block-NeRF

NeRF and other methods typically lead to significant artifacts when applying to large environments.

NeRF



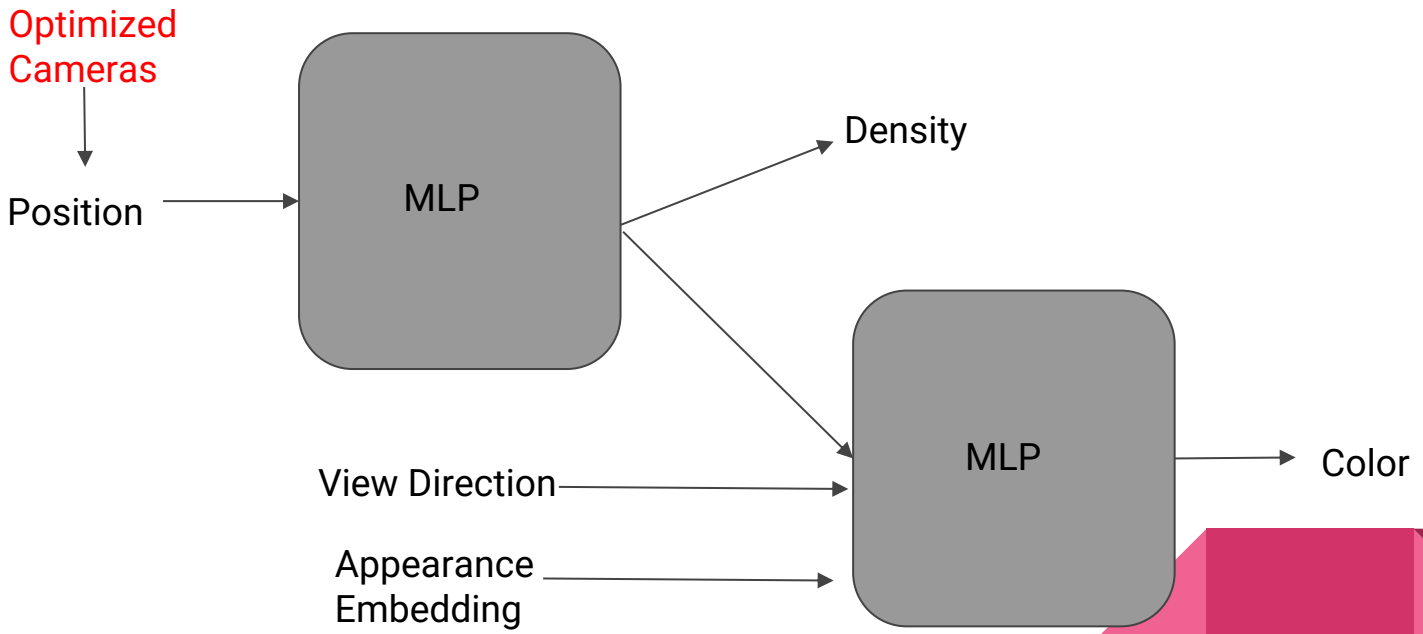
Modification



Appearance Embeddings

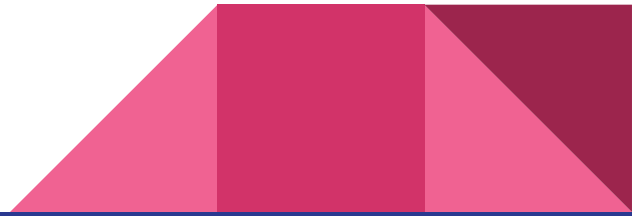
- allowed the NeRF to explain away several appearance-changing conditions, such as varying weather and lighting
- interpolated between different conditions, such as cloudy versus clear skies, or day and night

Modification



Learned Pose Refinement

- Pose offsets are learned per driving segment.
- Each offset has a translation and rotation component.
- Regularized pose offsets firstly learn a rough structure prior to modifying the poses.



Other Considerations

- **Exposure Input**

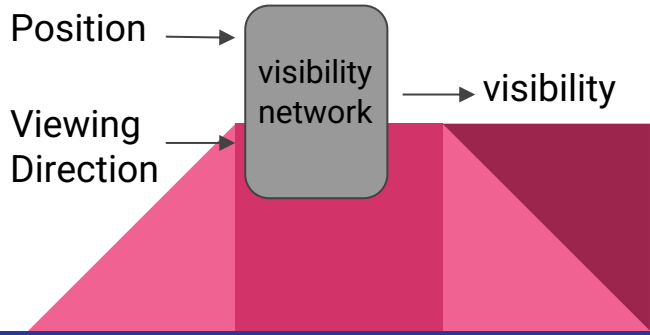
  Exposure information can compensate for the visual differences.

- **Transient Object**

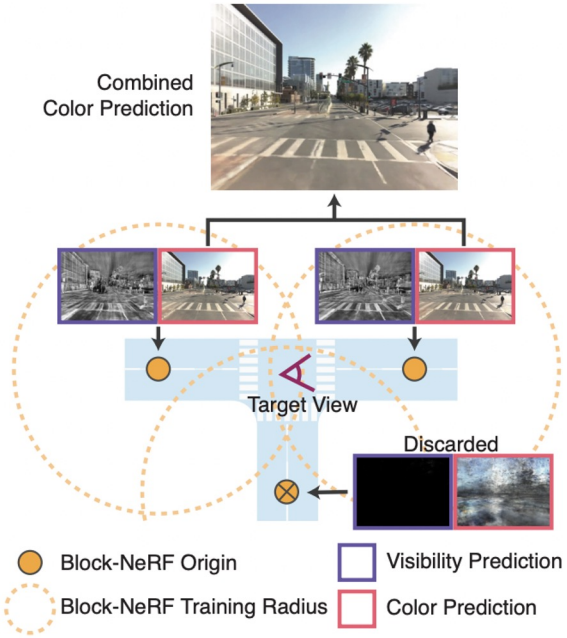
  Moving and transient objects are ignored for the assumption of consistency.

- **Visibility Prediction**

  When merging Block-NeRFs, visibility is important.



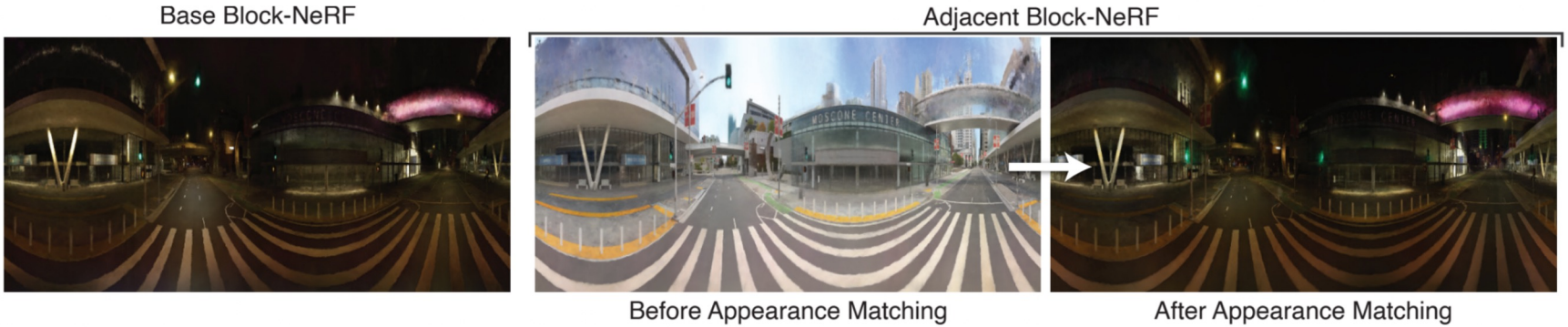
Merging



- Splitted each Block-NeRFs within radius
- Discarded the Block-NeRF with the lowest visibility
- Combined the remaining color renders

Appearance Matching

- Same appearance latent code typically leads to different appearances.
- Select a high-visibility matching position between two adjacent Block-NeRFs
- Freeze the network weights and optimize the appearance code of the target



Comparison

Result

- San Francisco Alamo Square Dataset

Limitations

- Missing transient objects
- Inability to handle dynamic objects
- Distant objects are blur
- Computationally expensive

Potential Solutions

- Missing transient objects
  - learning transient objects
- Inability to handle dynamic objects
  - directly modeling dynamic objects
- Distant objects are blur
  - different sampling method
- Computationally expensive
  - NeRF caching techniques or sparse voxel grids and multiple orders of magnitude

Conclusion

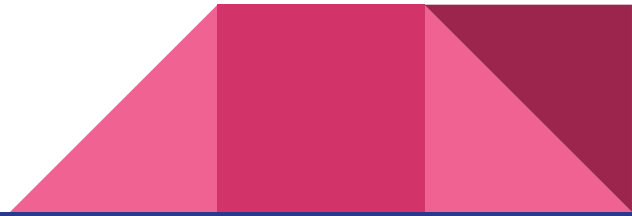
Block-NeRF enables large-scale scene reconstruction by representing the environment using multiple NeRFs.

Novelties:

- Separately trained NeRFs on the large scale environment
- Considered appearance embedding and exposure into the model
- Predicted Visibility for merging images

Future Improvements:

- Model capacities to consider more components(dynamic objects etc.)
- Less computational expense

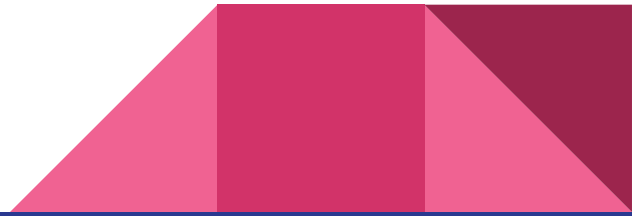


Questions

Quiz

Does color relate to the prediction of visibility of Block-NeRF?

No. Visibility is predicted by an independent network which takes position and viewing direction as inputs.



Quiz

Should we train the following images for Block-NeRF?

No. Temporal inconsistencies break the assumption of consistency.