Balanced Spherical Grid for Egocentric View Synthesis

Changwoon Choi, Sang Min Kim, Young Min Kim
3D Vision Lab @ Seoul National University

Poster Session THU-AM-011

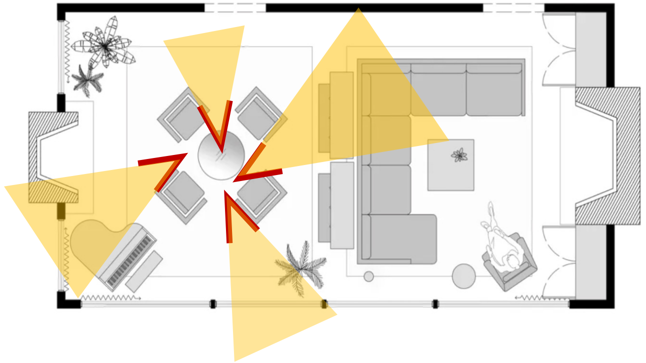
code&data   project page

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

EgoNeRF – Motivation & Goal

- NeRF & grid-based NeRFs [1, 2, 3] can visualize the photorealistic appearance of bounded scenes

Inward-facing views



Outward-looking views

Goal

- We propose a *practical* *solution* to reconstruct *large-scale* environments

[1] Anpei Chen et al., Tensorf: Tensorial radiance fields, ECCV 2022
[2] Cheng Sun et al., Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction, CVPR 2022
[3] Thomas Müller et al., Instant neural graphics primitives with a multiresolution hash encoding, ACM ToG 2022

EgoNeRF - Keys

1. Reconstruct NeRF of large-scale scenes from **egocentric video**

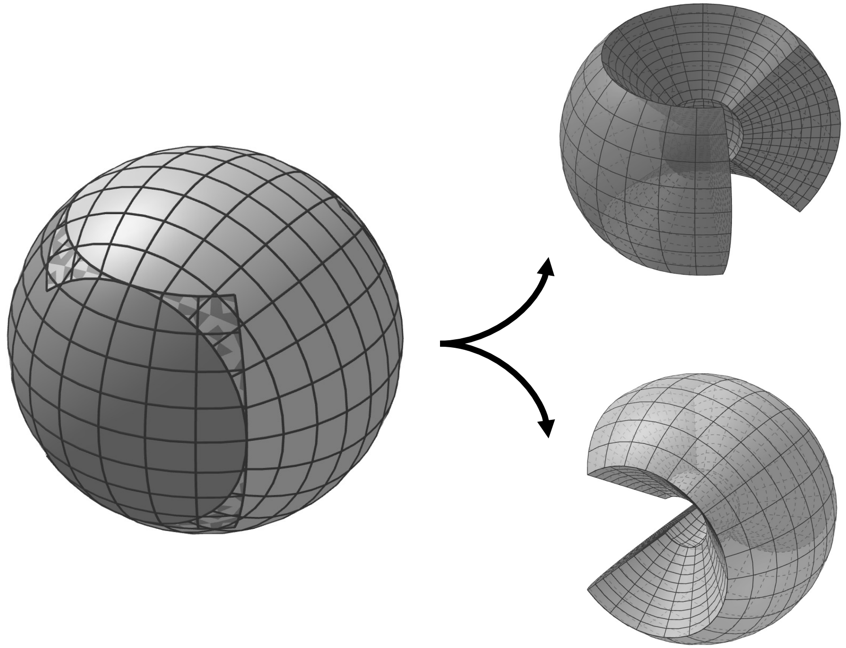
→

egocentric video

free-view rendering

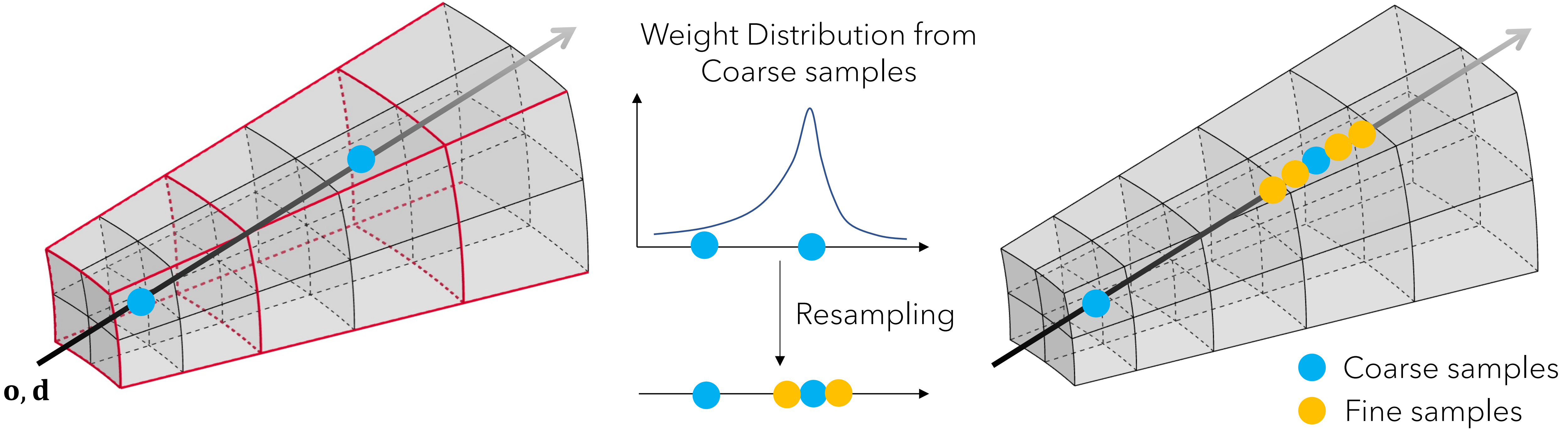
EgoNeRF - Keys

1. Reconstruct NeRF of large-scale scenes from **egocentric video**
2. Grid-based NeRF with **balanced spherical grid**



EgoNeRF - Keys

1. Reconstruct NeRF of large-scale scenes from **egocentric video**
2. Grid-based NeRF with **balanced spherical grid**
3. **Resampling technique** for grid-based methods



EgoNeRF

1. Reconstruct NeRF of large-scale scenes from egocentric video
2. Grid-based NeRF with balanced spherical grid
3. Resampling technique for grid-based methods

Capturing Large-Scale Scenes

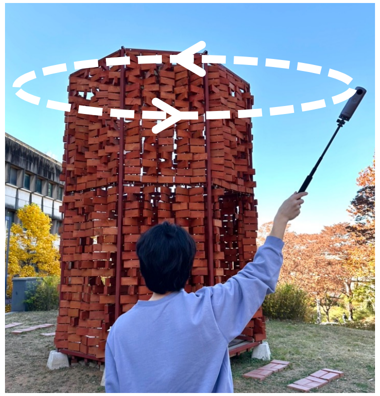
RGB-D Camera [4]

- ✗ Narrow field-of-view
- ✗ Hard to capture large-scale environments in a short period

Multi-camera rigs [5, 6]

- ✗ High-cost
- ✗ Inaccessible to non-expert users

Omnidirectional Video / Egocentric Video



- ✓ Short capturing time (less than 5s)
- ✓ Portable & low-cost
- ✓ Non-expert can capture easily

[4] Angela Dai et al., ScanNet: Richly-Annotated 3D Reconstruction of Indoor Scenes, CVPR 2017
[5] Michael Broxton et al., Deepview immersive light field video, SIGGRAPH 2020 Immersive Pavilion
[6] Albert Parra Pozo et al., An integrated 6dof video camera and system design, ToG 2019

Datasets - *OmniBlender*


BarberShop


Classroom


ItalianFlat


Restroom


BistroBike

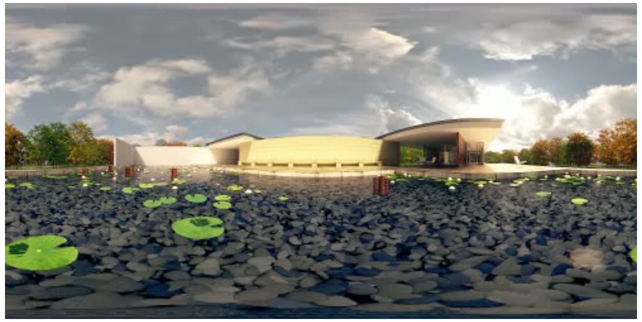

BistroSquare


FisherHut


LoneMonk


LOU


PavilionChair


PavilionPond

Datasets – *Ricoh360*


Bricks


Bridge


BridgeUnder


CatTower


Center

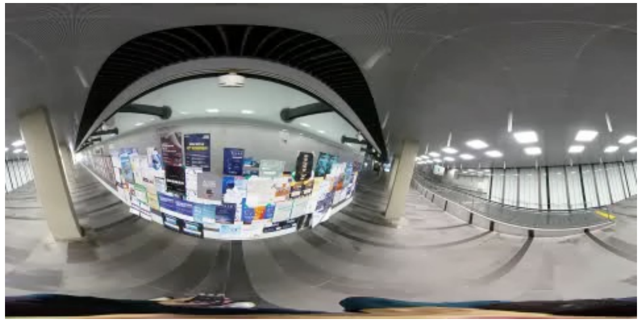

Farm


Flower


GalleryChair


GalleryPillar


Garden


Poster

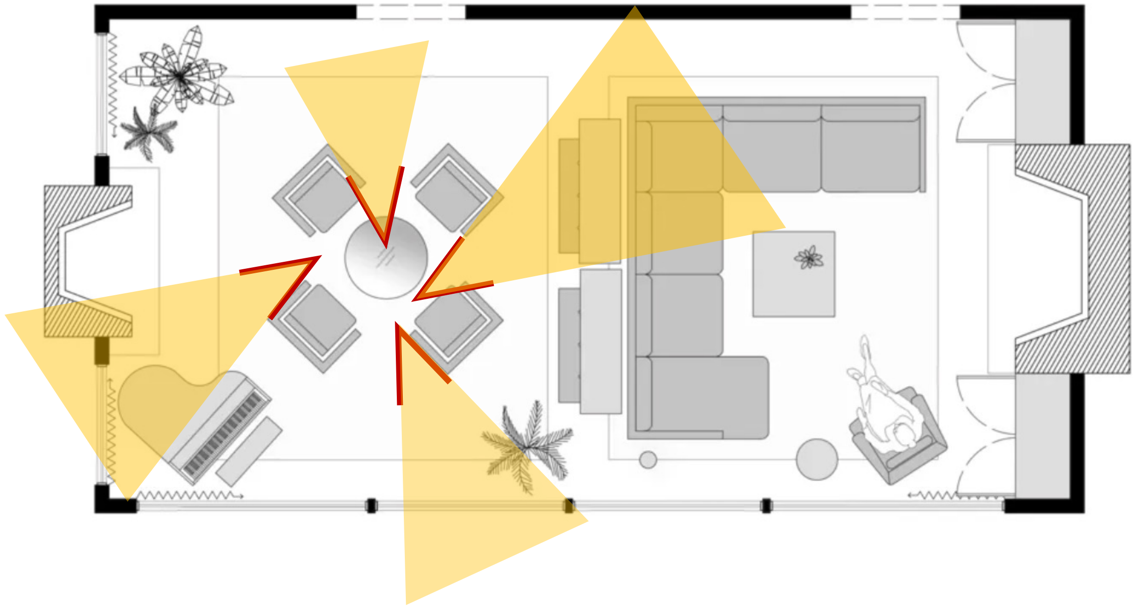
EgoNeRF

1. Reconstruct NeRF of large-scale scenes from egocentric video
2. Grid-based NeRF with balanced spherical grid
3. Resampling technique for grid-based methods

Grid-based Methods in Outward-looking Scene

- Inward-facing views vs Outward-looking views

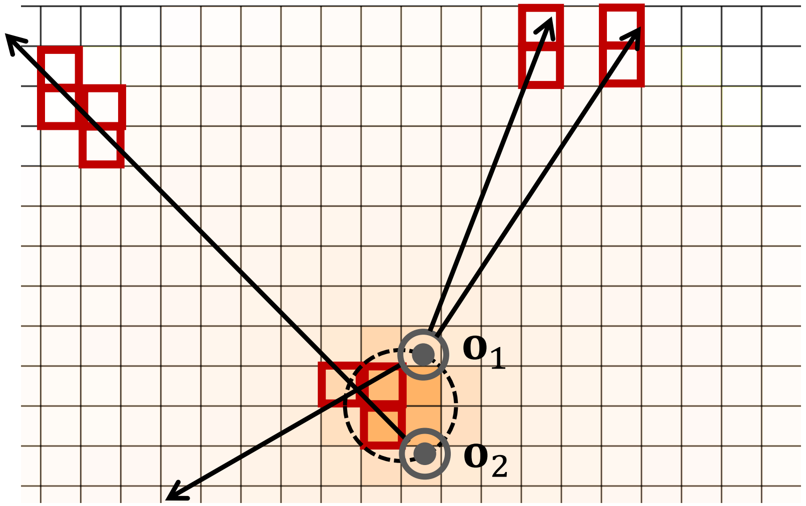
Inward-facing views



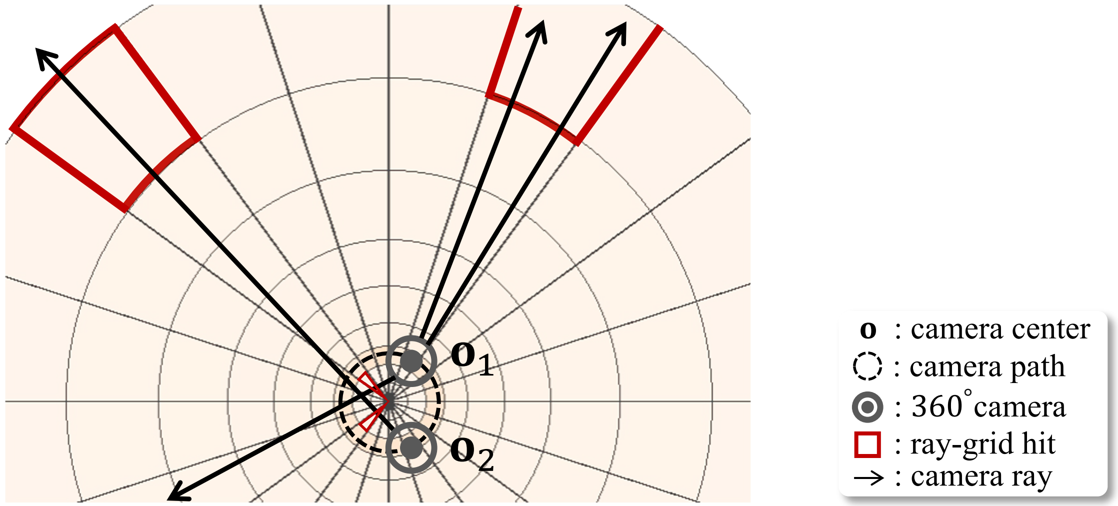
Outward-looking views

Grid-based Methods in Outward-looking Scene

- Cartesian Grid in Outward-looking Scenario
  - Uniform grid size regardless of distance from camera
  - Non-uniform ray-grid hits in outward-looking scenario
- Spherical Grid
  - Resolves aforementioned limitations



Cartesian Grid

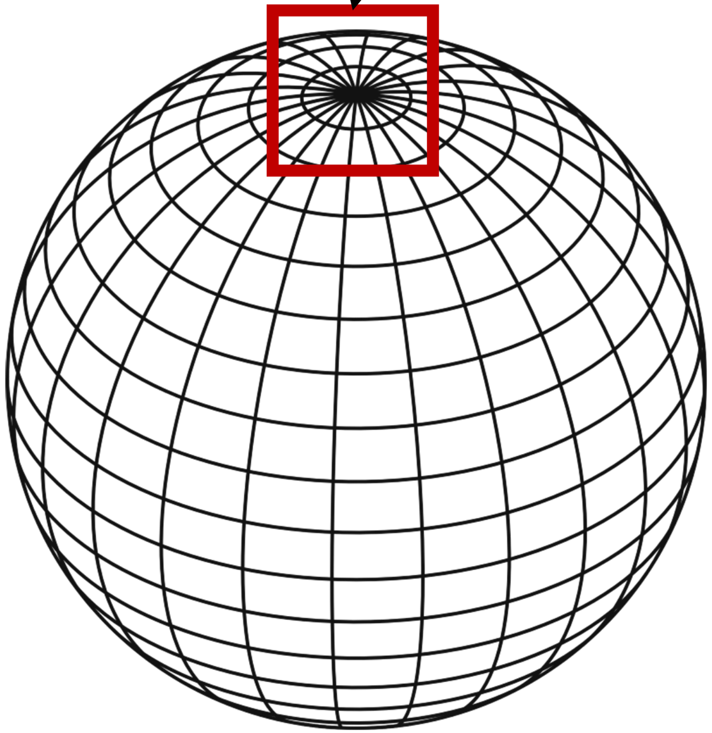


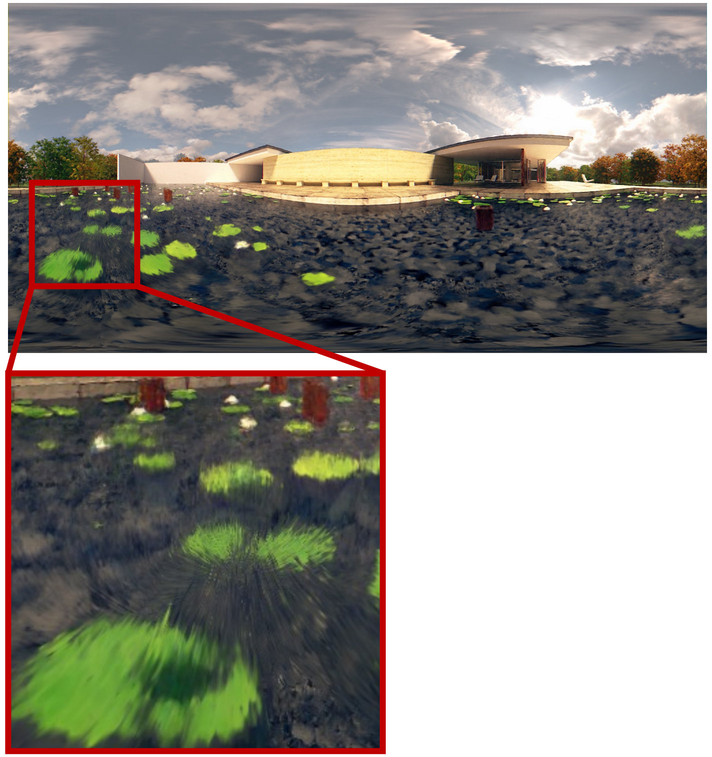
Spherical Grid

Balanced Spherical Grid

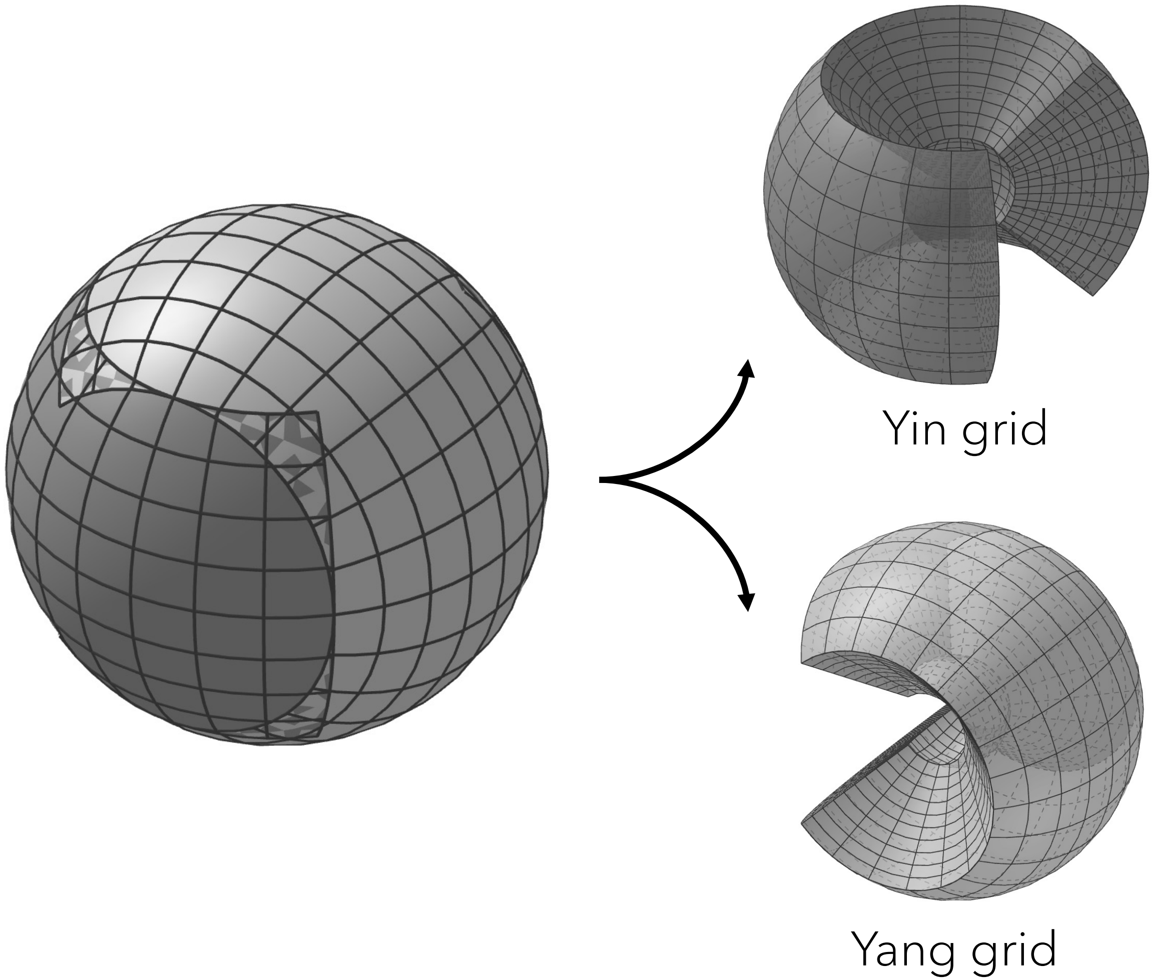
- Naïve Spherical Grid

Singularity at poles, cause artifacts





Balanced Spherical Grid



Angular Partition[7]

Yin grid: $\left(\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}\right) \cap \left(-\frac{3\pi}{4} \leq \phi \leq \frac{3\pi}{4}\right)$

Yang grid: $\begin{bmatrix} x^{\text{Yin}} \\ y^{\text{Yin}} \\ z^{\text{Yin}} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x^{\text{Yang}} \\ y^{\text{Yang}} \\ z^{\text{Yang}} \end{bmatrix}$

Partition: $\Delta\theta^y = \frac{\pi}{2}\frac{1}{N_\theta^y}, \Delta\phi^y = \frac{3\pi}{2}\frac{1}{N_\phi^y}, y \in \{\text{Yin}, \text{Yang}\}$

Radial Partition

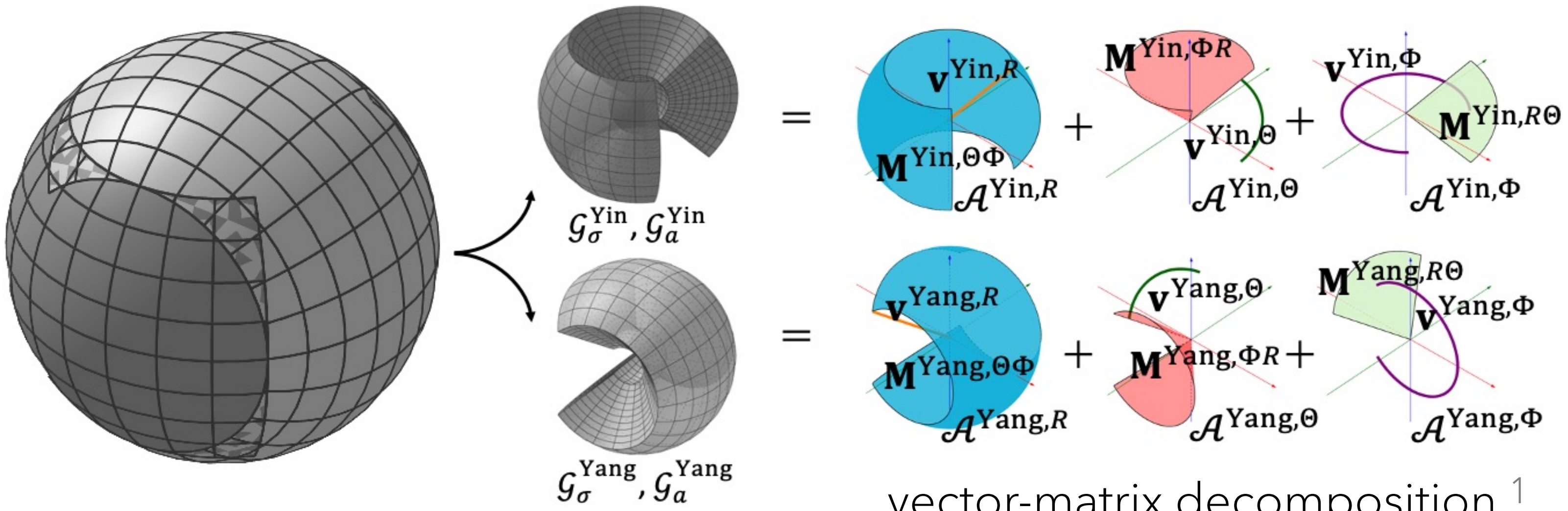
Exponential Partition

$$r_i^y = r_0 k^{i-1}, R_{\max} = r_0 k^{N_r^y - 1}$$

[7] Akira Kageyama and Tetsuya Sato, "yin-yang-grid": An overset grid and spherical geometry, Geochemistry 2004

Balanced Spherical Grid as a Radiance Field

- Consider Explicit Feature Grids as the Mapping Function instead of MLP
  - Density grid: $\mathcal{G}_\sigma \in \mathbb{R}^{2N_r^y \times N_\theta^y \times N_\phi^y}$
  - Appearance grid: $\mathcal{G}_a \in \mathbb{R}^{2N_r^y \times N_\theta^y \times N_\phi^y \times C}$
- Querying Density and Color
  - $\sigma(\mathbf{x}) = \mathcal{T}(\mathcal{G}_\sigma, \mathbf{x}), c(\mathbf{x}, \mathbf{d}) = f_{\text{MLP}}(\mathcal{T}(\mathcal{G}_a, \mathbf{x}), \mathbf{d})$ ($\mathcal{T}$: Trilinear Interpolation)



vector-matrix decomposition [1]

[1] Anpei Chen et al., Tensorf: Tensorial radiance fields, ECCV 2022

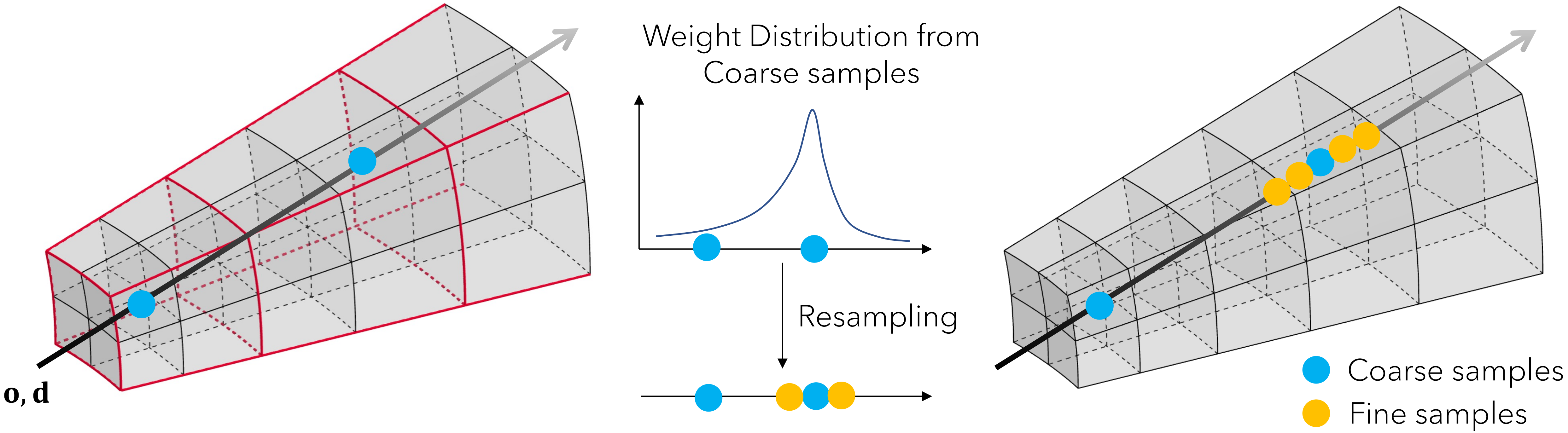
EgoNeRF

1. Reconstruct NeRF of large-scale scenes from egocentric video
2. Grid-based NeRF with balanced spherical grid
3. Resampling technique for grid-based methods

Resampling Technique

- Sample N_c coarse points from a coarse grid
  - Don't need to allocate additional memory for coarse grid

$$\mathcal{G}_\sigma^c = K * \mathcal{G}_\sigma,\ \ K\text{: non-learnable kernel (e.g. avg pool)}$$



Optimization

- Training Objective

$$\hat{C} = \sum_{i=1}^{N} \tau_i \left(1 - e^{-\sigma(\mathbf{x}_i)\delta_i}\right) c(\mathbf{x}_i, \mathbf{d}) + \tau_{N+1} c_{\mathbf{env}}(\mathbf{d})$$

Environment map $\mathcal{E}$

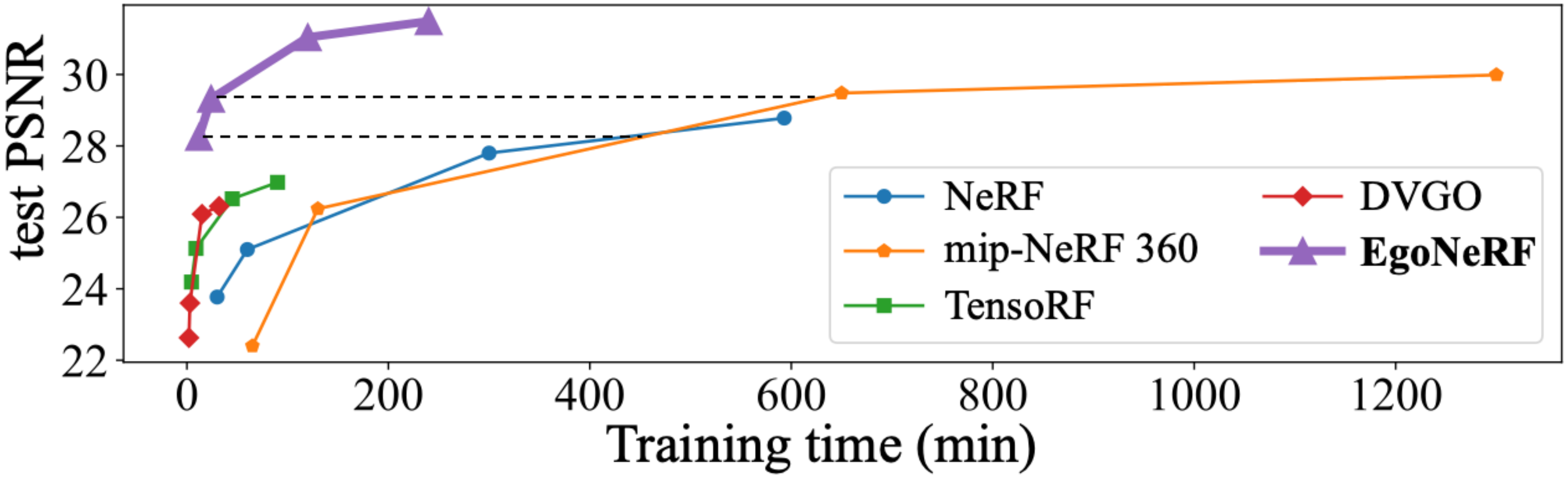
$$\mathcal{L} = \frac{1}{|\mathcal{R}|} \left\| \hat{C}(\mathbf{r}) - C(\boldsymbol{r}) \right\|$$

w/o Env map

w/ Env map

Quantitative Results

- Fast training & rendering time
- High reconstruction quality at large-scale outward-looking scenes



Qualitative Results

2D Map

Camera

Training trajectory

[1] Anpei Chen et al., Tensorf: Tensorial radiance fields, ECCV 2022
[2] Cheng Sun et al., Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction, CVPR 2022
[8] Jonathan T. Barron et al., Mip-nerf360: Unbounded anti-aliased neural radiance fields, CVPR 2022

Conclusion

- EgoNeRF takes casual egocentric video of large scene as input
- EgoNeRF takes ad

s (fast speed) and improved perform cal Grid in large scene

Changwoon Choi

Sang Min Kim

Young Min Kim

3D LAB VISION @

code&data

project page