ManifoldNeRF: View-dependent Image Feature Supervision for Few-shot Neural Radiance Fields

○ Daiju Kanaoka[1,2], Motoharu Sonogashira[2], Hakaru Tamukoh[1,3], Yasutomo Kawanishi[2]
[1]Kyushu Institute of Technology
[2]RIKEN Guardian Robot Project
[3]Research Center for Neuromorphic AI Hardware

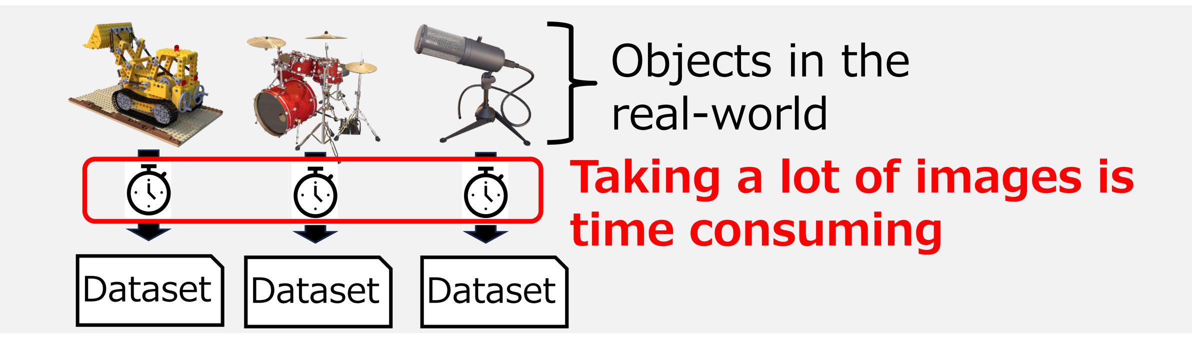
BMVC 2023

Introduction

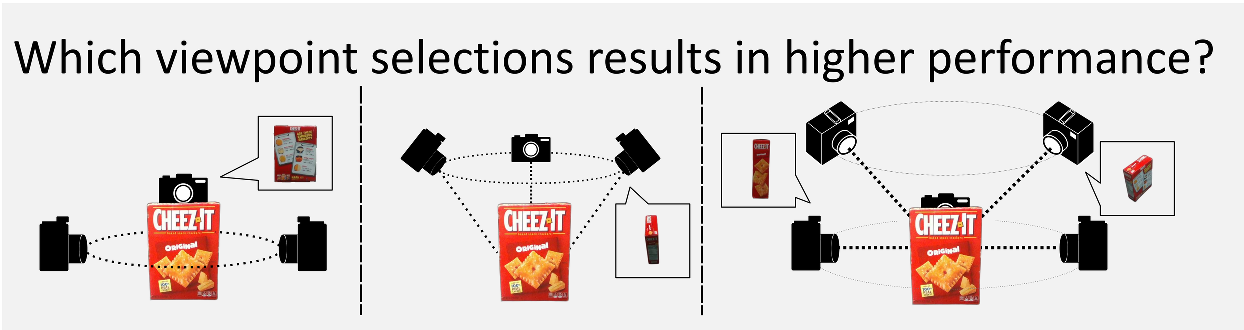
It is difficult to use NeRF[1] in rea-world applications bceause of several problems.

Example of the application) A robot observes an object and obtains its 3D representation.

Problem1: Requires time to take a lot of images.



Problem2: Appropriate viewpoints are nonclear.



Objectives

- Obtain high quality 3D representation from only a few images.
- Clarify which viewpoint patterns are highly performance for the proposed methods considering real-world applications.

[1] B.Mildenhall et al., ECCV, 2020.

DietNeRF[2]

- Assumption: **Feature vectors of images of an object at arbitrary viewpoints in the same scene should be consistent.**
- Loss calculation at unknown viewpoints are impossible since there are no GT images.
- The method enables loss calculation anywhere using the feature vector.

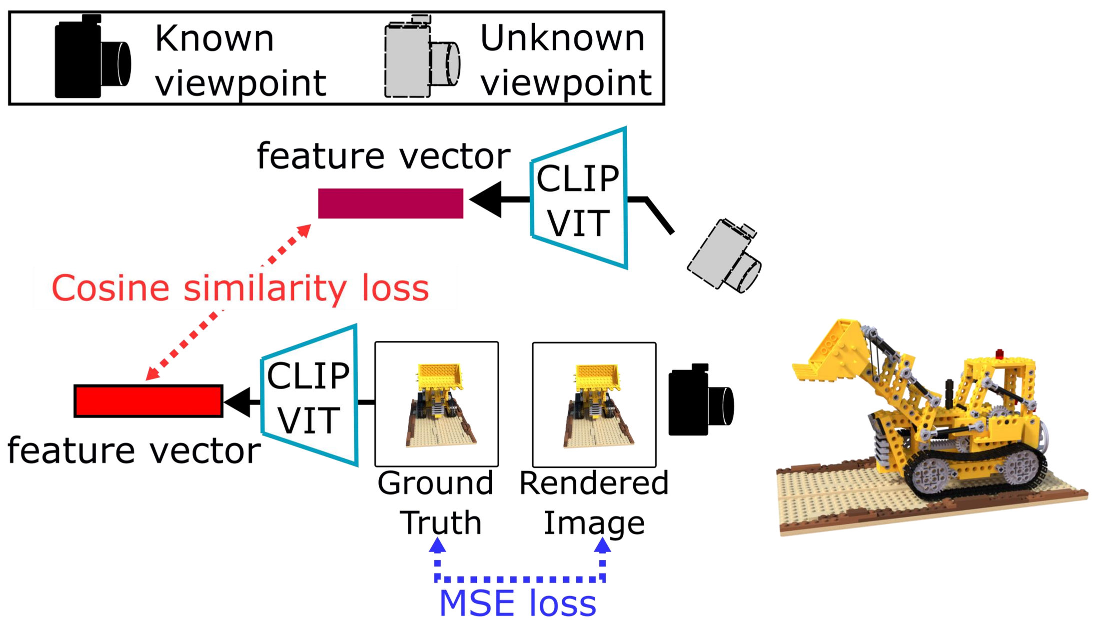
Loss function

$$L = L_{MSE} + \lambda(1 - L_{cosine})$$

L_{MSE} : Mean squared error

L_{cosine} : Cosine similarity

λ : Scaling parameter

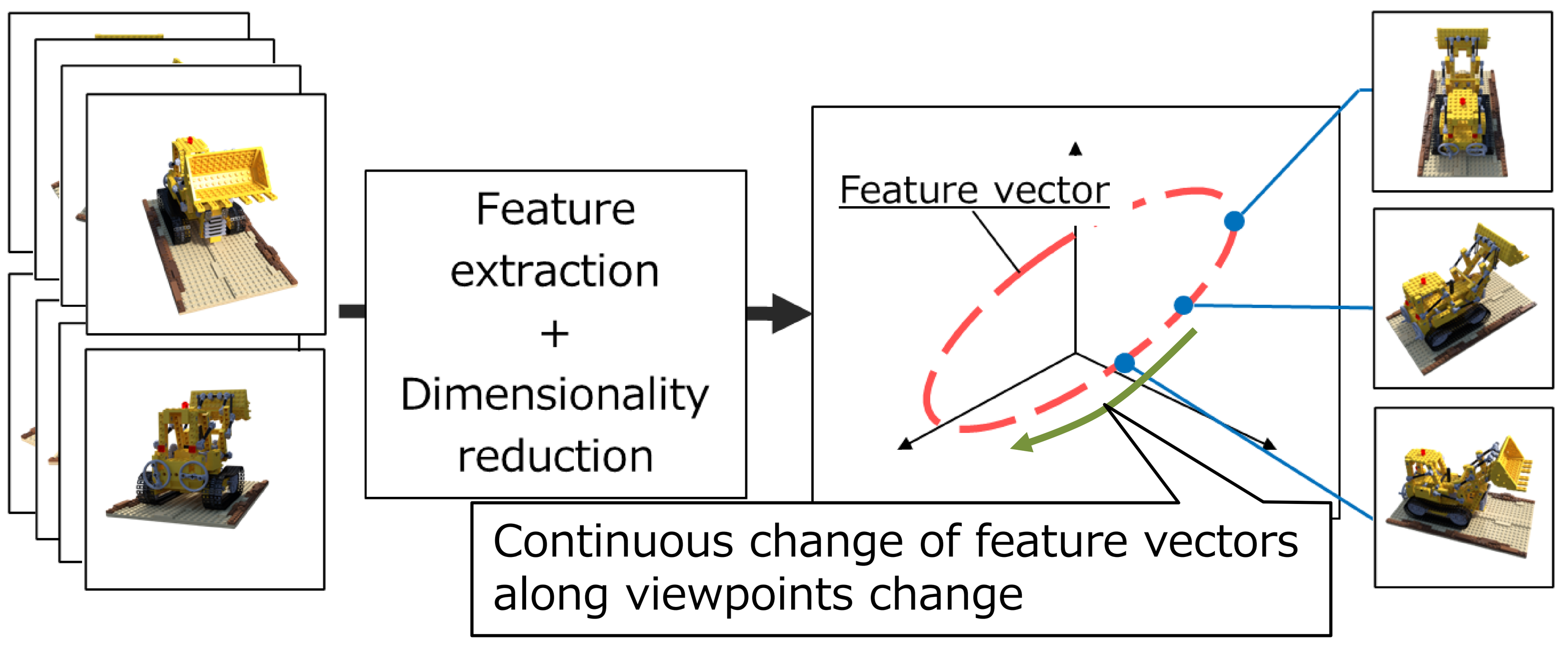


Are the feature vectors of all viewpoints really similar? ➡ No!

[2] A.Jain et al., ICCV, 2021.

Idea from Parametric Eigenspace method[3]

The feature vector changes continuously as the viewpoint of the image changes continuously.



Features at novel viewpoints can be interpolated along with the viewpoint change.

[3] H. Murase and S. K. Nayar, IJCV, 1995

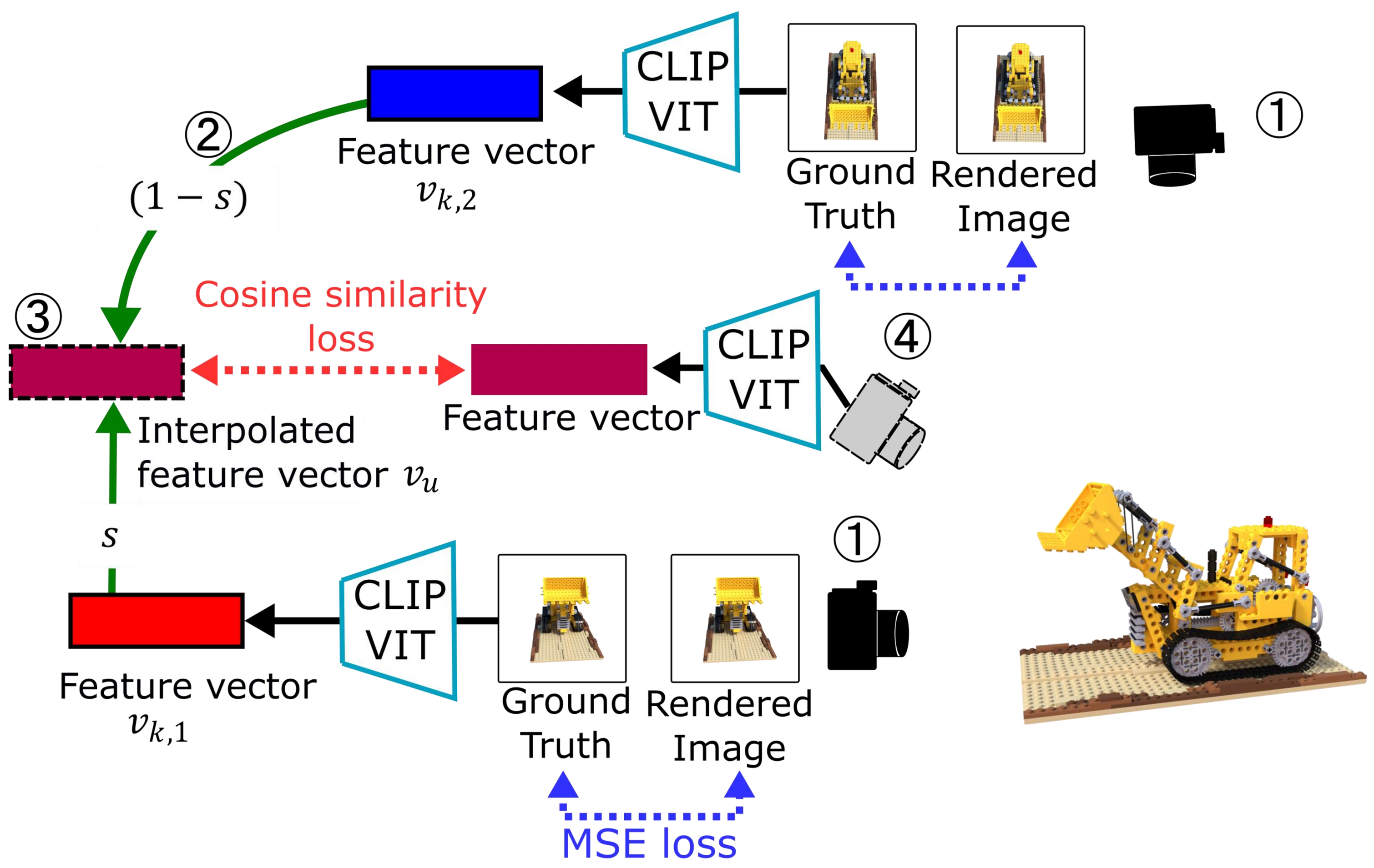
Proposed method: ManifoldNeRF

- Assumption: **The feature vector varies continuously as the viewpoint changes, the interpolated feature can be used as GT features.**

➡ This enables loss calculation using the feature vectors close to the actual feature vector at unknown viewpoints.

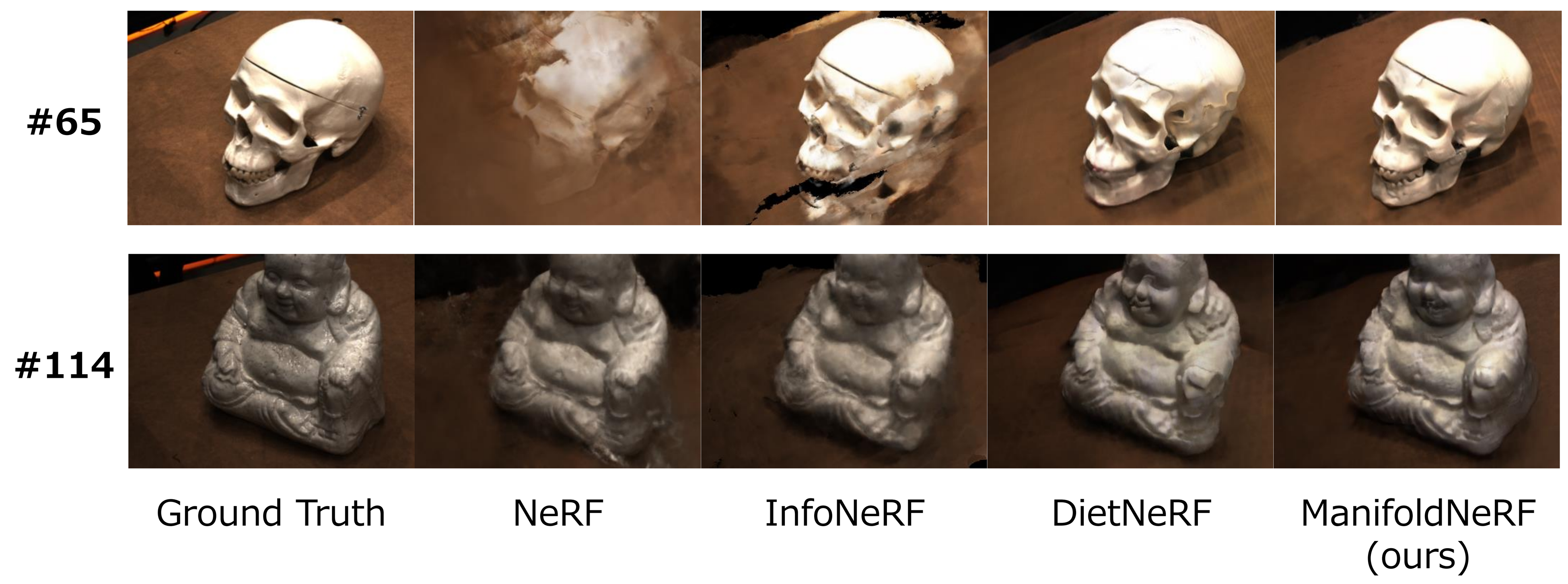
The procedure of calculating pseudo ground truth

①Select two nearby known viewpoints
②Sample a random value for division ratio s in (0,1)
③Calculate pseudo ground truth by linear interpolation
④Put a camera virtually at the position of ration s



Result of DTU MVS dataset

In this experiment, we selected 8 images sampled from the dataset for training each object.

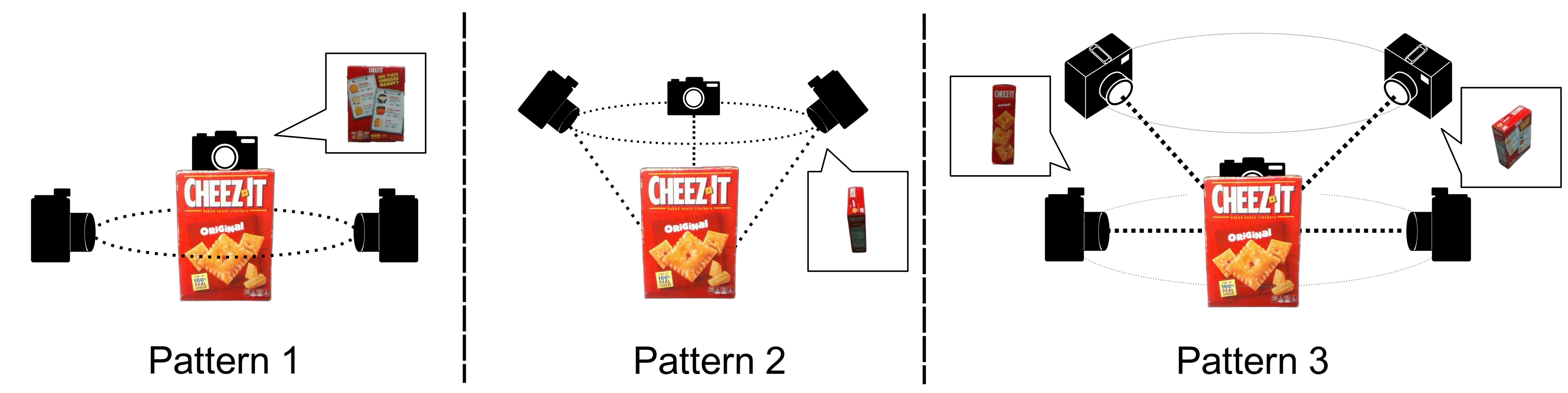


#65	PSNR↑	SSIM↑	LPIPS↓
NeRF[1]	11.970	0.481	0.527
InfoNeRF[4]	14.786	0.484	0.431
DietNeRF[2]	20.883	0.698	0.352
ManifoldNeRF (ours)	**22.197**	**0.702**	**0.302**

#114	PSNR↑	SSIM↑	LPIPS↓
NeRF[1]	18.691	0.636	0.396
InfoNeRF[4]	21.382	0.611	0.364
DietNeRF[2]	20.861	0.673	0.337
ManifoldNeRF (ours)	**23.202**	**0.742**	**0.299**

[4] M.Kim et al., CVPR, 2022.

Evaluation of viewpoint selections for real-world application



ManifoldNeRF	PSNR↑	SSIM↑	LPIPS↓
Pattern1	17.983	0.849	0.147
Pattern2	21.143	0.871	0.114
Pattern3	**23.203**	**0.899**	**0.074**

Viewpoints should be selected uniformly spaced hemispherically around an object for high performance