OrthoPlanes: A Novel Representation for Better 3D-Awareness of GANs

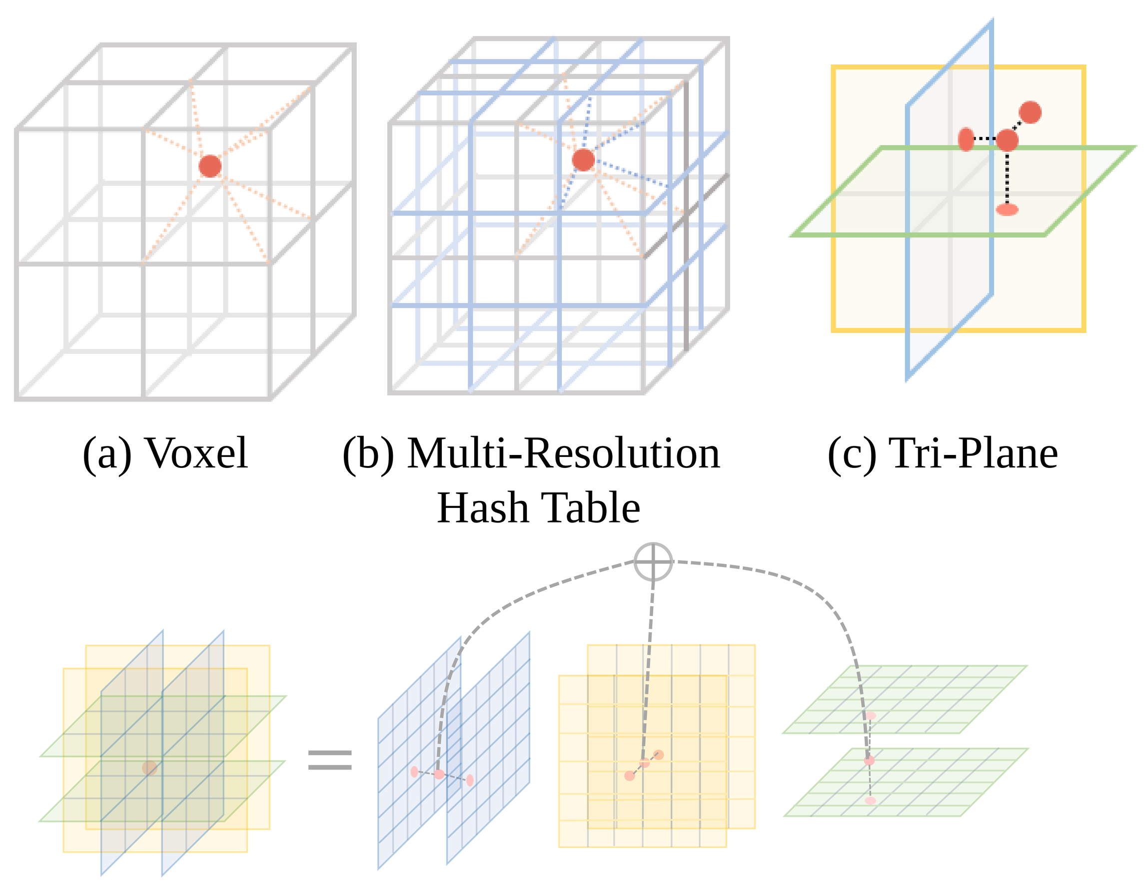
Honglin He,[*,1,2] Zhuoqian Yang,[*,1,3] Shikai Li,[1] Bo Dai,[1] Wayne Wu,[+,1]

[1]Shanghai AI Laboratory [2]Tsinghua University [3]School of Computer and Communication Sciences, EPFL

Motivation and Approach

Inverse rendering aims at recovering 3D world from 2D images. In this work, we are interested in constructing a powerful 3D representation that can be applied to a wide range of tasks, especially 3D content generation.

Current approaches that widely used, are either inefficient (Deep Voxel), structurally complex (Hash Grid) or spatially ambiguous (Triplane).



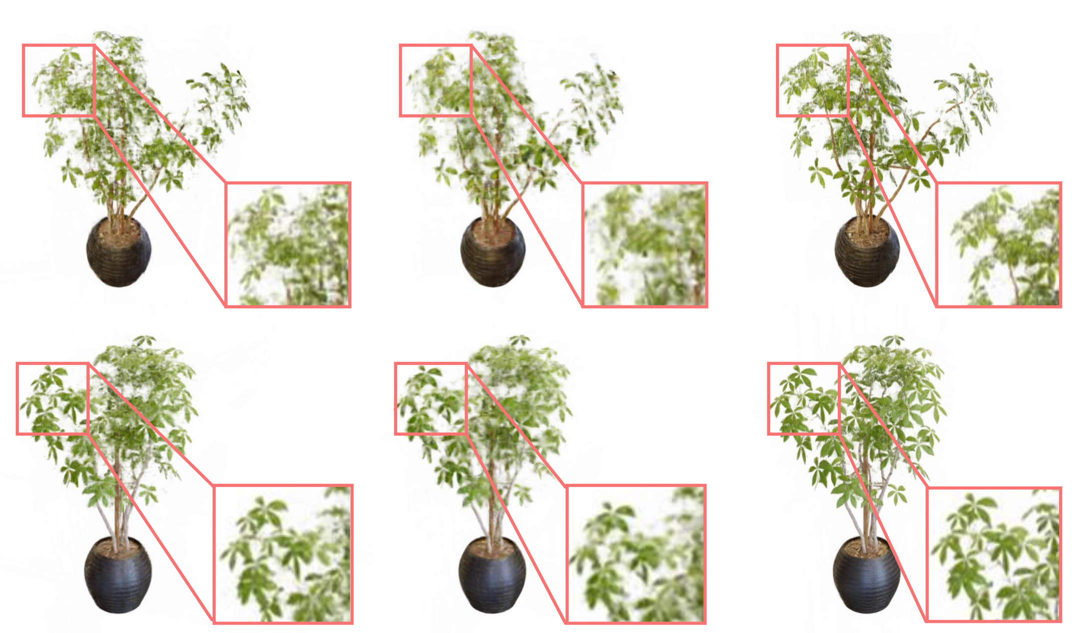
(a) Voxel (b) Multi-Resolution Hash Table (c) Tri-Plane

(d) Ours (OrthoPlanes)

We propose an efficient and expressive 3D representation, *Orthoplanes*, in which scenes are stored by three orthogonal plane-groups , each plane is anchored along its projection direction, and distributed evenly along the axis, storing compressed features of the scene while preserving information associated with projection distance.

NeRF Overfitting with Orthoplanes

We demonstrate the expressiveness of orthoplanes with single-scene overfitting (SSO) experiments. Our model archives better PSNR in all cases compared to triplane. And our approach recovers the color tone of *ficus*, especially asymmetric parts like sparse leaves and protruding branches.

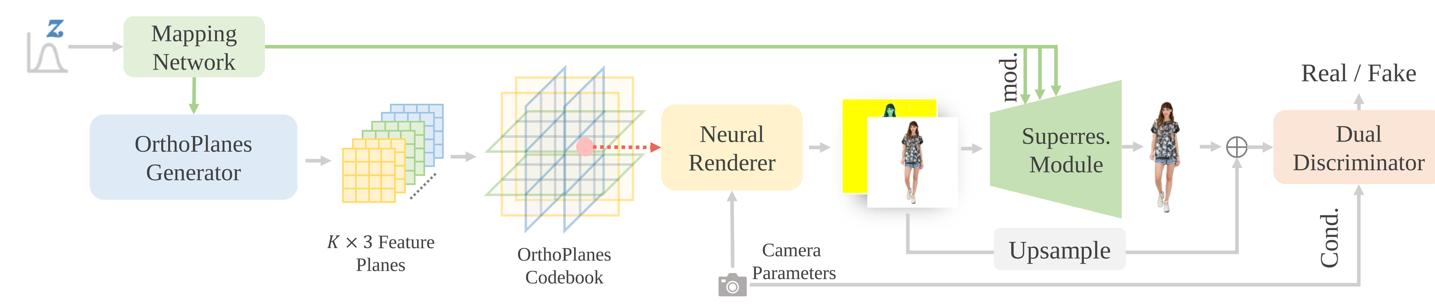


Triplane- 3 x 2H x 2H x C | Triplane- 3 x H x H x 4C | OrthoPlanes- 3 x 4 x H x H x C

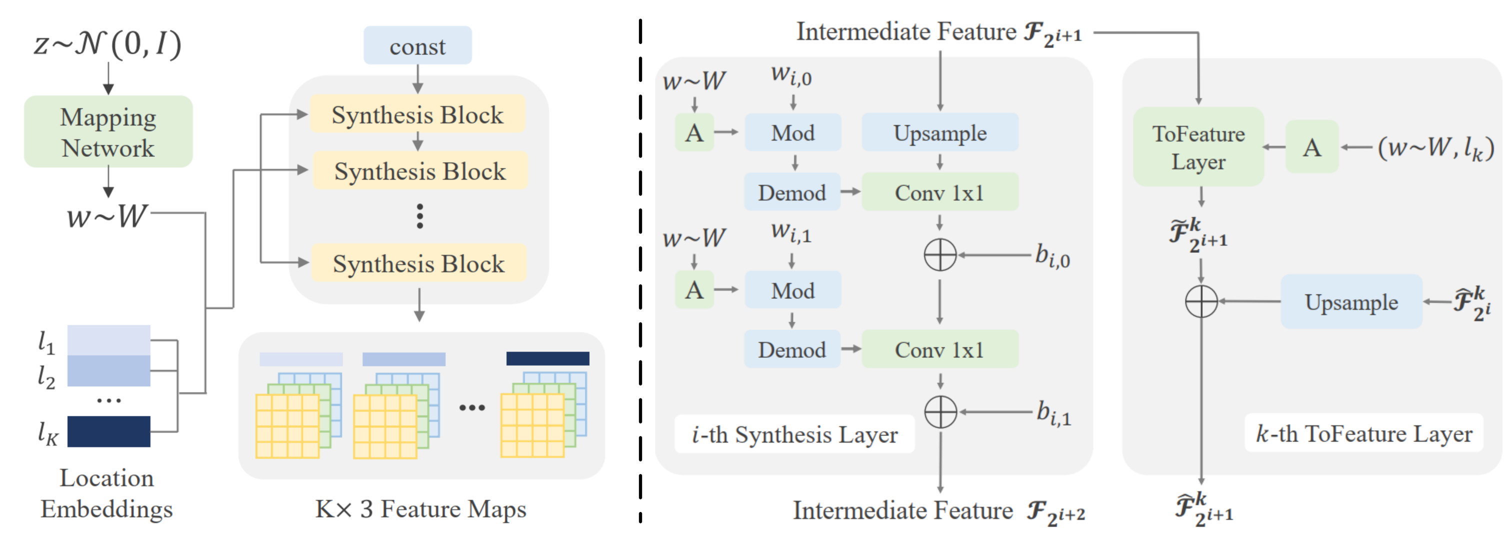
	lego	*ficus*	*family*	*caterpillar*
Tri-plane $(H^2, 4C)$	26.56	23.13	30.59	22.03
Tri-plane $((2H)^2, C)$	28.66	23.50	30.48	22.32
Ours $(H^2, C; K = 4)$	31.12	26.83	32.06	23.93

3D GAN with Orthoplanes

We use orthoplanes in both of unconditional and conditional 3D-aware image synthesis, and give our unconditional GAN framework here.



Based on StlyeGAN2, we introduce the design of location-embedding, modulating the feature images generated by the backbone, which is as efficient as EG3D.



Results: Qualitative Results on Unconditional GAN Task

Generators trained on diverse datasets all obtain view-consistency rendered results with fine geometry details.

Results: Qualitative Results on Conditional GAN Task

Based on inverse LBS, we use Orthoplanes as the proxy of pose-canonical human avatar, we can generate pose-controllable avatars based on the model trained on 2D dataset SHHQ.

Results: Quantitative Comparision

We compare our approach with other State-Of-The-Art methods on diverse datasets.

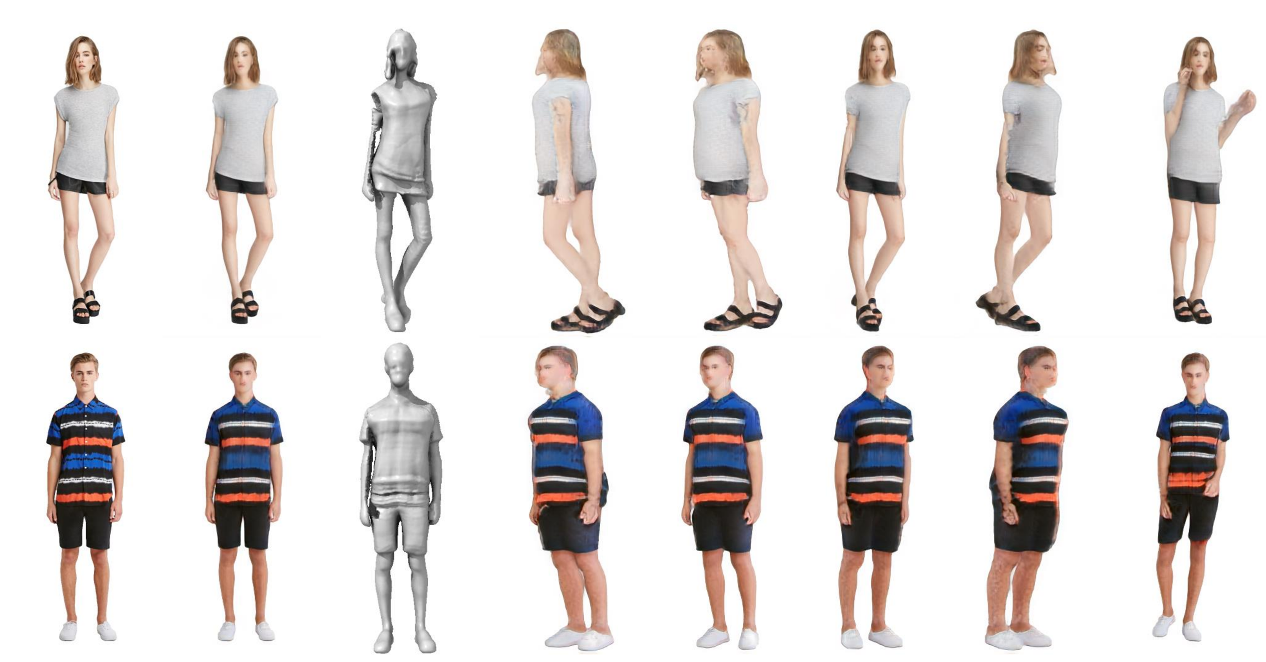
	FFHQ			AFHQv2-Cat	
	FID ↓	KID ↓	ID ↑	FID ↓	KID ↓
pi-GAN 128^2	29.9	35.73	0.67	16.0	14.92
StyleSDF 256^2	11.5	2.65	-	12.8*	4.47*
StyleNeRF 512^2	7.80	2.20	-	13.2*	3.60*
GMPI 512^2	8.29	4.54	0.74	7.79	4.74
EG3D 512^2	4.70‡	1.32	0.77	2.77†	0.41†
Ours-S 512^2	4.11‡	1.05	0.71	2.82†	0.46†
Ours-R 512^2	4.01‡	1.23	0.73	-	-

- ∗ Metrics are reported on the whole AFHQv2 dataset.
- † Trained with adaptive discriminator augmentation.
- ‡ Trained with camera condition.

SHHQ	FID ↓	KID ↓	Depth ↓
StyleSDF 512^2	33.29	25.2	0.036
StyleNeRF 512^2	7.60	3.96	-
EG3D 512^2	5.79	2.26	0.082
Ours 512^2	4.18	2.05	0.082

Applications: Single Image Reconstruction

Through pivotal tuning inversion (PTI) , our model can be used to reconstruct 3D shapes from single views of objects. In this manuscript, we provide results of human body reconstruction with known pose conditions. We use off-the-shelf pose estimator to extract body-pose parameters and camera parameters.



Input | Recon RGB | Recon Shape | Novel Views | Novel Pose

References

[1] Eric R Chan, Connor Z Lin, Matthew A Chan, Koki Nagano, Boxiao Pan, Shalini De Mello, Orazio Gallo, Leonidas J Guibas, Jonathan Tremblay, Sameh Khamis, et al. Efficient geometry-aware 3d generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16123–16133, 2022.

[2] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

* Equal contribution. + Correpondence author.