World Tracing

Generative Pixel-Aligned Geometry Beyond the Visible

Hao Zhang[1,2*] Mohamed El Banani[1] Jen-Hao Cheng[1] Paul Zhang[1] Yi Hua[1] Ben Mildenhall[1] Christoph Lassner[1] Narendra Ahuja[2] Gengshan Yang[1]

[1] World Labs [2] University of Illinois Urbana-Champaign * Internship at World Labs

Paper · demos WeChat

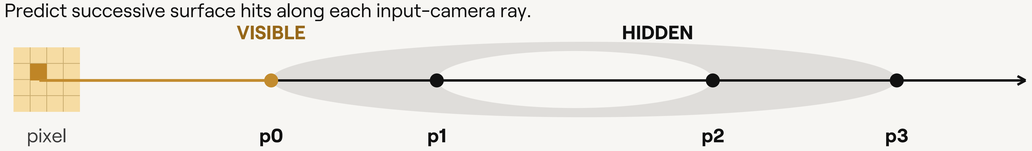
World Labs · University of Illinois Urbana-Champaign

NeurIPS 2026 · SPOTLIGHT

01 / GENERATIVE RAY TRACING

See through the image.

Predict successive surface hits along each input-camera ray.



INVALID DEPTH FILLING

p0	p1	p2	p3	–	–
p0	p1	p2	p3	p3	p3

No hit? Repeat the last valid XYZ.
Same surface set; dense foreground targets.

Why no validity mask?

8.14% → 0.60%

Valid area / full image: L0 → L5 (paper)

- Sparse masks can collapse.
- XYZ / validity gradients can conflict.

No validity head; read termination from layer gaps.

Input alpha still defines support; outside-alpha endpoint loss is masked.

Frozen MoGe + noisy XYZ → WT-DiT: Layer · Ray · Global attention
Direct XYZ flow · Layer-aware noise · RGBD mixing · Temporal extension

02 / SCENE COMPLETION

Reveal what the image hides.

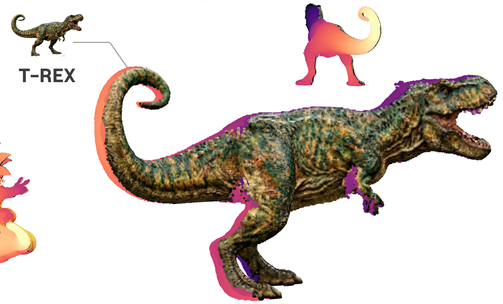
Same camera and scale. RGB: visible · magma: hidden depth.



03 / OBJECTS

Quite a cast. One image each.









04 / GEOMETRY IN MOTION

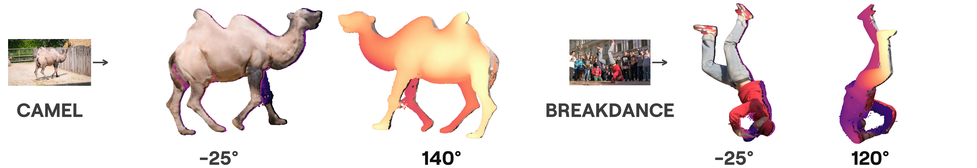
WT-D · one video, evolving multilayer geometry



REAL CAPTURE / DAVIS

Real inputs. Two inspection angles per prediction.

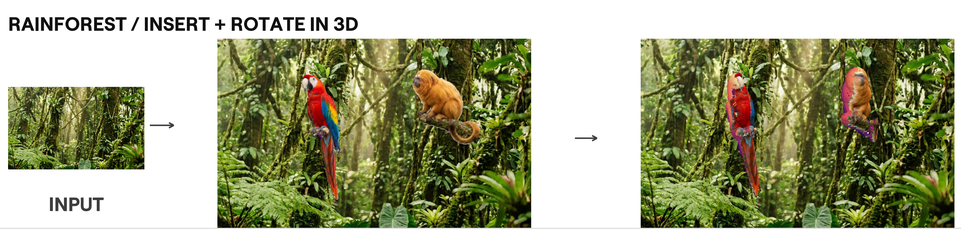
Foreground masks provided



05 / Give the scene new company.

3D SCENE EDITING · Place in 2D → align all 6 layers

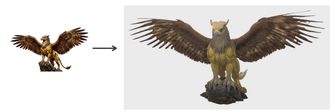




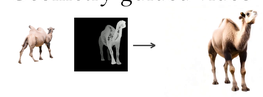
05 / MAKE IT USEFUL

Large: oblique · small: reverse

Textured meshes



Geometry-guided video



06 / EXPLORE THE LAYERS

Beyond the visible.

Inspect · edit · animate

13-inch iPad · landscape

34 × 26 cm clear mounting area

Display: layer termination + outlier filtering · magma: hidden depth · editing: MoGe + WT.

Paper: WT mean 8 · AbsRel .0089 · CD-L1 .0243