Visual Recreation

with *Data* and *Physics*

Shenlong Wang

Apr 2, 2024

Immersive experience for “what if” questions.

What if…

KITTI-360

Immersive experience for ”what if” questions.

What if… driving at night?

Immersive experience for ”what if” questions.

What if… jaywalkers present?

KITTI-360

Immersive experience for ”what if” questions.

What if… driving on a smoggy day?

Immersive experience for ”what if” questions.

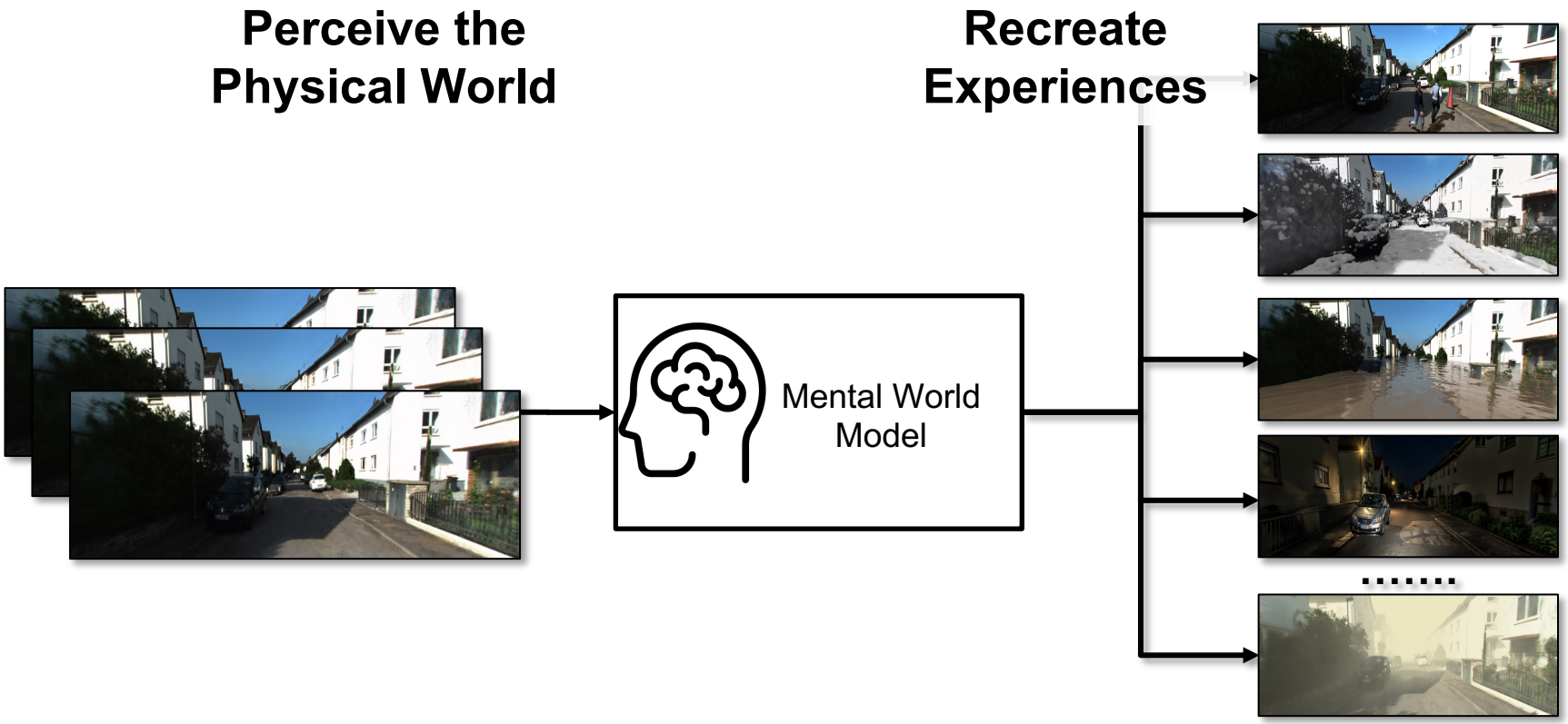
What if… the street is covered by snow?

KITTI-360

Immersive experience for ”what if” questions.

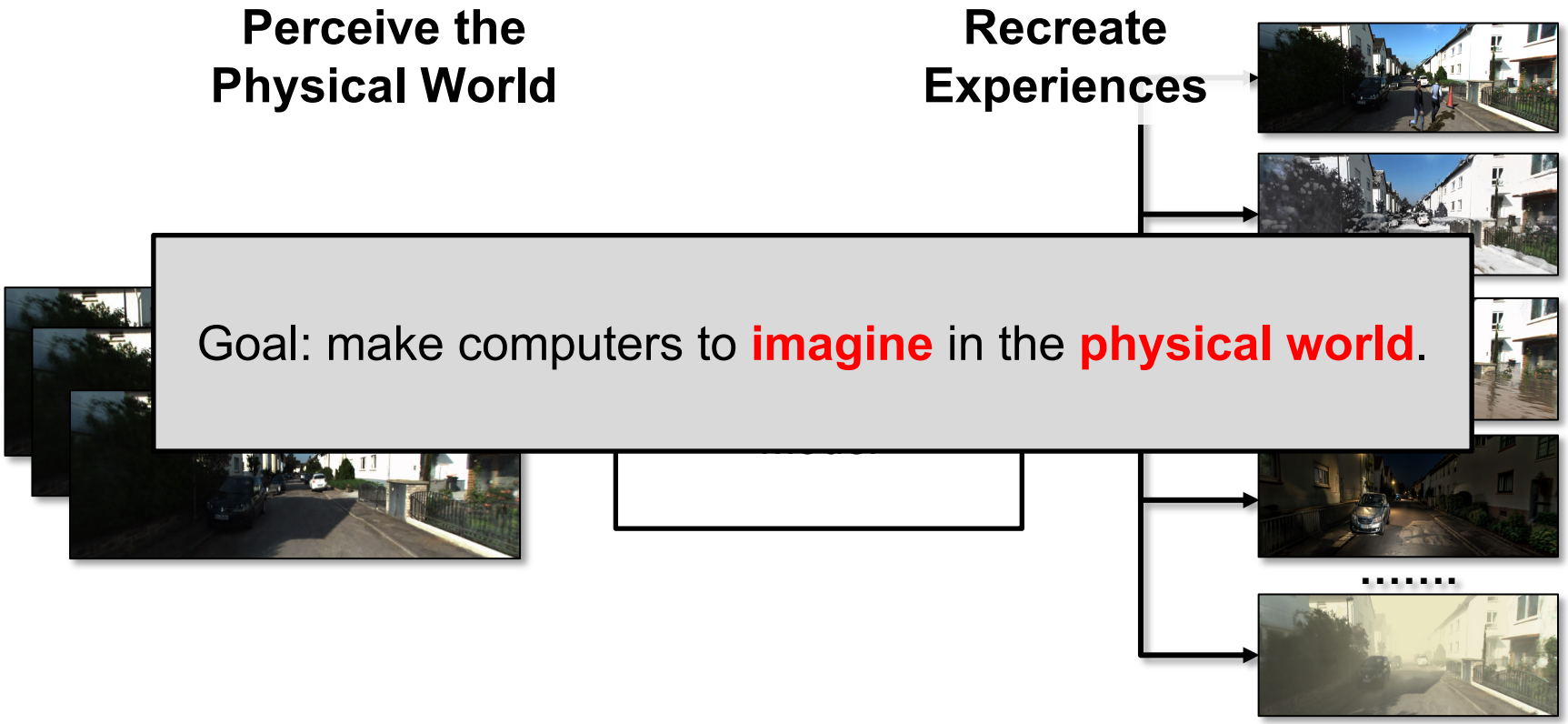
What if… the street is flooded?

Recreative mind [1]



Recreative imagination: an ability to experience or think about the world from a perspective different from the one that experience presents -- Gregory Currie and Ian Ravenscroft (2002)

Recreative mind [1]

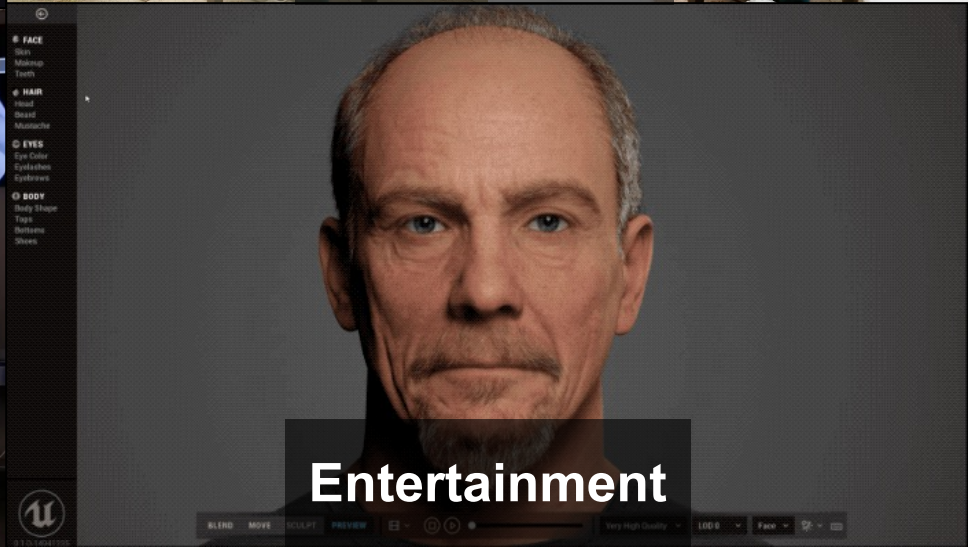


Recreative imagination: an ability to experience or think about the world from a perspective different from the one that experience presents -- Gregory Currie and Ian Ravenscroft (2002)

Teleconference

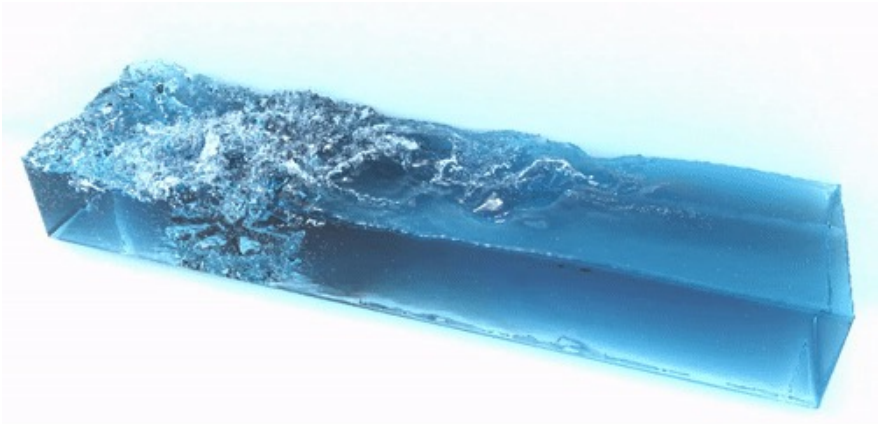
origin
flood
smog
snow
Simulation

Healthcare


FACE
Skin
Makeup
Teeth
HAIR
Head
Beard
Mustache
EYES
Eye Color
Eyelashes
Eyebrows
BODY
Body Shape
Tops
Bottoms
Shoes
Entertainment

What we have done so far

Physical simulation can simulate, but it is not based on real scenes, and necessitates user-created assets.



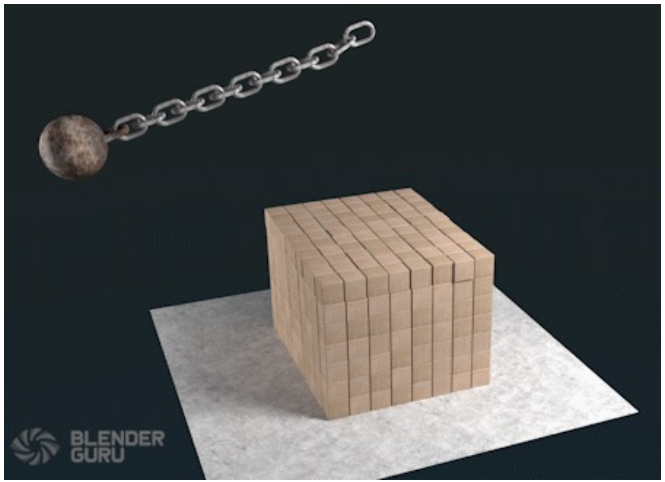


Image credit: SPlisHSPlasH, BlenderGuru

What have done so far

Video Effects (VFX): can be ultra realistic but requires professionals.

Image credit: Life of PI, Game of Thrones

What we have done so far

Generative models can edit but are data-hungry and lack physical grounding.

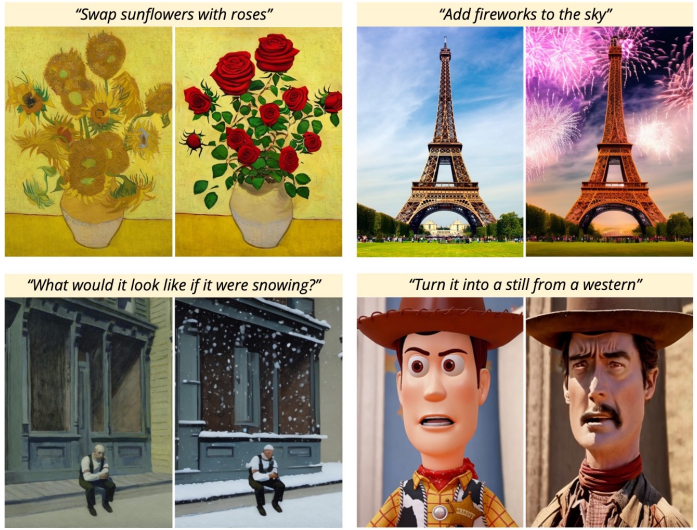


Image credit: Instruct Pix2Pix, Instruct NeRF2NeRF, Runway Gen-2

What we have done so far

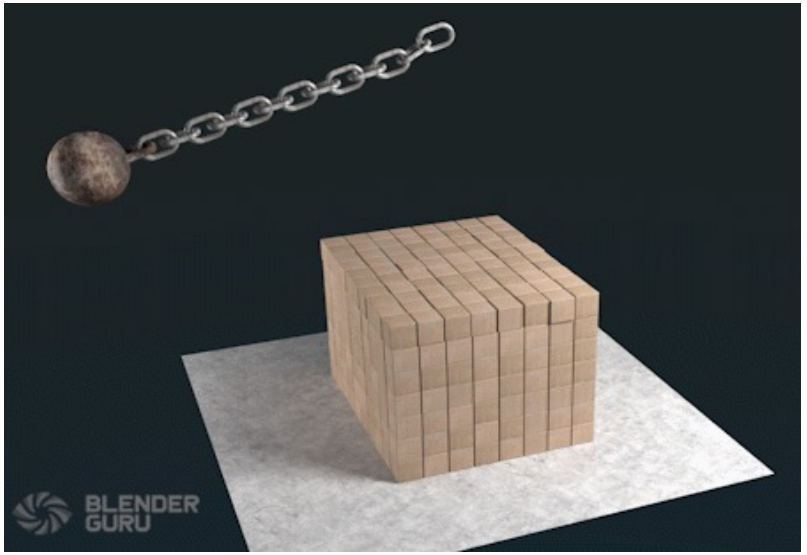
(Neural) 3D modeling can encode rich appearance, geometry and semantics, but has limited editability.

Image credit: 3D Gaussian Splatting, Panoptic Lifting

Key insights

Data-driven

Model-based



Key insights

Data-driven

Model-based

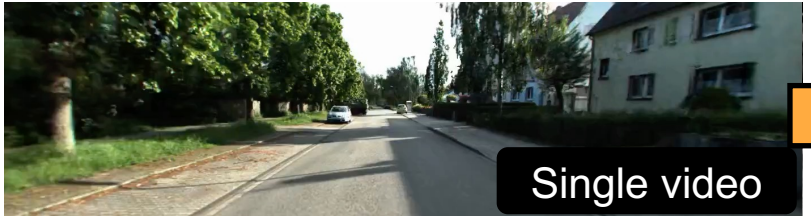
Approach: **data-driven** and **model-based** knowledge for simulation over real-world data.

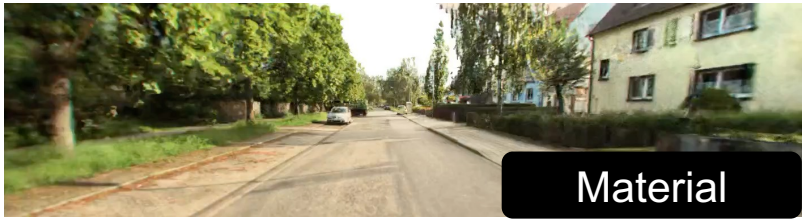
Can we simulate from and to the reality?

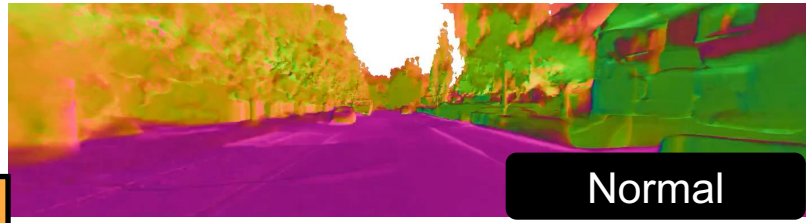
Lighting

Dynamics

Relighting



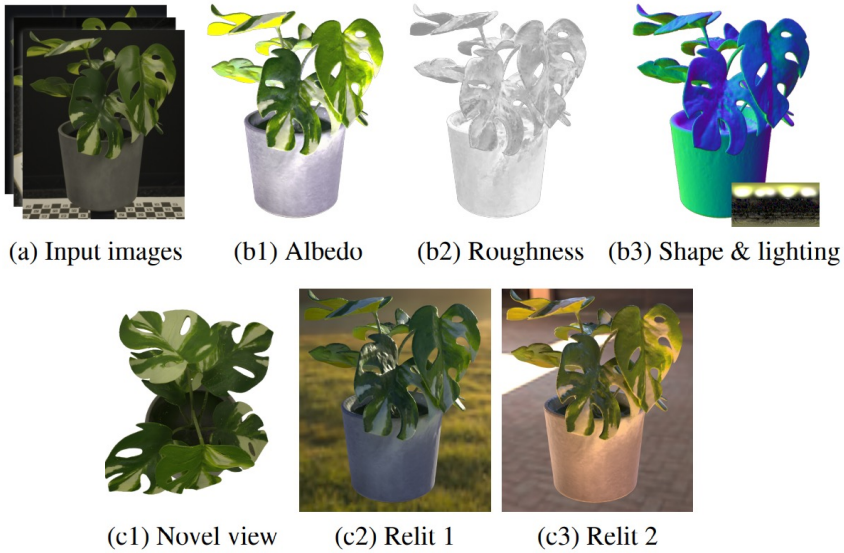




https://urbaninverserendering.github.io/

Key challenge: relightable scene representation

Object-level, surrounding views



Multi-illumination or known geometry

Image credit: NeuralPIL, NeRFOSR

Key challenge: relightable scene representation

Object-level, surrounding views

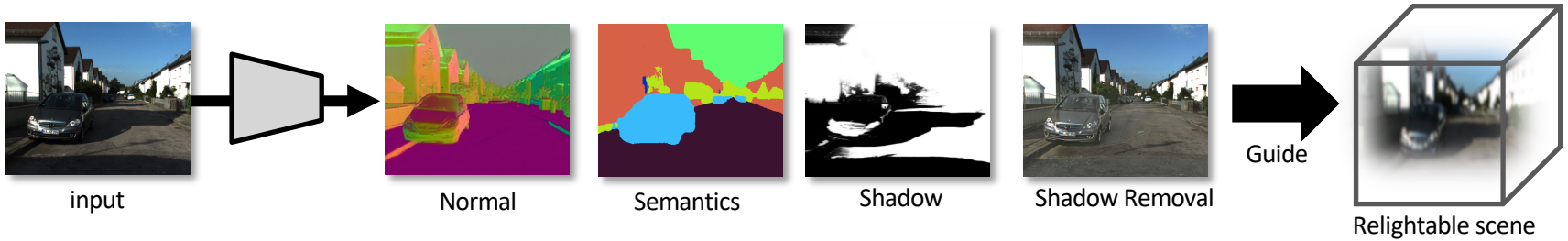
Multi-illumination or known geometry

Goal: **relightable** large scene models from a single monocular video.

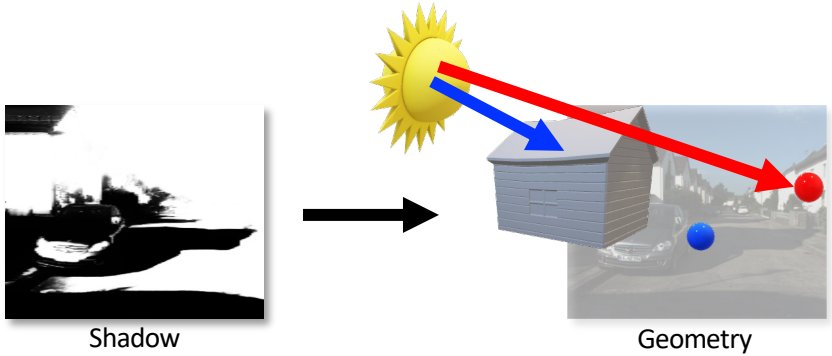
Image credit: NeuralPIL, NeRFOSR

UrbanIR: Urban Scene Inverse Rendering

- **Data-driven monocular cues** provide strong prior / regularization



- Explicit modeling **physics** of **shadow** improves geometry



Zhi-Hao Lin, Bohan Liu, Yi-Ting Hu, Anand Bhattad, David Forsyth, Jia-Bin Huang, Shenlong Wang. UrbanIR: Large-Scale Urban Scene Inverse Rendering from a Single Video, arXiv. 2023

Reconstruction

Reconstruction (Day)

Night simulation

IRIS: Inverse Rendering of Indoor Scenes from Low Dynamic Range Images by Zhi-Hao Lin, Jia-Bin Huang, Zhengqin Li, Zhao Dong, Christian Richardt, Tuotuo Li, Michael Zollhöfer, Johannes Kopf, Shenlong Wang, Changil Kim, arXiv, https://irisldr.github.io/

Can we model and simulate weather effects?

ClimateNeRF

https://climatenerf.github.io/

Extreme Weather Synthesis in NeRF

Snow

Climate Impact

Style image

Multi-view Input Images

Yuan Li*, Zhi-Hao Lin*, David Forsyth, Jia-Bin Huang, Shenlong Wang, "ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis", ICCV 2023.

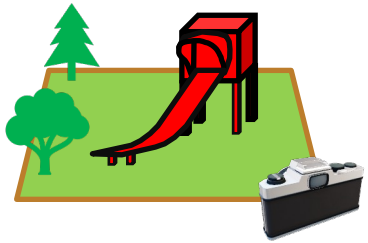
Extreme Weather Synthesis in NeRF

Snow

Climate Impact

Style image

Multi-view Input Images



3D Scene Modeling

Yuan Li*, Zhi-Hao Lin*, David Forsyth, Jia-Bin Huang, Shenlong Wang, "ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis", ICCV 2023.

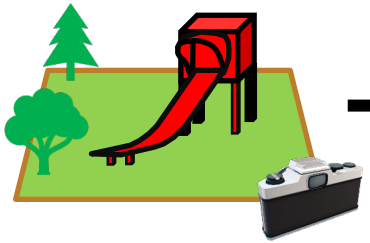
Extreme Weather Synthesis in NeRF

Snow

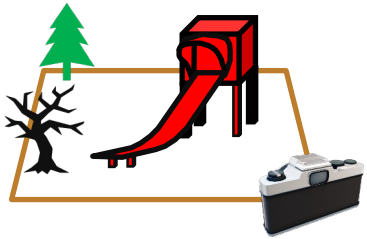
Climate Impact

Style image

Multi-view Input Images



3D Scene Modeling



Neural Scene Stylization

Yuan Li*, Zhi-Hao Lin*, David Forsyth, Jia-Bin Huang, Shenlong Wang, "ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis", ICCV 2023.

Extreme W

Snow

Climate Impact

Sty

Multi-view Input Imag

Yuan Li*, Zhi-Hao Lin*, David Forsyth, Jia-Bin Huang, Shenlong Wang, "ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis", ICCV 2023.

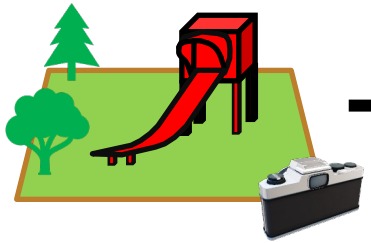
Extreme Weather Synthesis in NeRF

Snow

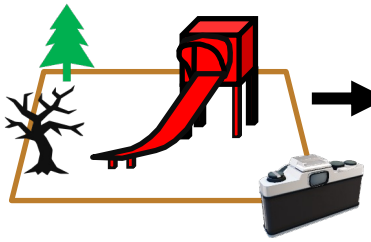
Climate Impact

Style image

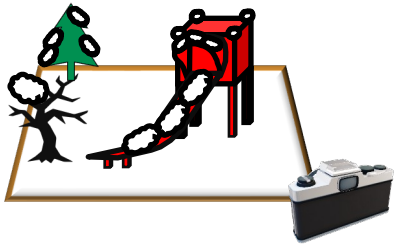
Multi-view Input Images



3D Scene Modeling



Neural Scene Stylization



Physical-Inspired Simulation

Yuan Li*, Zhi-Hao Lin*, David Forsyth, Jia-Bin Huang, Shenlong Wang, "ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis", ICCV 2023.

WESTERN
PACIFIC
713

Controllability

Possible to incorporate realistic weather projection

Can we simulate dynamic actors?

Input Video

Reconstructed Object

Simulated Results

Chen et al.. GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving, CVPR 2021 (best paper award candidate)

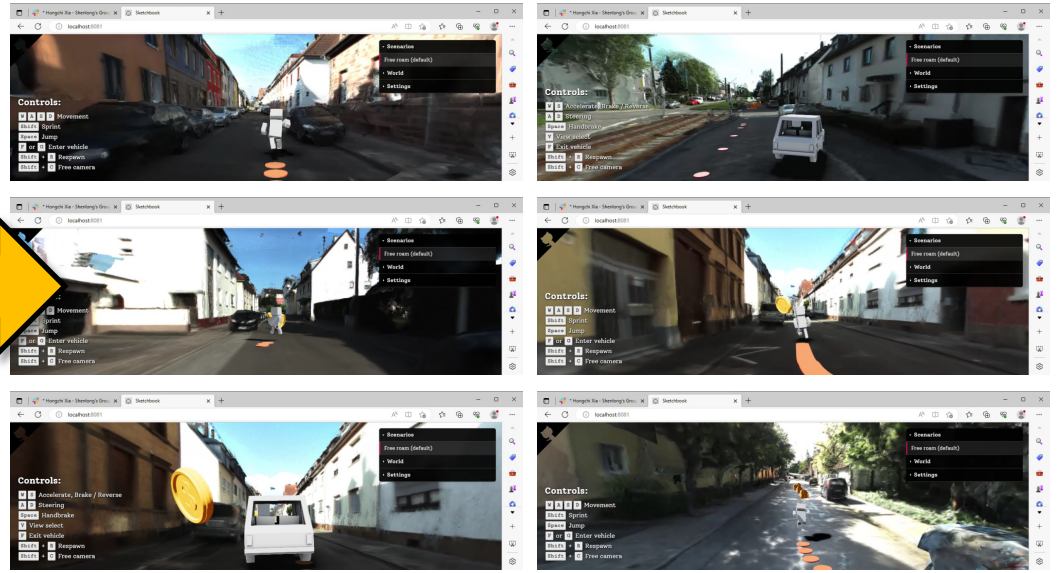
COOK
COMPANY
FedEx
0.5x

Eleventh St
0.5x

Can we turn a video to a game?

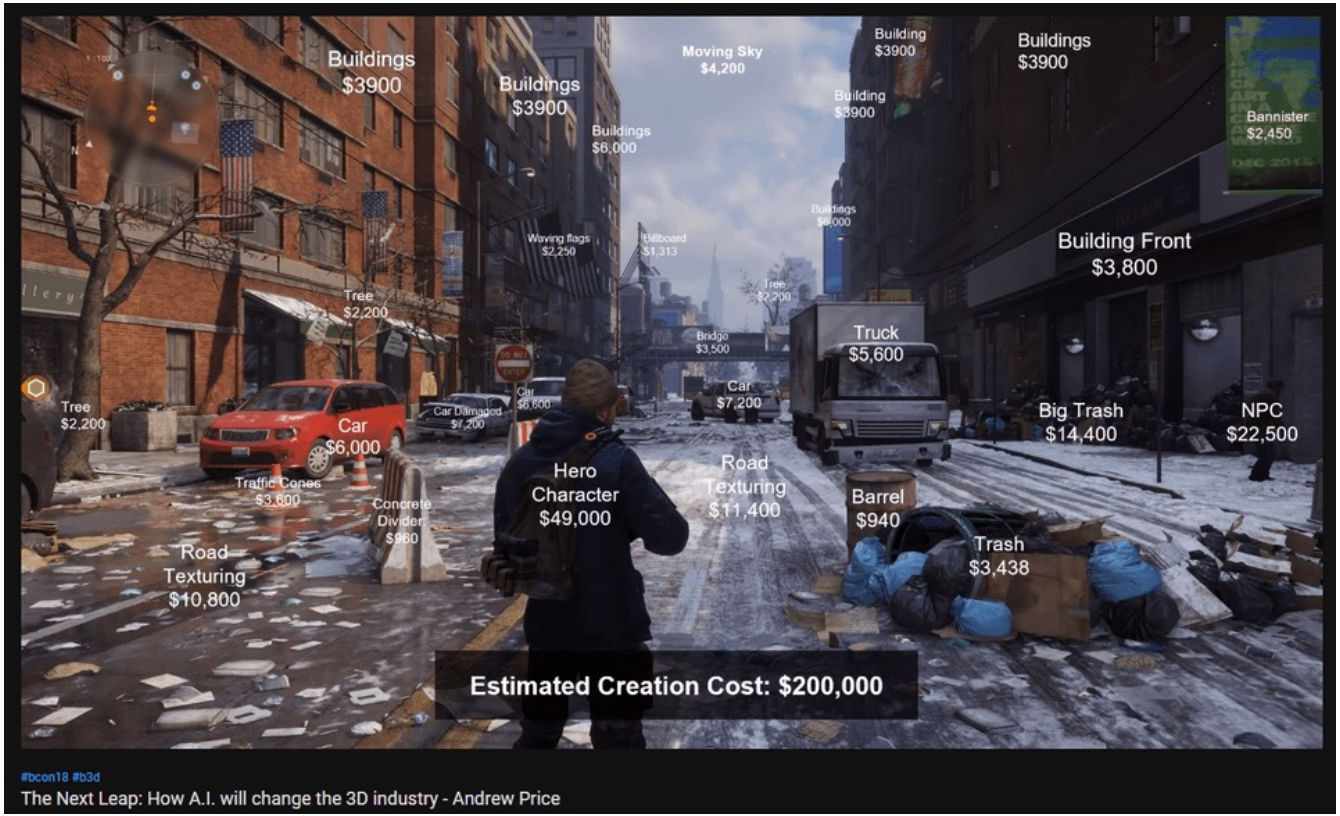
Input: single video

Output: real-time, realistic, interactive environment



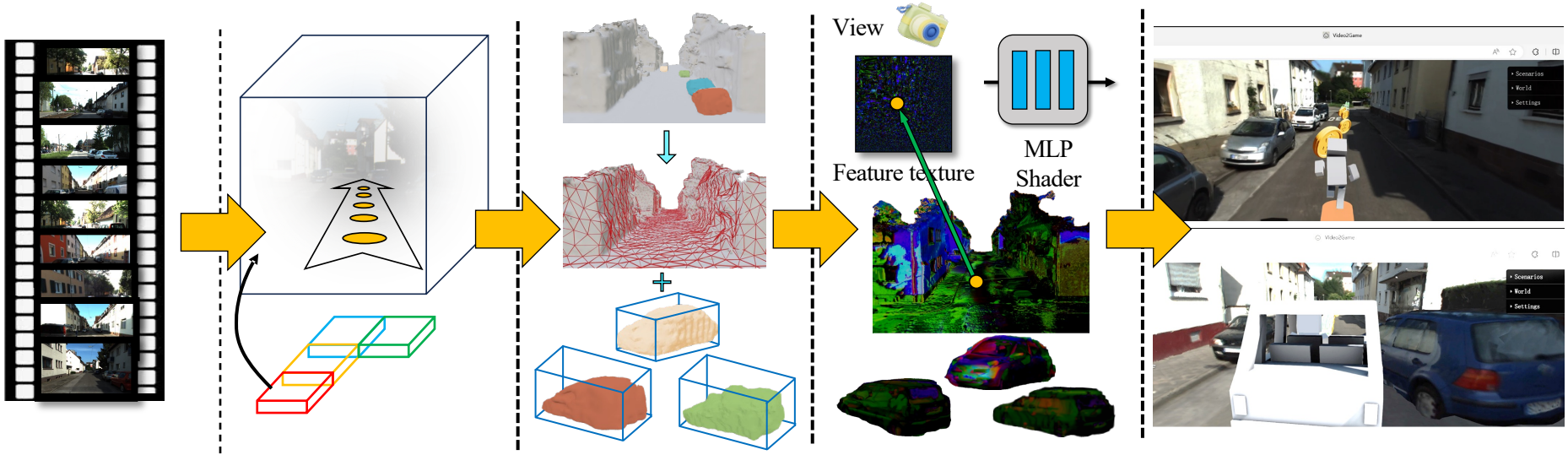
Hongchi Xia, Zhi-Hao Lin, Wei-Chiu Ma, Shenlong Wang, Video2Game: Real-time, Interactive, Realistic and Browser-Compatible Environment from a Single Video, CVPR 2024

Motivation: everything comes with a price tag



https://www.reddit.com/r/gamedev/comments/9zdn7a/putting_a_price_tag_on_game_assets_in_a_screenshot/

Overall framework

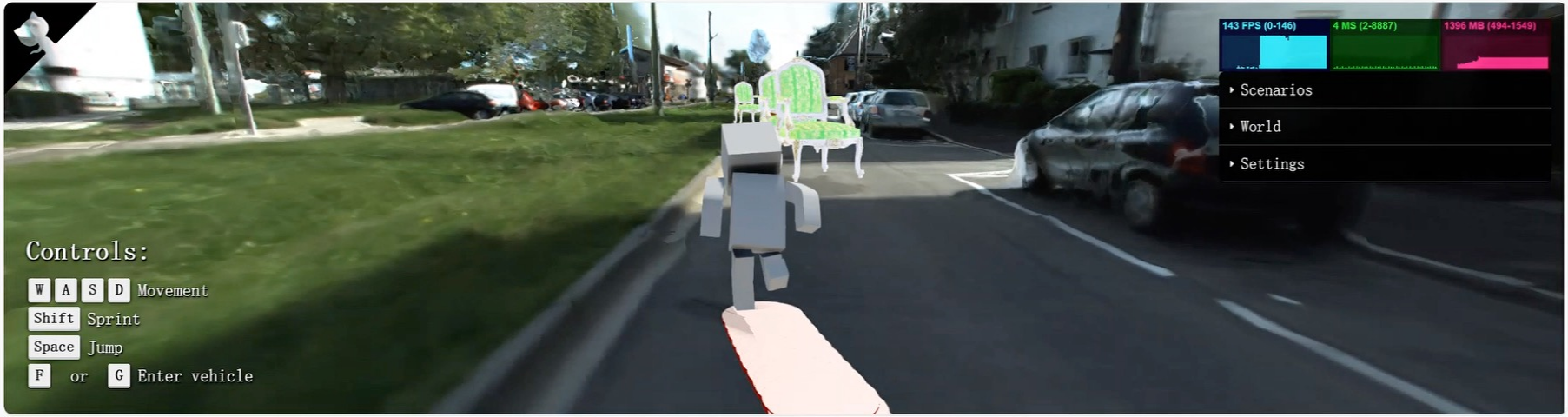


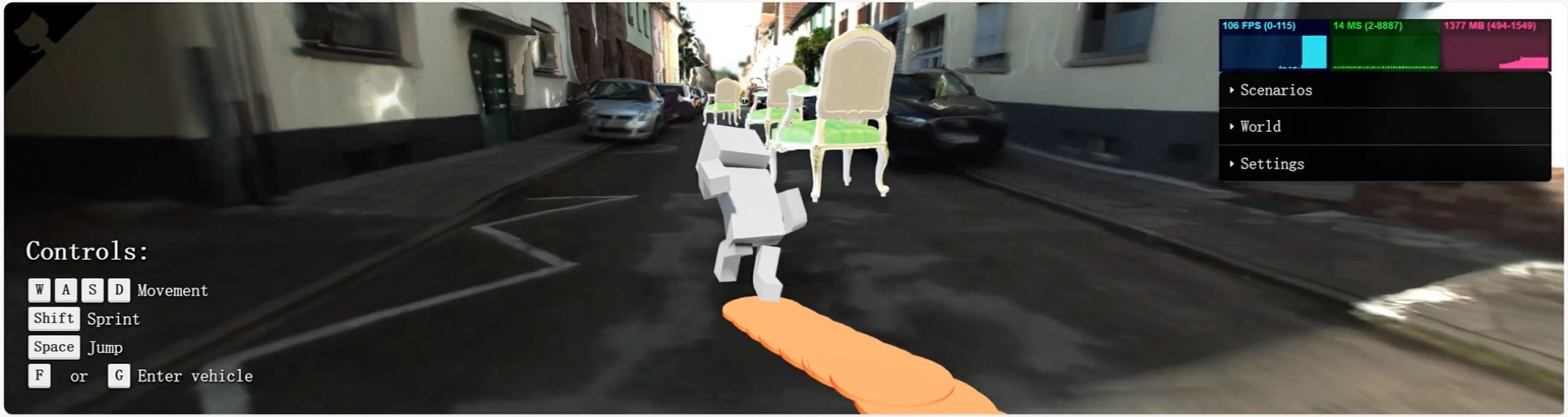
Input video | Base NeRF | Rigid-body dynamics | Neural-textured mesh | Game interactive Environment

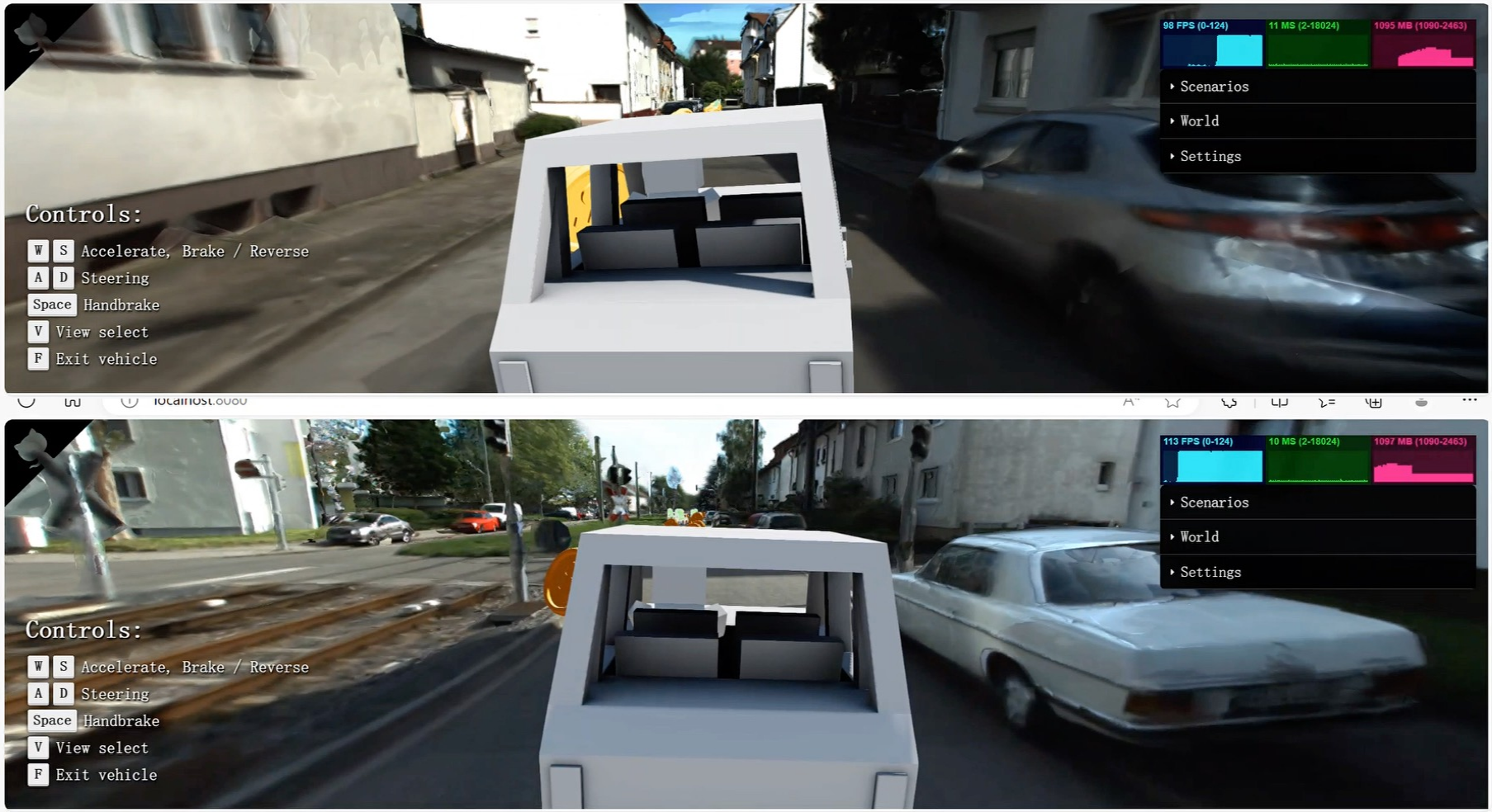
Hongchi Xia, Zhi-Hao Lin, Wei-Chiu Ma, Shenlong Wang, Video2Game: Real-time, Interactive, Realistic and Browser-Compatible Environment from a Single Video, CVPR 2024

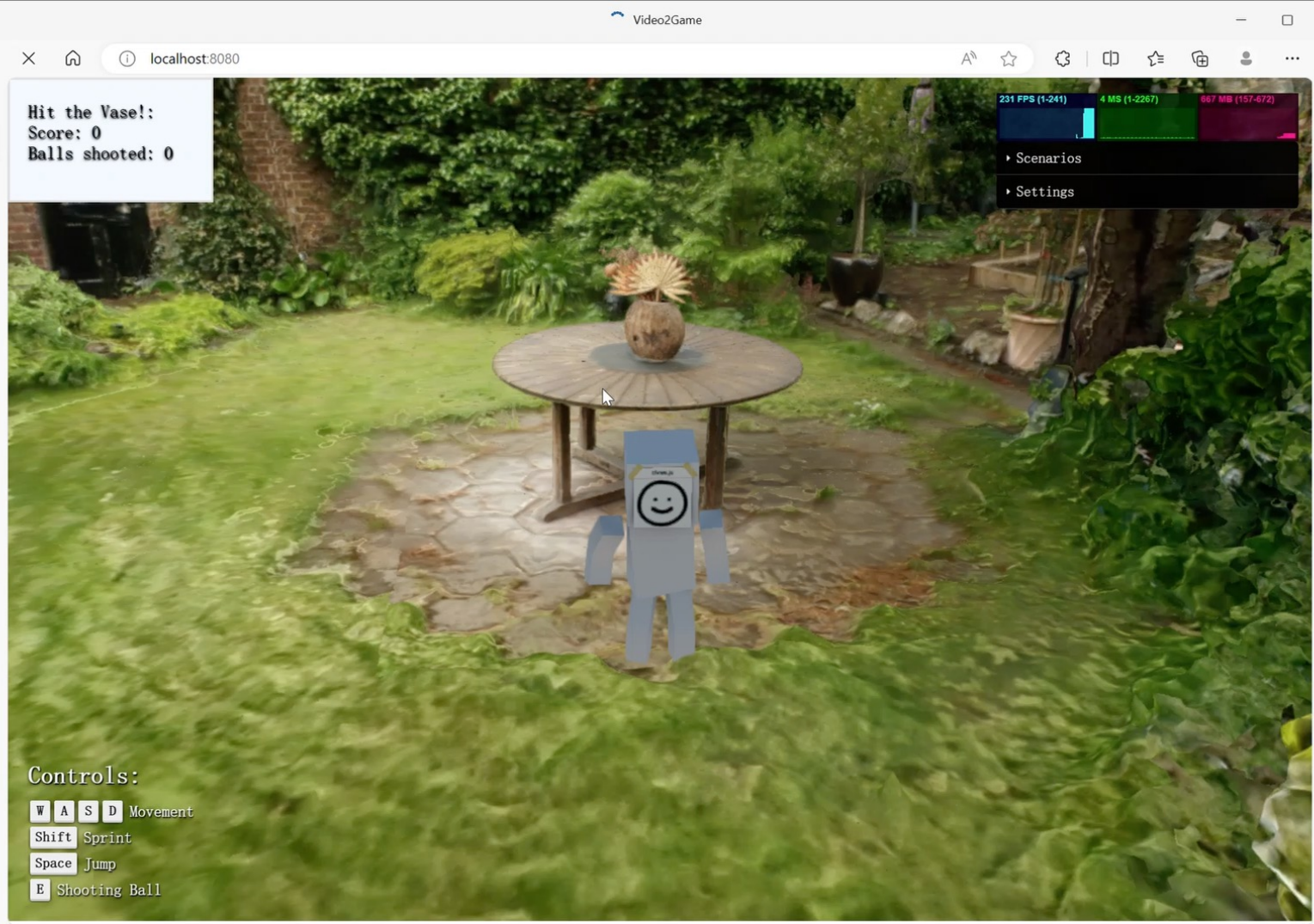
Video2Game

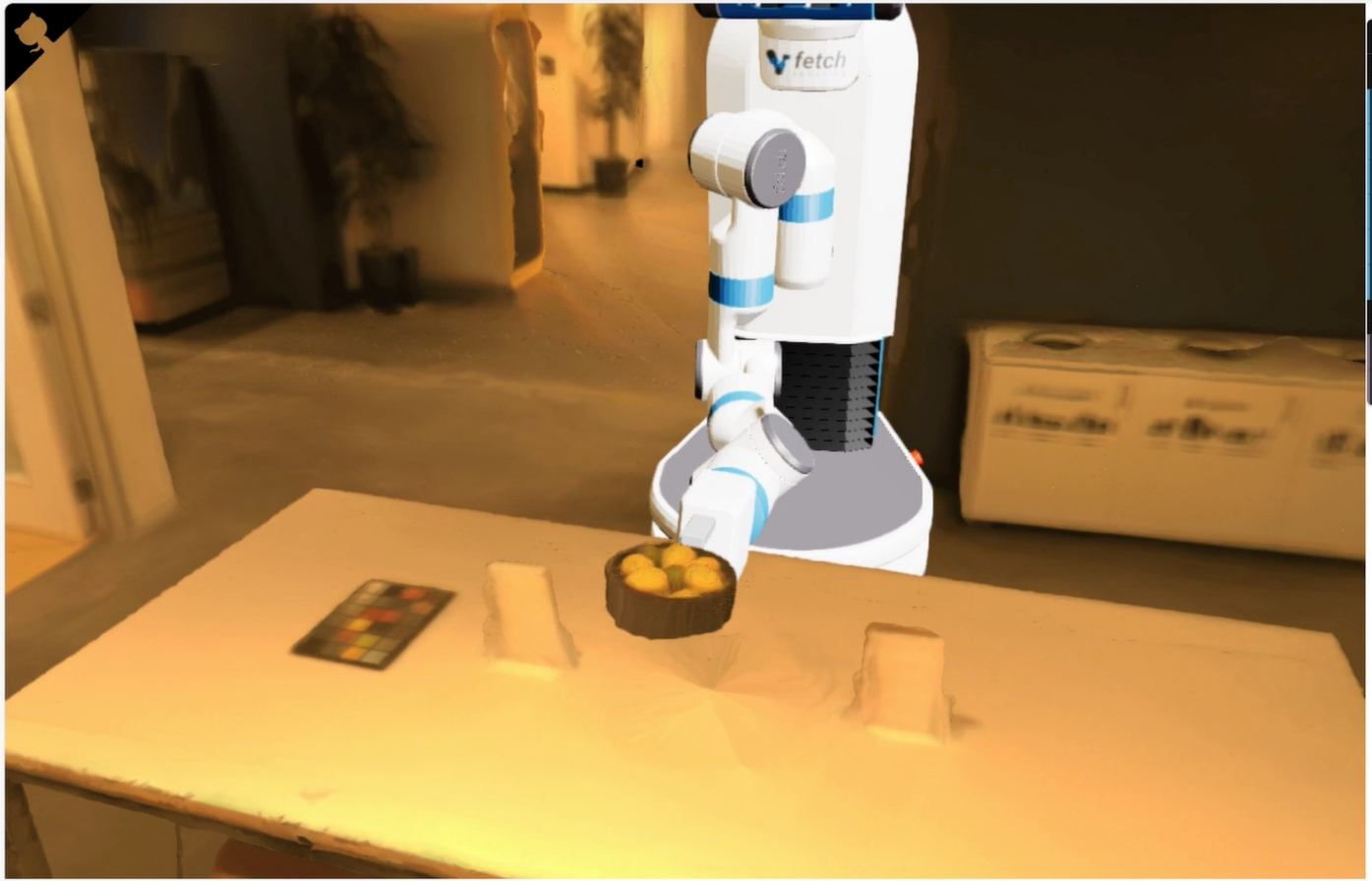
localhost:8080






Video2Game
localhost:8080
98 FPS (0-124)
11 MS (2-18024)
1095 MB (1090-2463)
Scenarios
World
Settings
Controls:
W S Accelerate, Brake / Reverse
A D Steering
Space Handbrake
V View select
F Exit vehicle
113 FPS (0-124)
10 MS (2-18024)
1097 MB (1090-2463)
Scenarios
World
Settings
Controls:
W S Accelerate, Brake / Reverse
A D Steering
Space Handbrake
V View select
F Exit vehicle


Video2Game
localhost:8080
Hit the Vase!:
Score: 0
Balls shooted: 0
231 FPS (1-241)
4 MS (1-2267)
667 MB (157-672)
Scenarios
Settings
Controls:
W A S D Movement
Shift Sprint
Space Jump
E Shooting Ball



Scenarios

World

ascend	0
radius	0.71
fix_camera_y	14
fetch_frame	0
camera_lowy	0.2
stretch	0

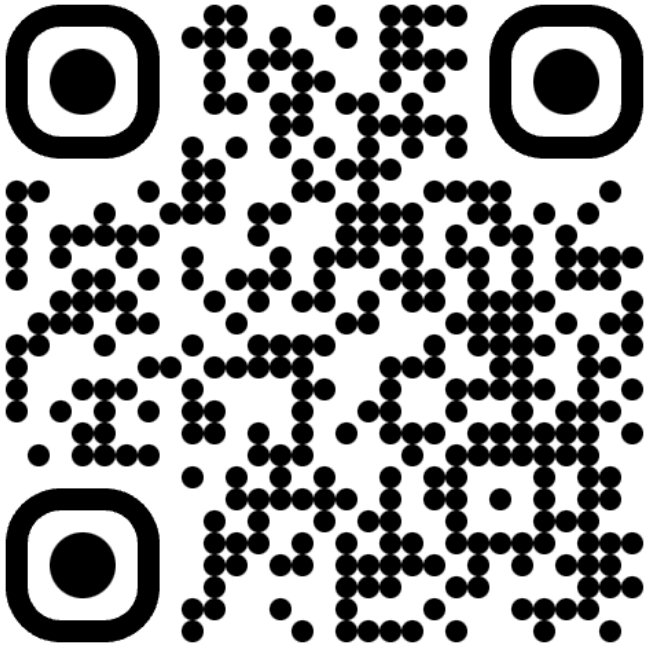
Settings

Debug_Physics	☐
Debug_FPS	☐

Controls:

W A S D Movement

Demo



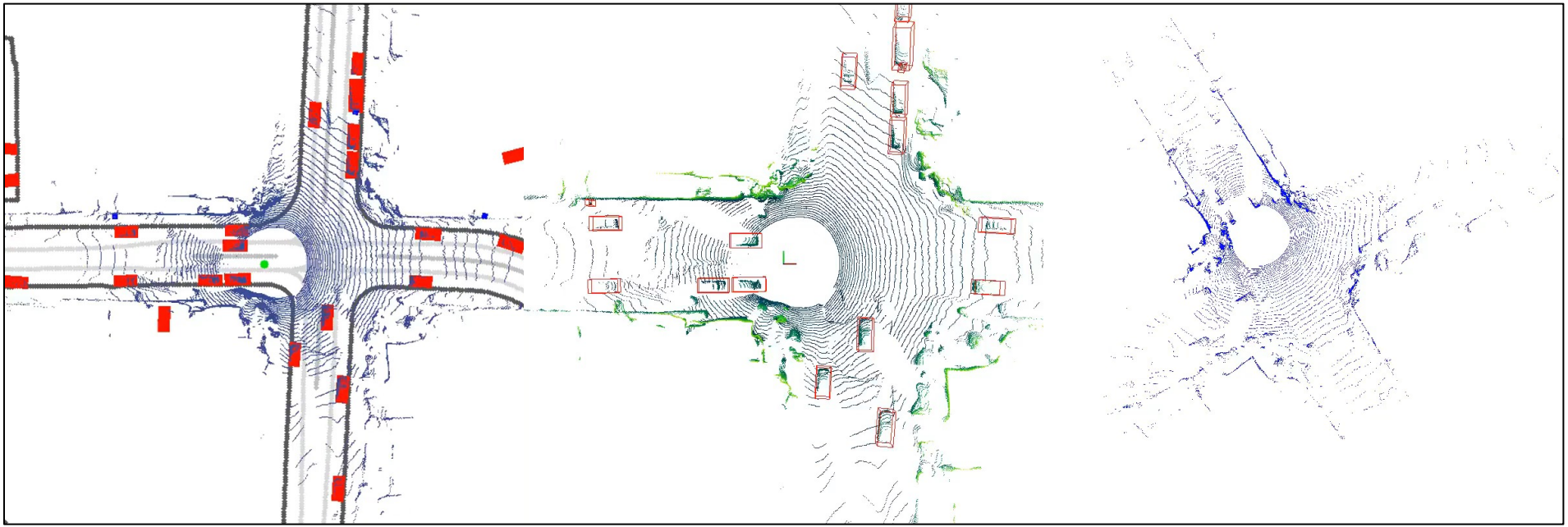
- https://video2game.github.io/

Today's talk

Modeling: *Data-driven perception + inverse physics*

Generation: *Physics-based simulation + generative prior*

Can we ground physics in generative AI?



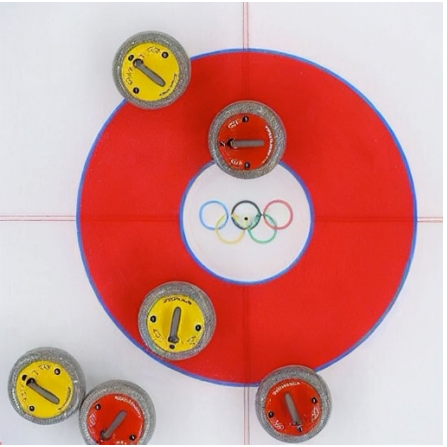
Layout-aware

Realistic

Coherent in space & time

LidarDM, arXiv soon

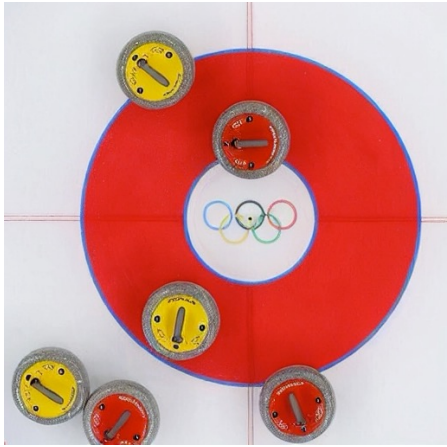
Can we ground physics in generative AI?



SEINE

I2VGEN-XL

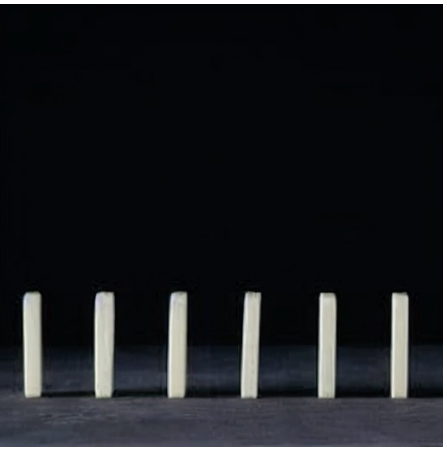
DynamiCrafter



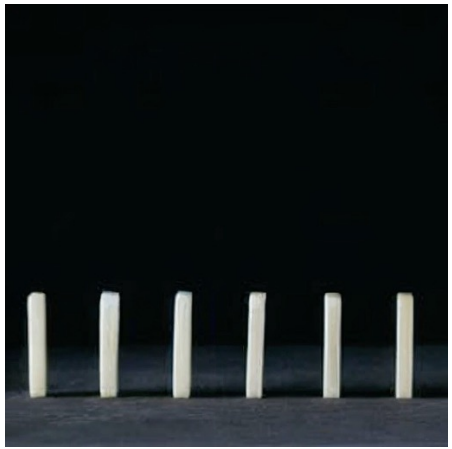
Ours

PhysGen, arXiv soon

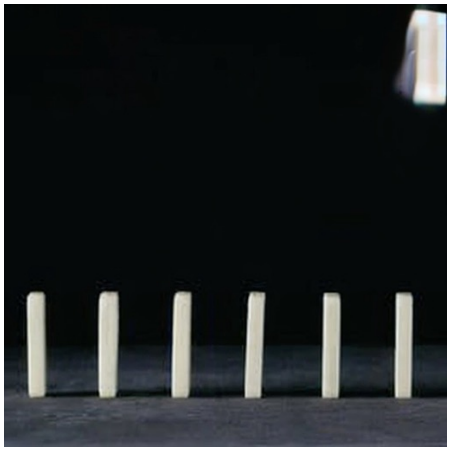
Can we ground physics in generative AI?



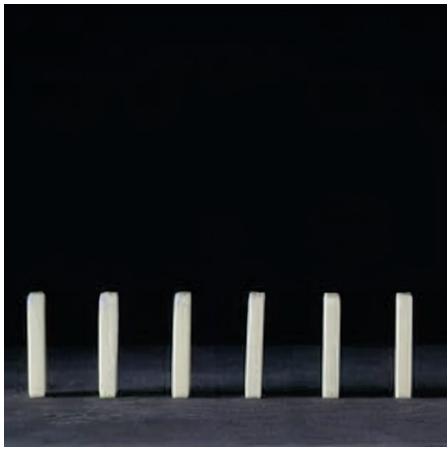
SEINE



I2VGEN-XL



DynamiCrafter

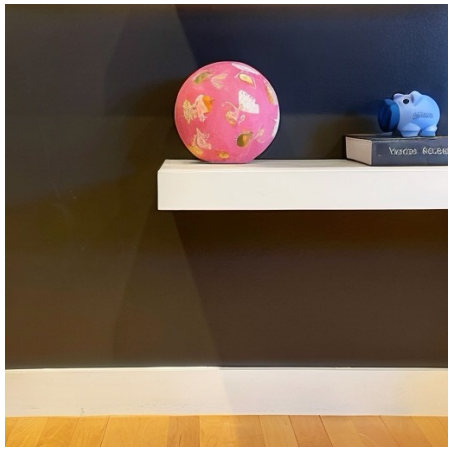


Ours

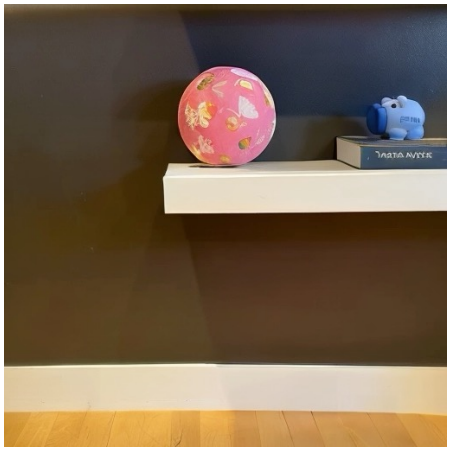
PhysGen, arXiv soon

Can we ground physics in generative AI?

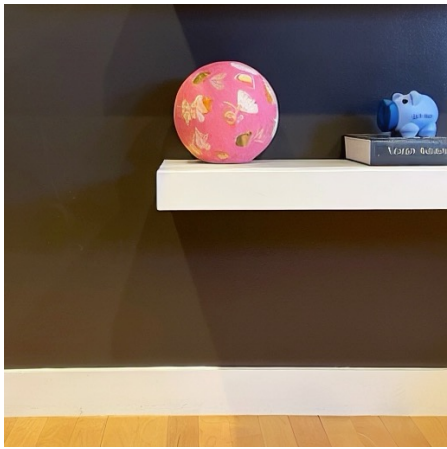
SEINE



I2VGEN-XL



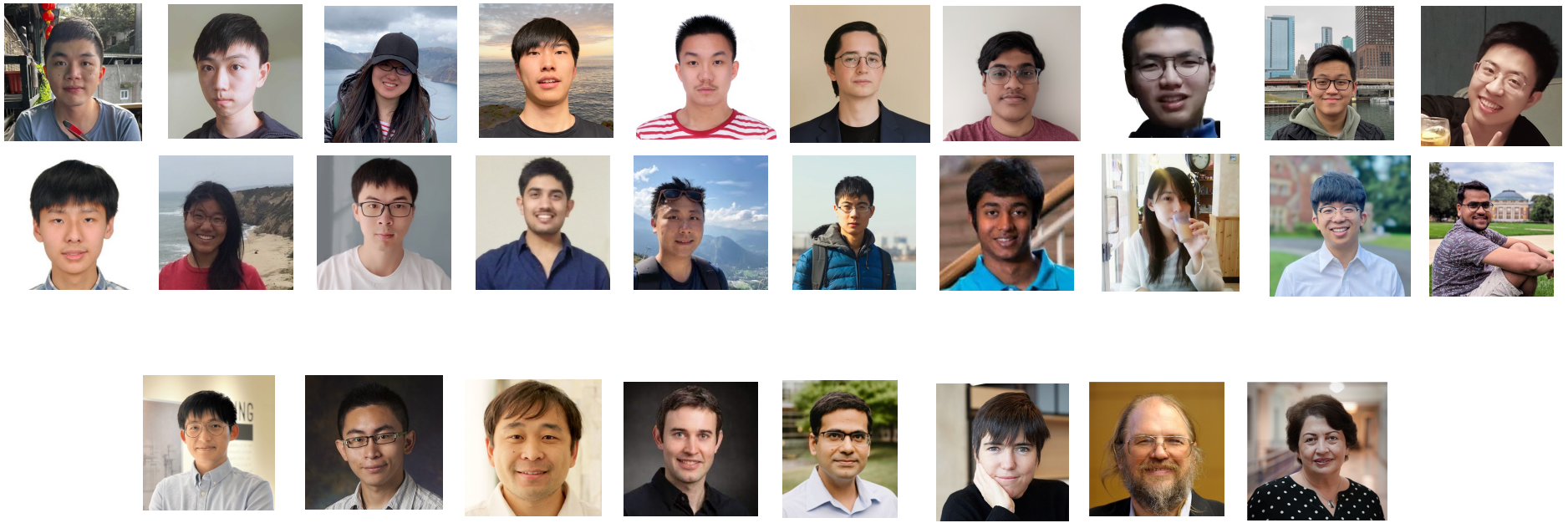
DynamiCrafter



Ours

PhysGen, arXiv soon

Acknowledgement



Acknowledgement

Today's talk

Modeling: *Data-driven perception + inverse physics*

Generation: *Physics-based simulation + generative prior*