CS598 Fall 2024: 3D Vision

Shenlong Wang

Aug 27, 2024

Welcome!

About me

- **Shenlong, Email:** `shenlong@Illinois.edu`
- **Assistant Professor @ CS, Siebel 3336**

https://virtual-correspondence.github.io/

About Zhi-Hao

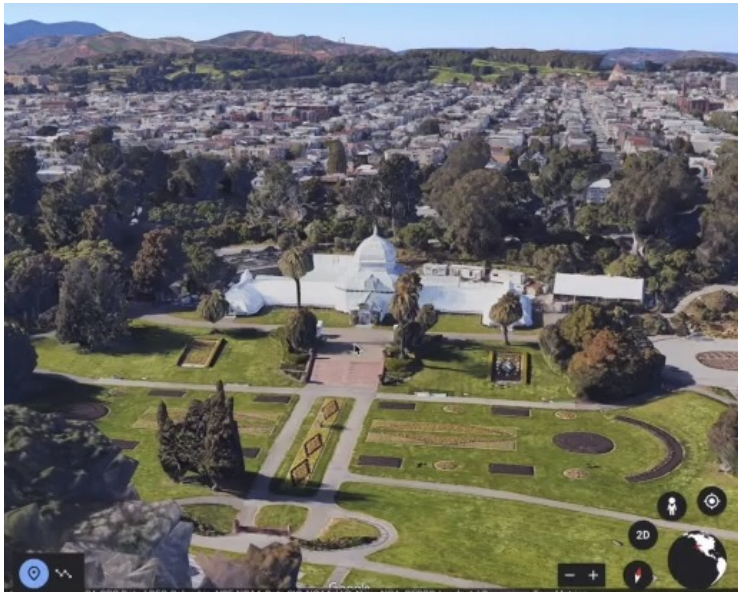
- Zhi-Hao Lin, 3rd PhD @ CS
- BS, MS @ National Taiwan University
- Internship @ Meta, Nvidia

- Why do I like 3D vision:

  it's the starting point for creating worlds in AR/VR—who wouldn't want to design their own reality?

2014: Beginning of my PhD at Toronto

.. Curious how Google Earth 3D works..



Google earth

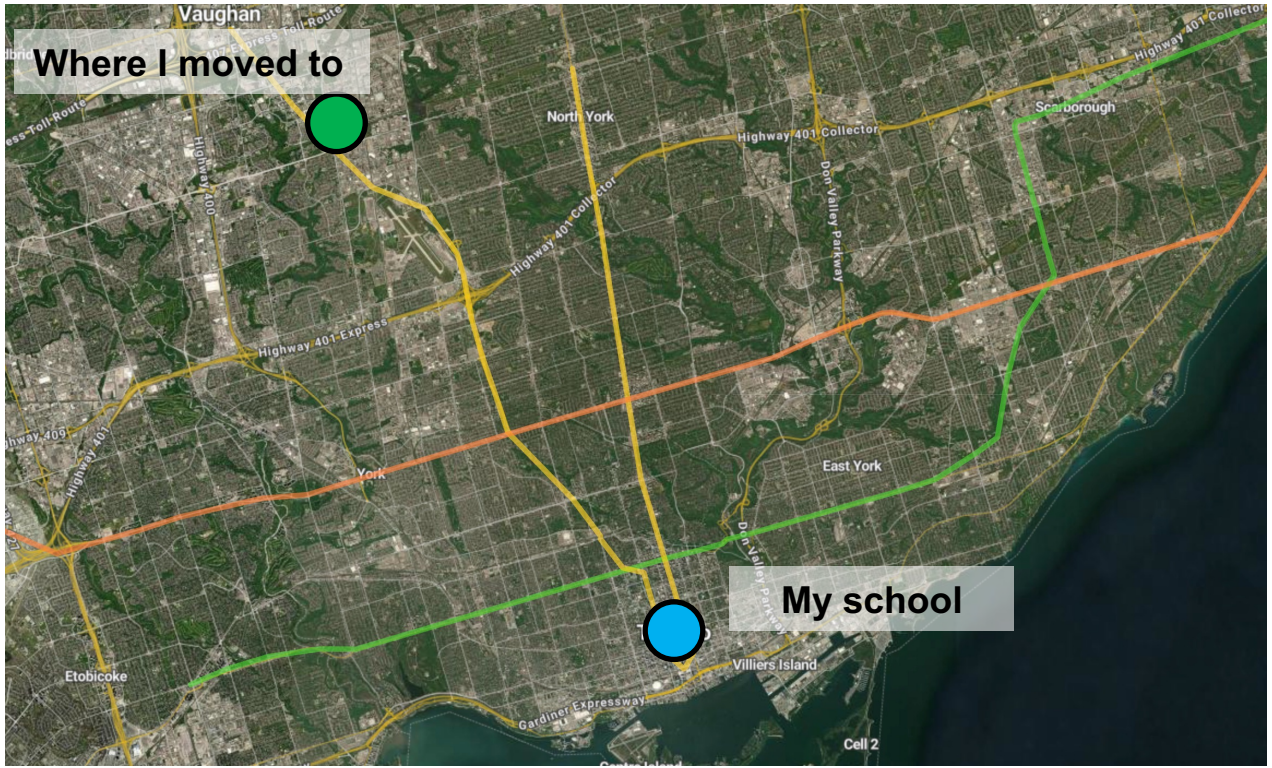
Found that it requires thousands of images!

Noah Snavely

Let’s do 3D reasoning from a single image!

Holistic 3D Scene Understanding from a Single Geo-tagged Image, CVPR 2015

2014: Toronto began to be expensive..

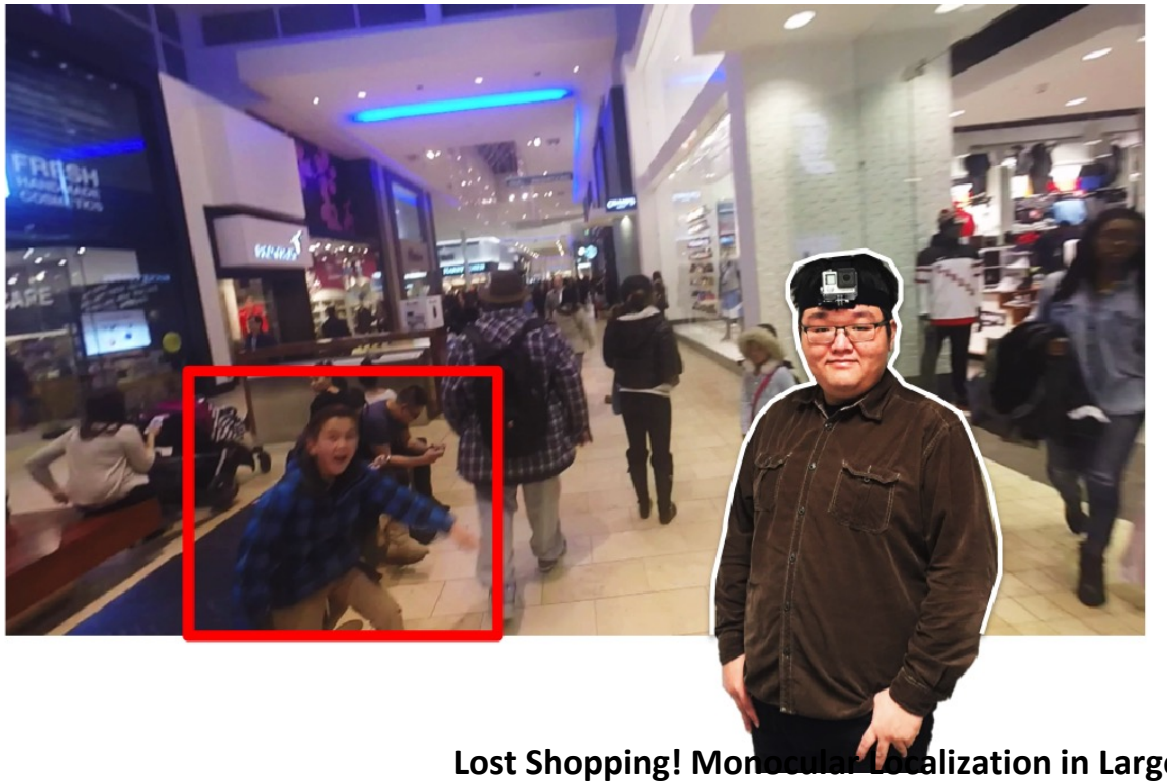


2014: Toronto began to be expensive..

Getting lost in malls..

Lost Shopping! Monocular Localization in Large Indoor Spaces, ICCV 2015

Collect data to study the 3D indoor localization..

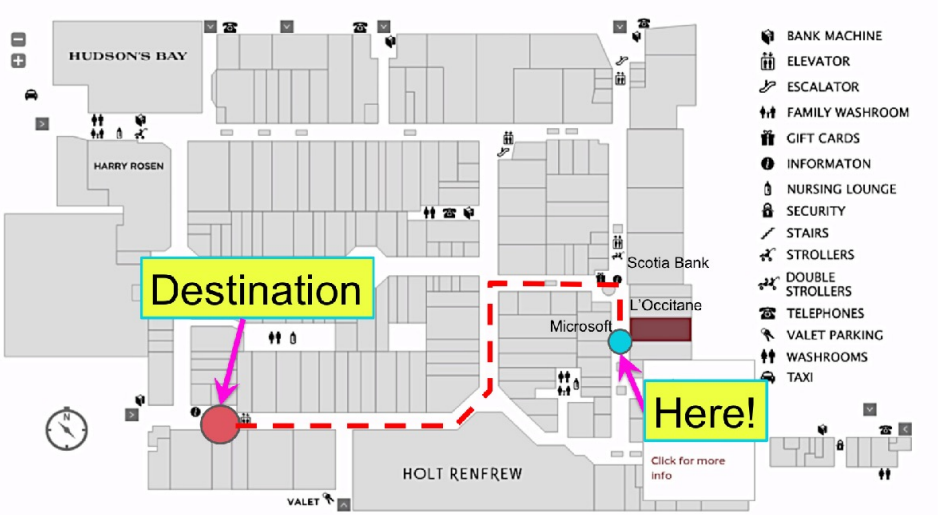


Lost Shopping! Monocular Localization in Large Indoor Spaces, ICCV 2015

Get an algorithm for it..

Step 1: Look Around & Reason 3D

Step 2: Check the Floorplan!



Lost Shopping! Monocular Localization in Large Indoor Spaces, ICCV 2015

2015: Attracted by Microsoft HoloLens..

Hololens, Microsoft

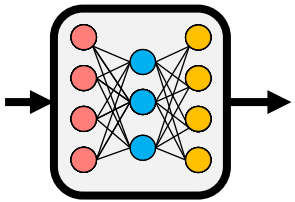
2016: Did my first internship there on teleportation

Holoportation: Virtual 3D Teleportation in Real-time, UIST 2016

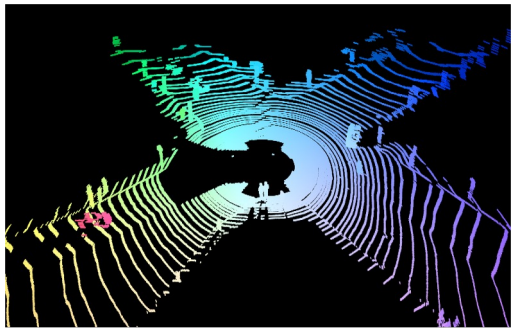
2017: Started building self-driving cars…

2018: Developed Convolution on 3D Point Cloud

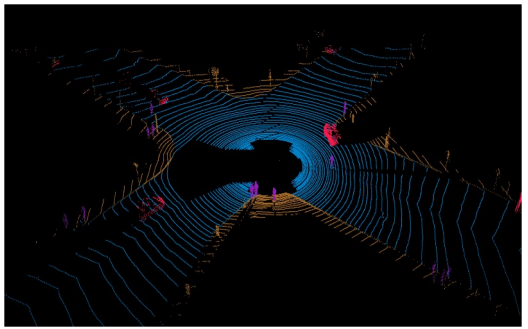
3D Point Cloud



Motion Estimation

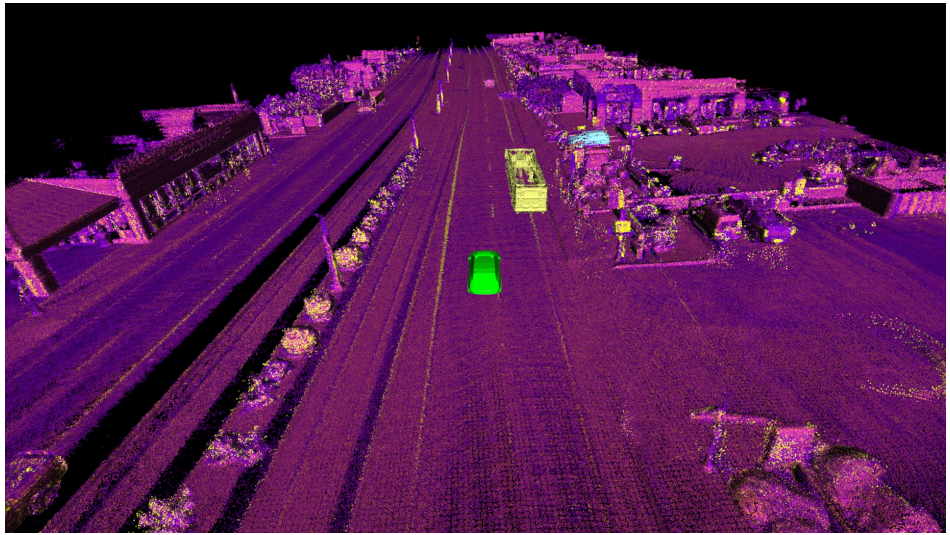


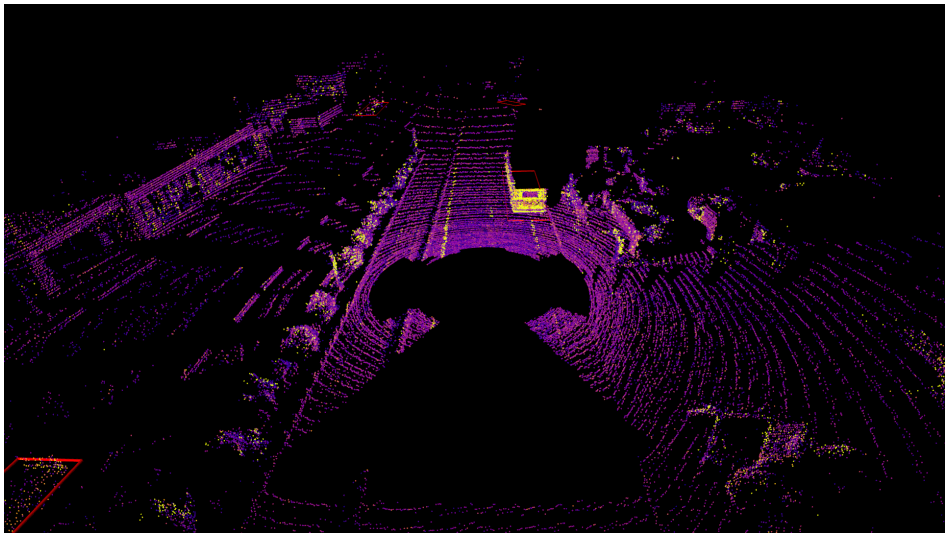
Semantic Segmentation



Deep Parameteric Continuous Convolutional Neural Networks, CVPR 2018

2020: Autonomy Need Simulation!





LidarSIM: Realistic LiDAR Simulation by Leveraging the Real World, CVPR 2020

2020: Autonomy Need Simulation!

Input Video

Reconstructed Object

Simulated Results

GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving, CVPR 2021

2022: Came to Illinois and started working with an amazing group of talents!

Make 3D world editable

GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving, CVPR 2021

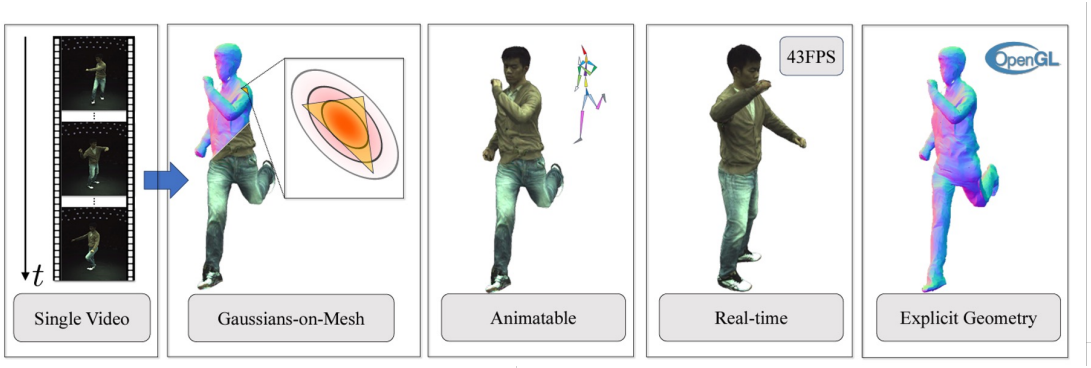
Make 3D world playable

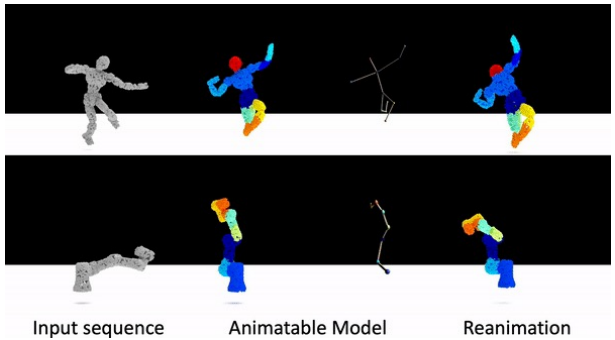
GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving, CVPR 2021

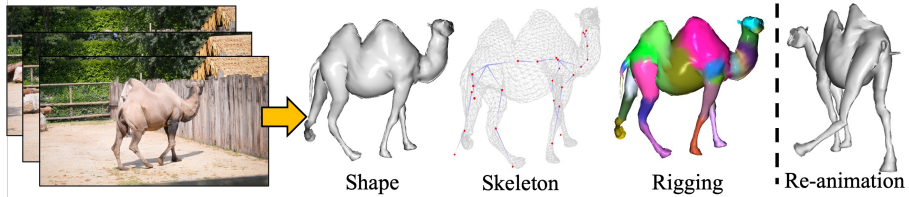
Make 3D world Actionable

GeoSim: Realistic Video Simulation via Geometry-Aware Composition for Self-Driving, CVPR 2021

Reason Dynamic Objects in 3D World





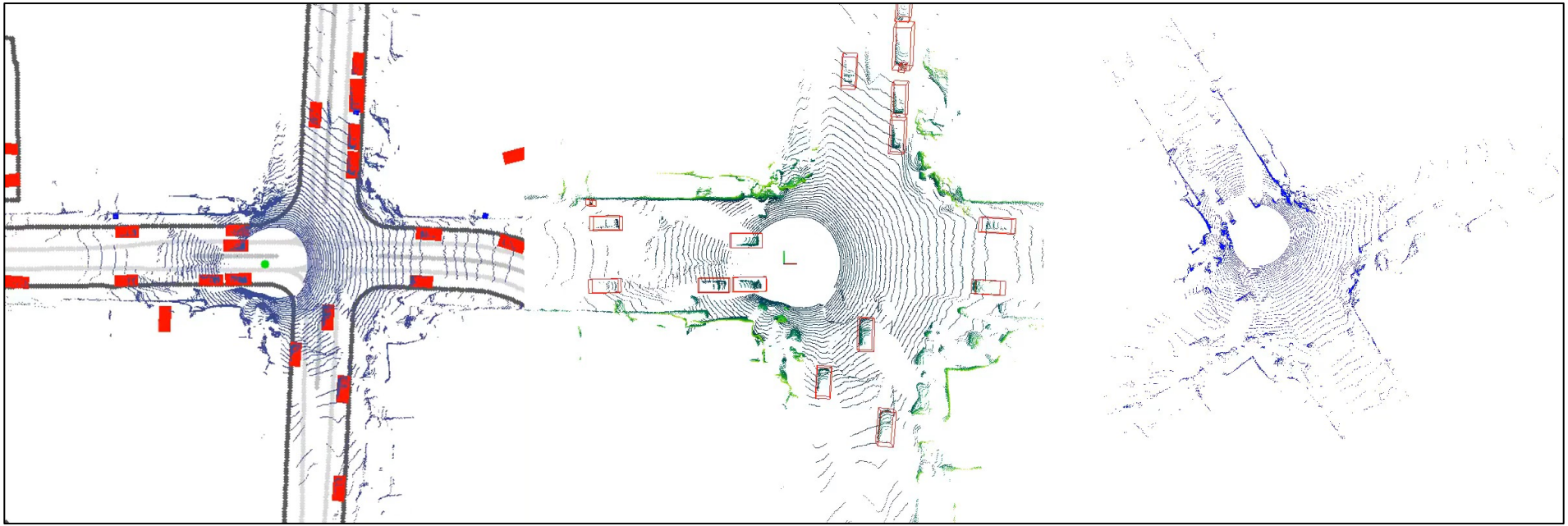


GoMAvatar: Efficient Animatable Human Modeling from Monocular Video Using Gaussians-on-Mesh, CVPR 2024

CASA: Category-agnostic Skeletal Animal Reconstruction, NeurIPS 2023

Building Rearticulable Models for Arbitrary 3D Objects from 4D Point Clouds, CVPR 2023

Generating virtual world for autonomy



LidarDM: Generative LiDAR Simulation in a Generated World, arXiv 2024

Make 3D world filled with imagination

AutoVFX: Physically Realistic Video Editing from Natural Language Instructions

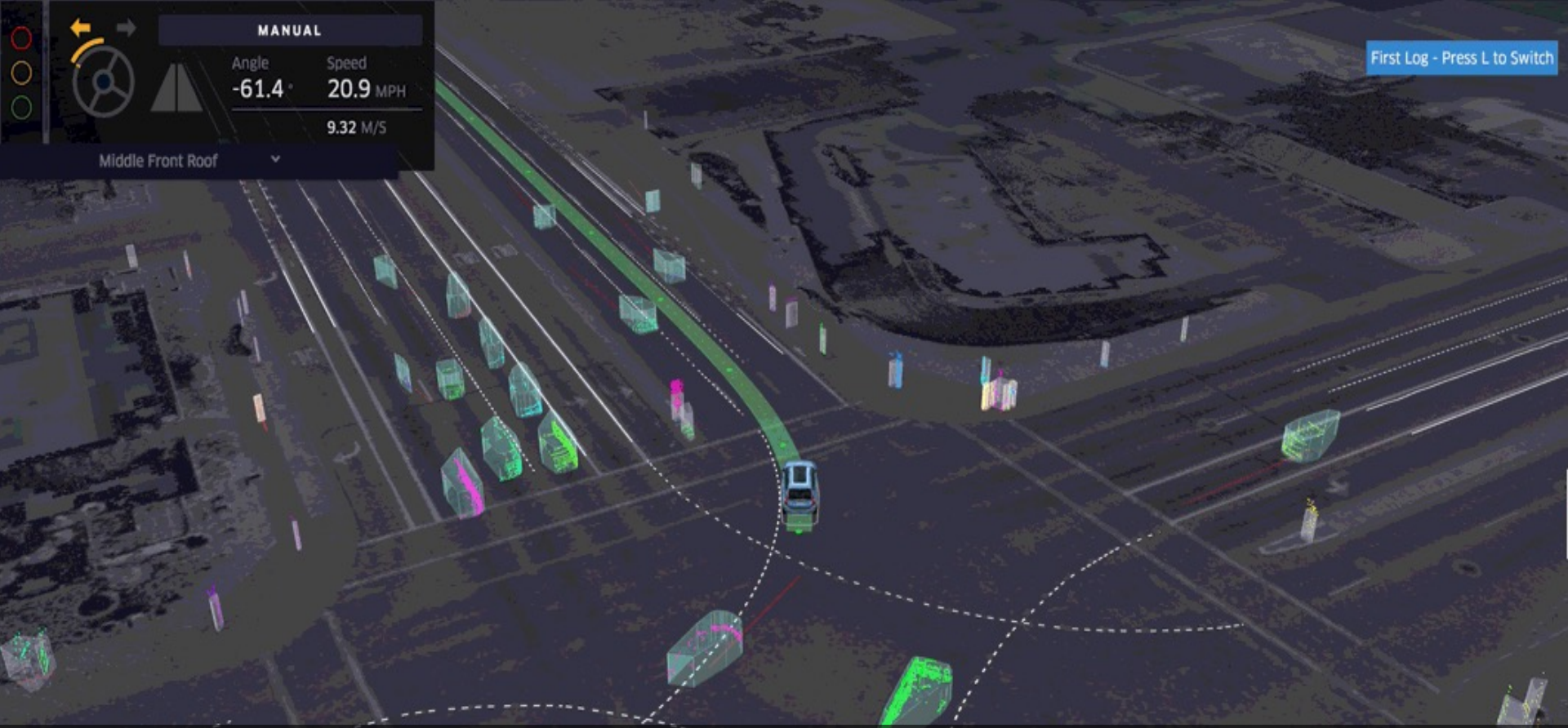
Supplementary Video

Today’s Agenda

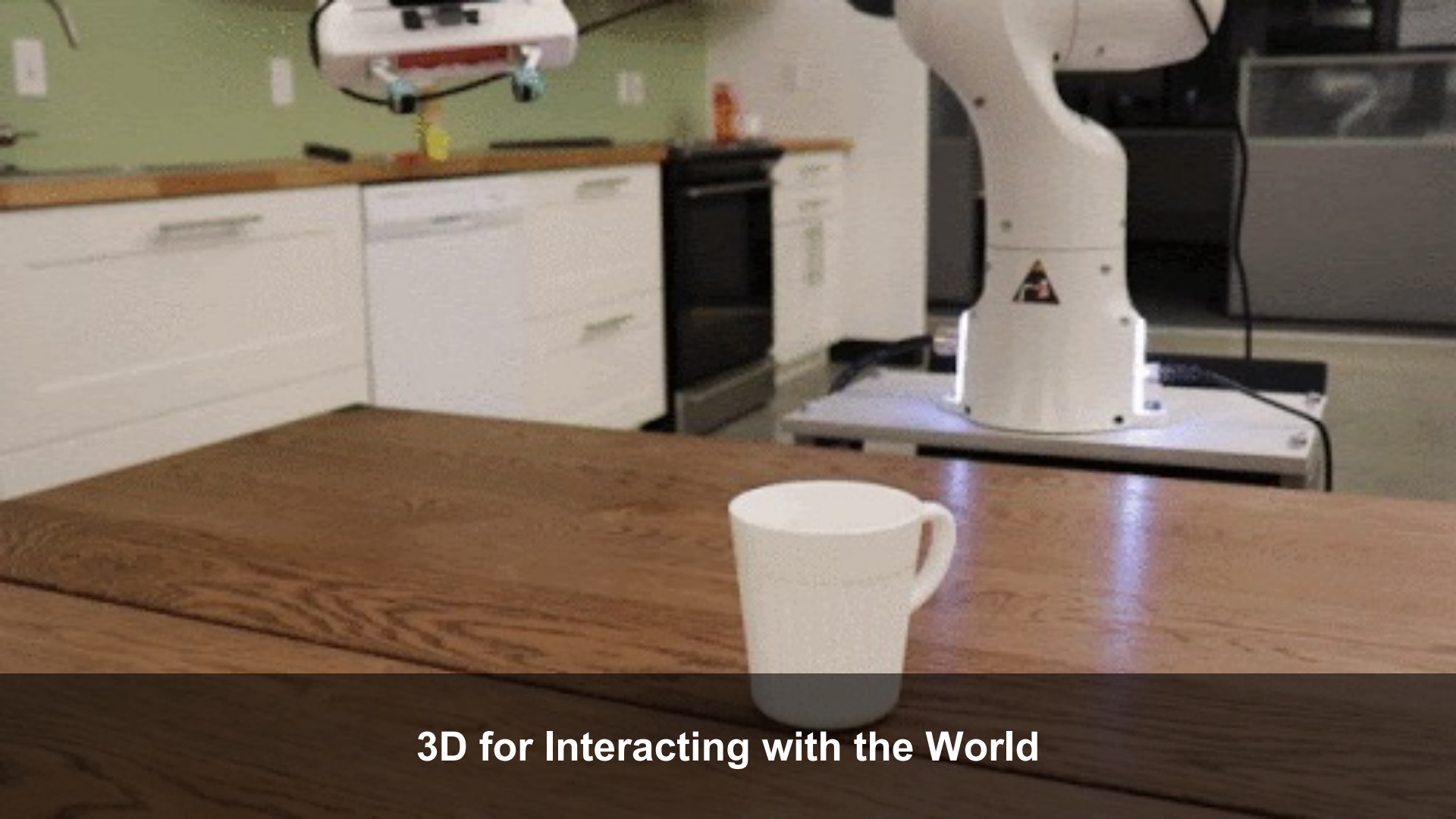
- Why 3D Vision?
- What will be covered?
- Logistics (Role-playing!)

Why 3D vision?

We live in a 3D World



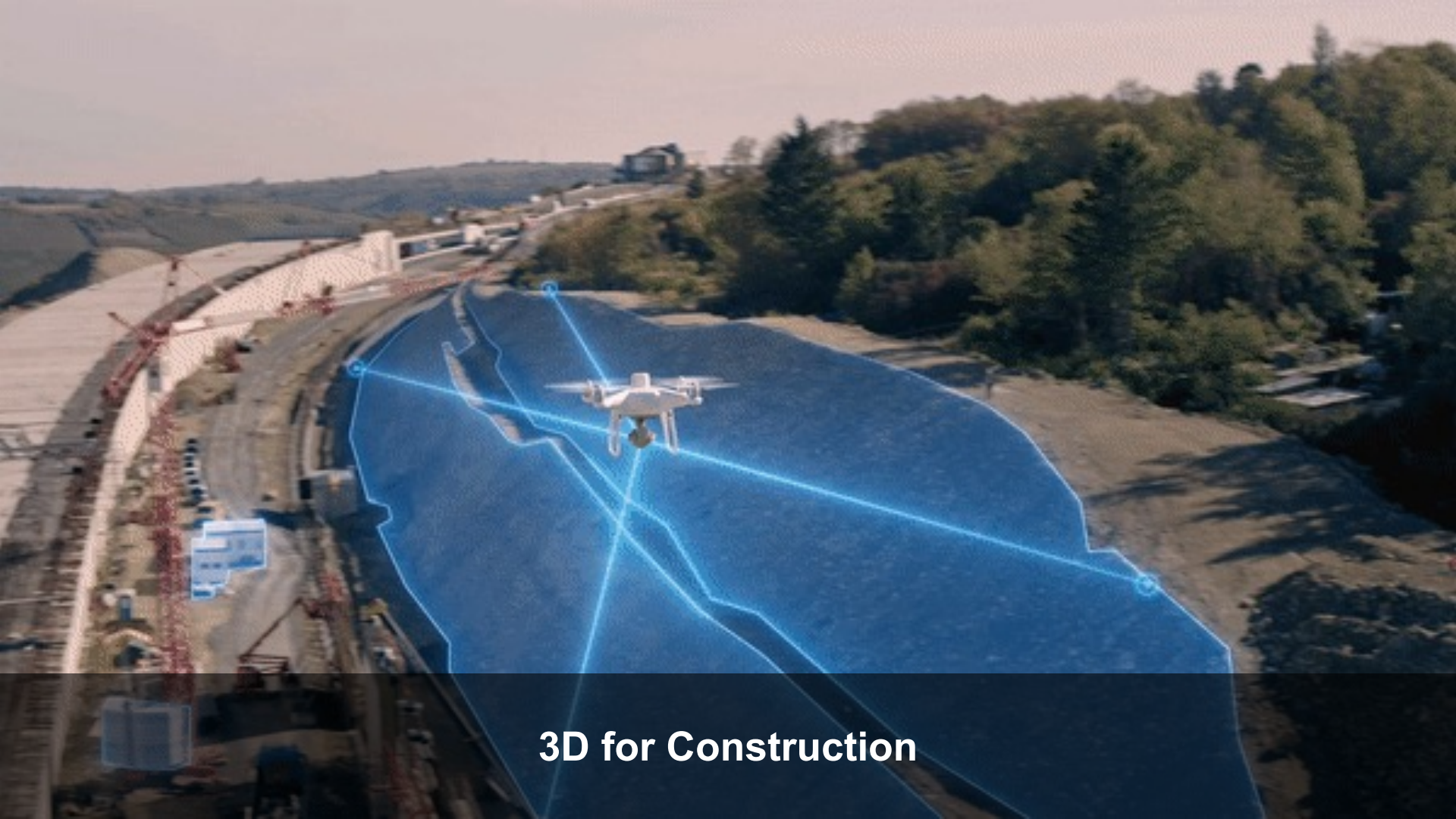
3D for Autonomous Driving


3D for Interacting with the World


3D for Scientific Discovery

3D for AR / VR

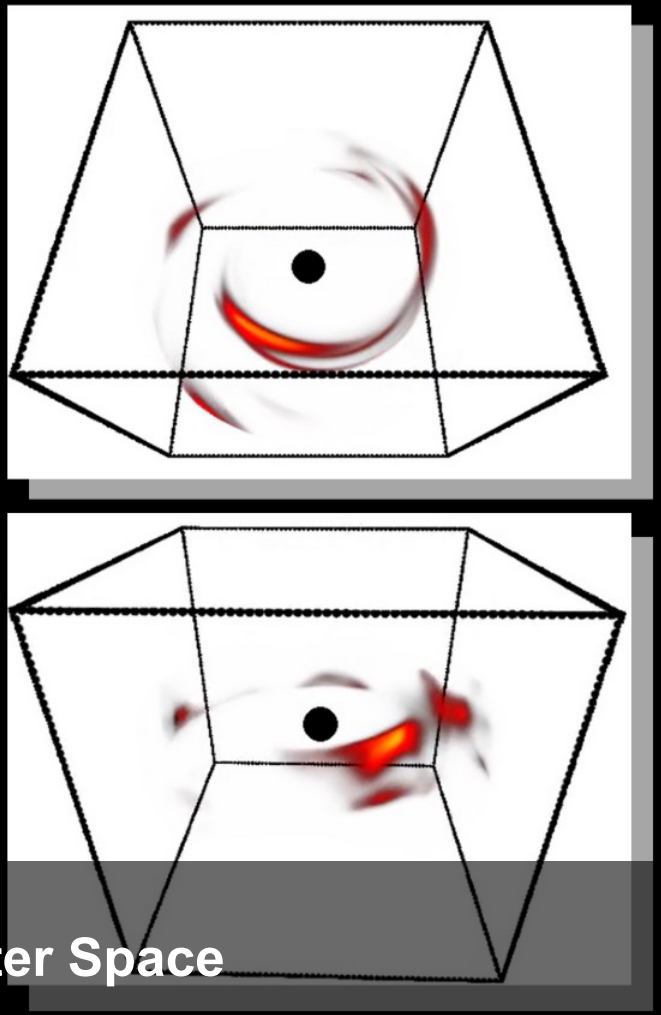
3D for Gaming and Arts


3D for Construction

3D for Agriculture


3D for Social Goods

Scientists use AI to reconstruct energetic flare blasted from Milky Way's supermassive black hole, Space.com



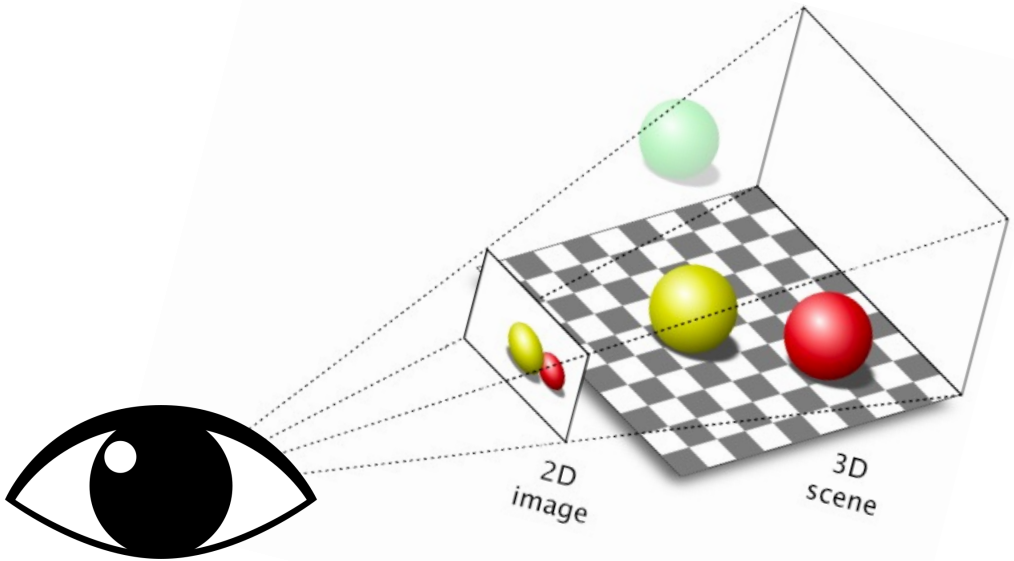
3D for Understanding the Outer Space


3D for Capturing Memories

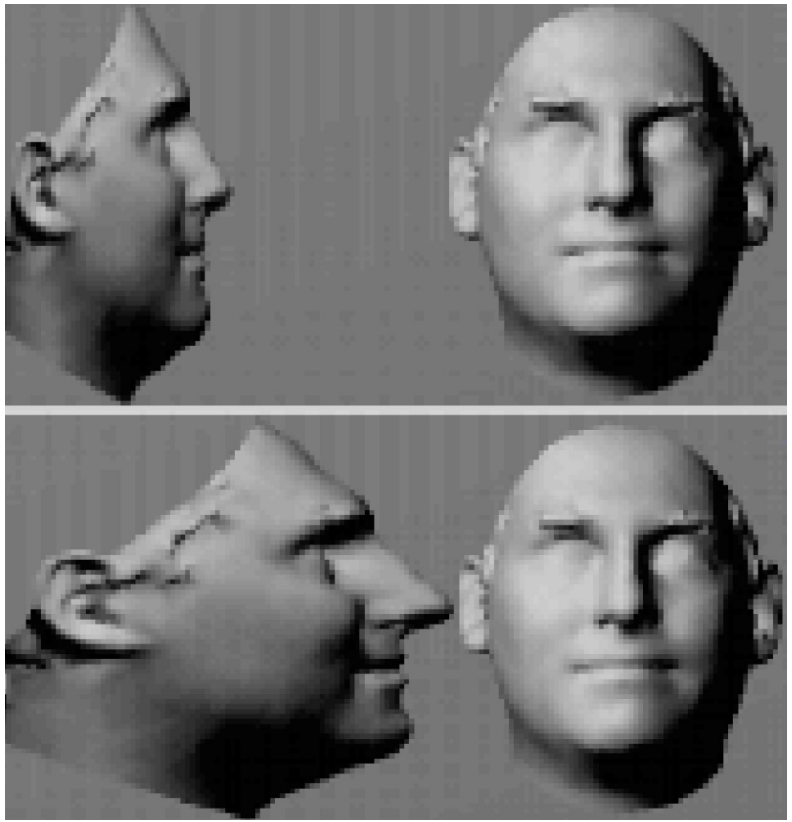

3D for Better Connecting People

Why 3D vision is challenging?

We (and our cameras) see the 3D world from 2D

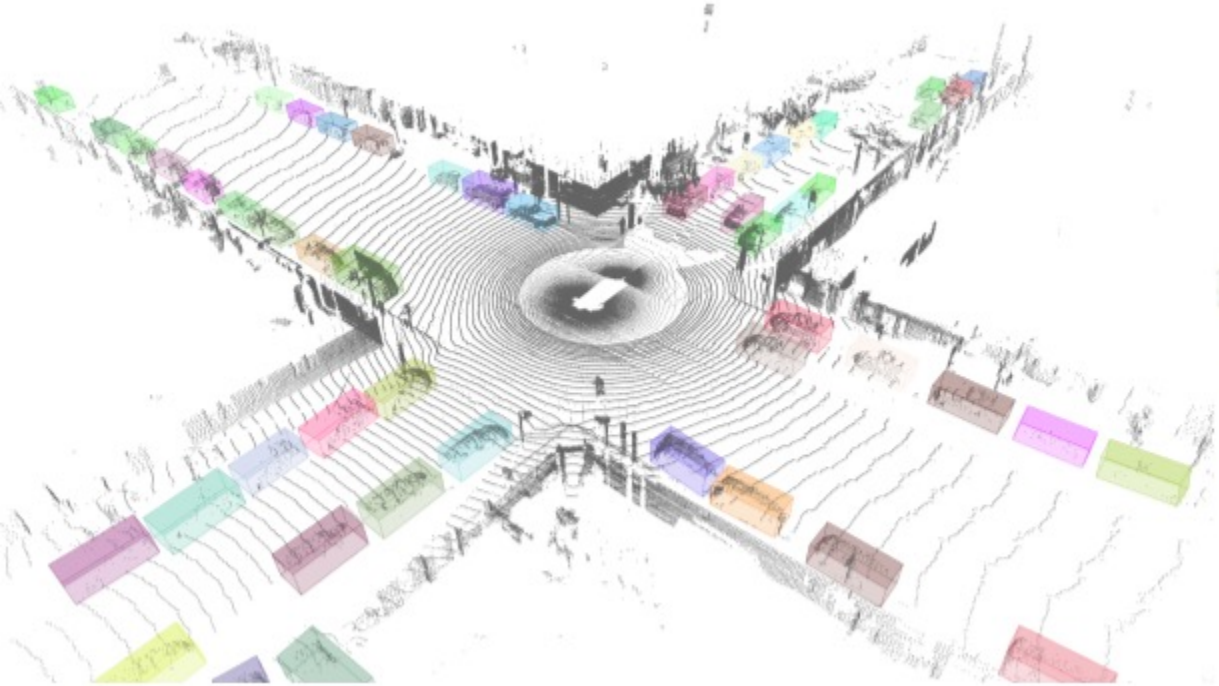


But that becomes ill-posed..

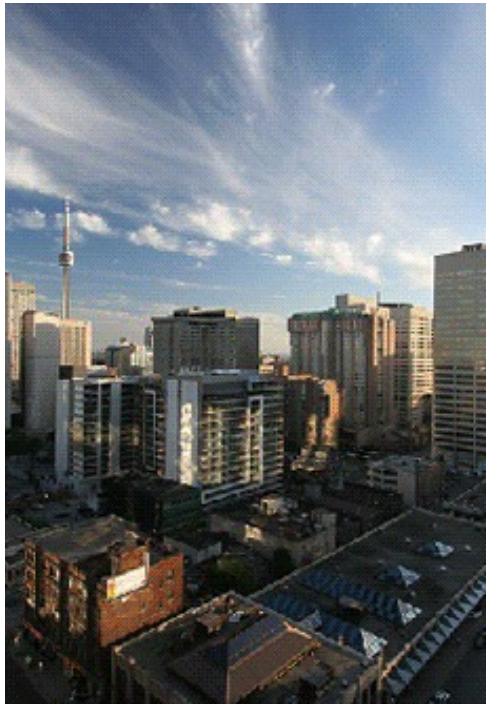


And brings illustrations..

Modeling 3D world needs understanding semantics



And handle many other factors.



Illumination

Materials

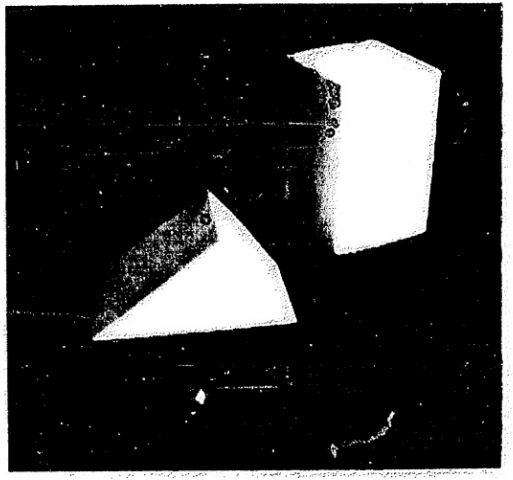
Dynamics and Motion

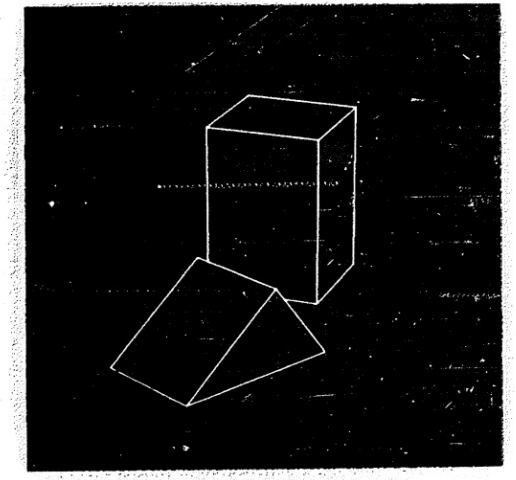
MACHINE PERCEPTION OF THREE-DIMENSIONAL SOLIDS

by

LAWRENCE GILMAN ROBERTS

Submitted to the Department of Electrical Engineering on May 10, 1963, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.





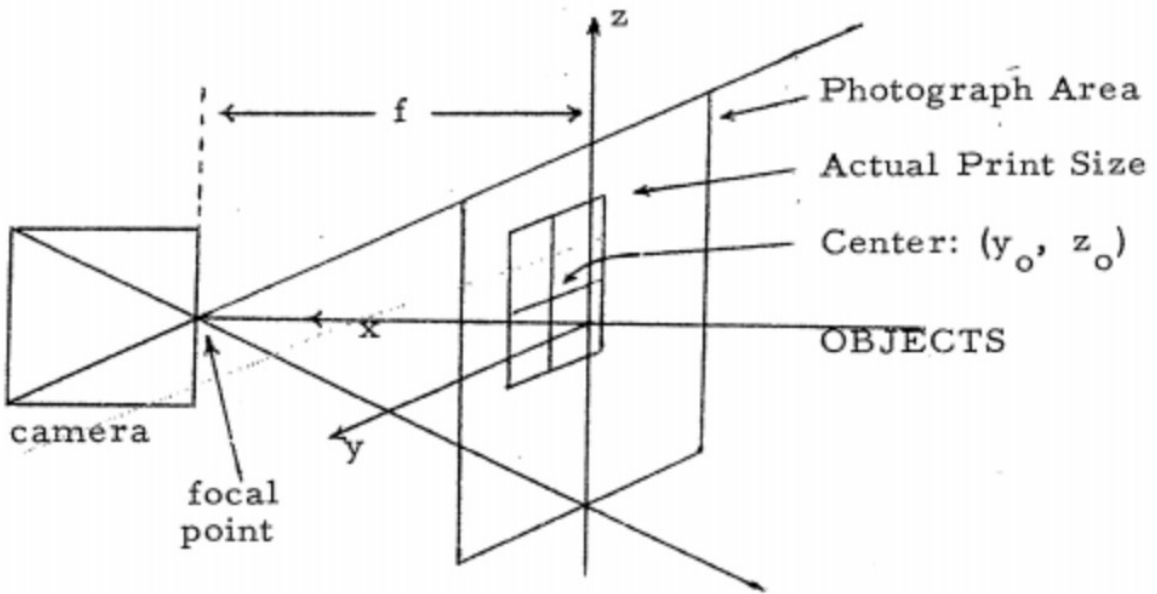


Figure 1: Camera Transformation

The Lumigraph

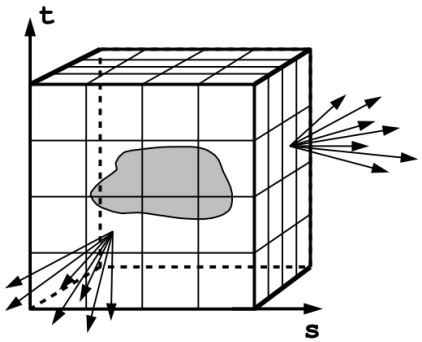


Figure 1: The surface of a cube holds all the radiance information due to the enclosed object.

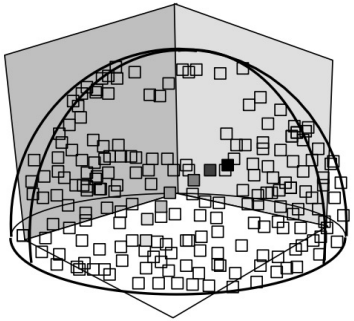


Figure 11: The user interface for the image capture stage displays the current and previous camera positions on a viewing sphere. The goal of the user is to "paint" the sphere.

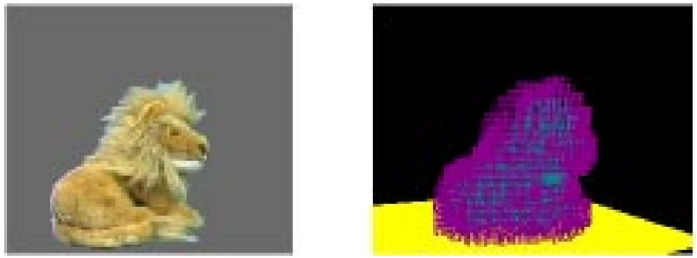
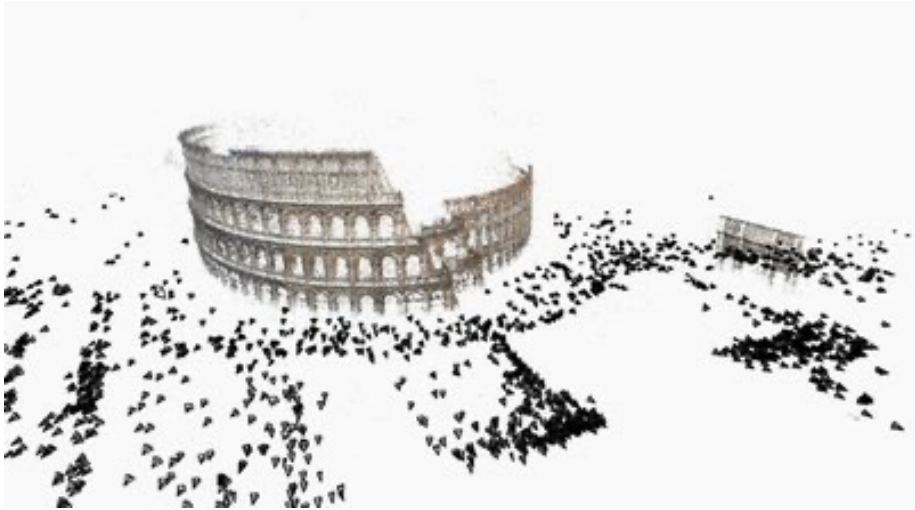


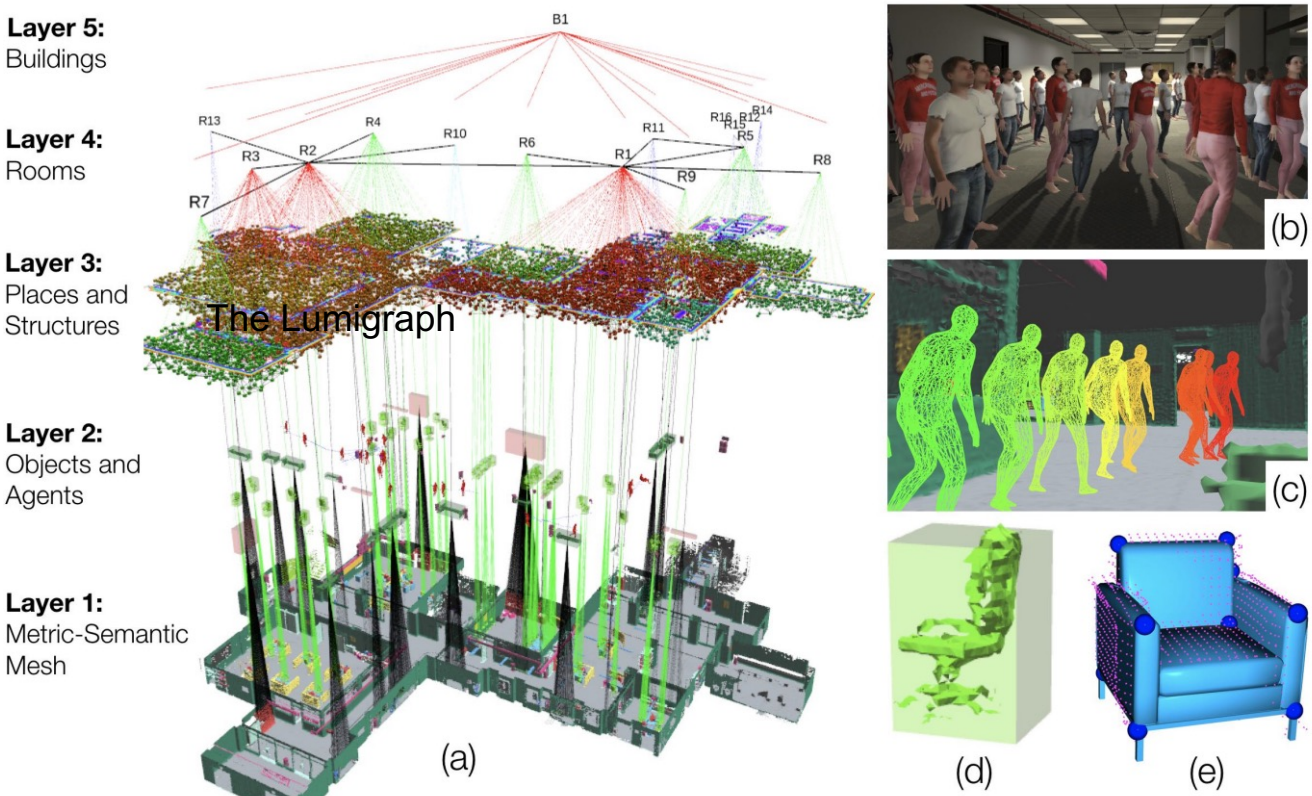
Figure 12: Segmented image plus volume construction

Building Rome in a Day



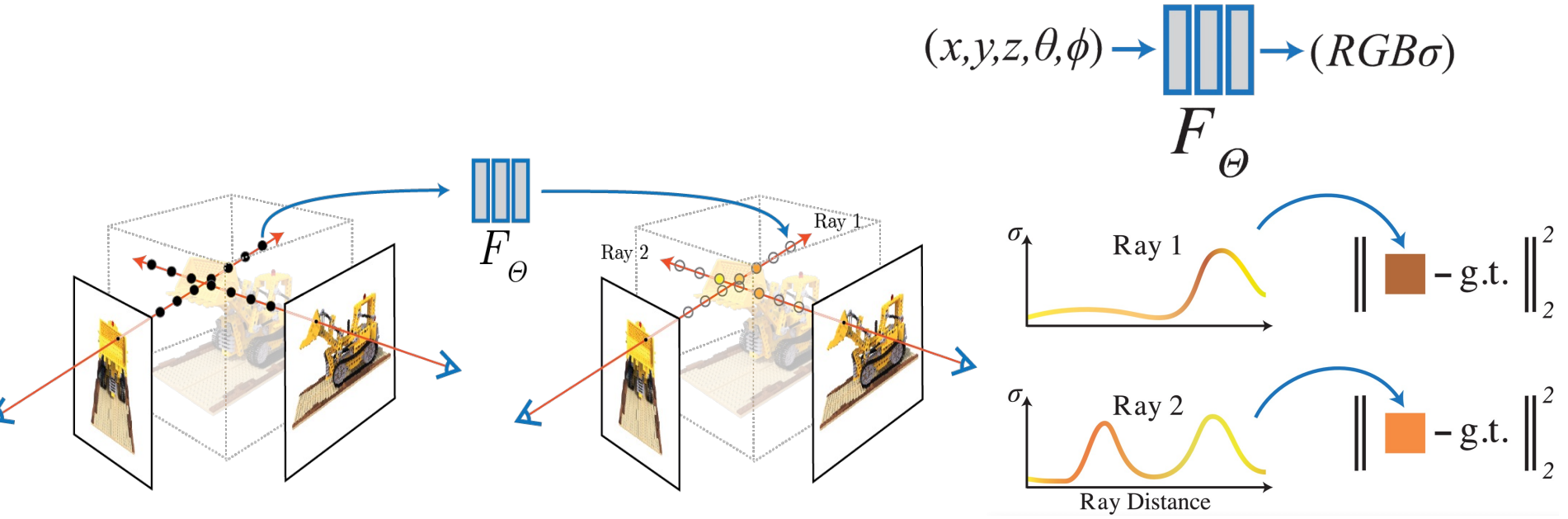
Building Rome in a day

3D Perception



3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans, TRO 2023

Neural Radiance Field



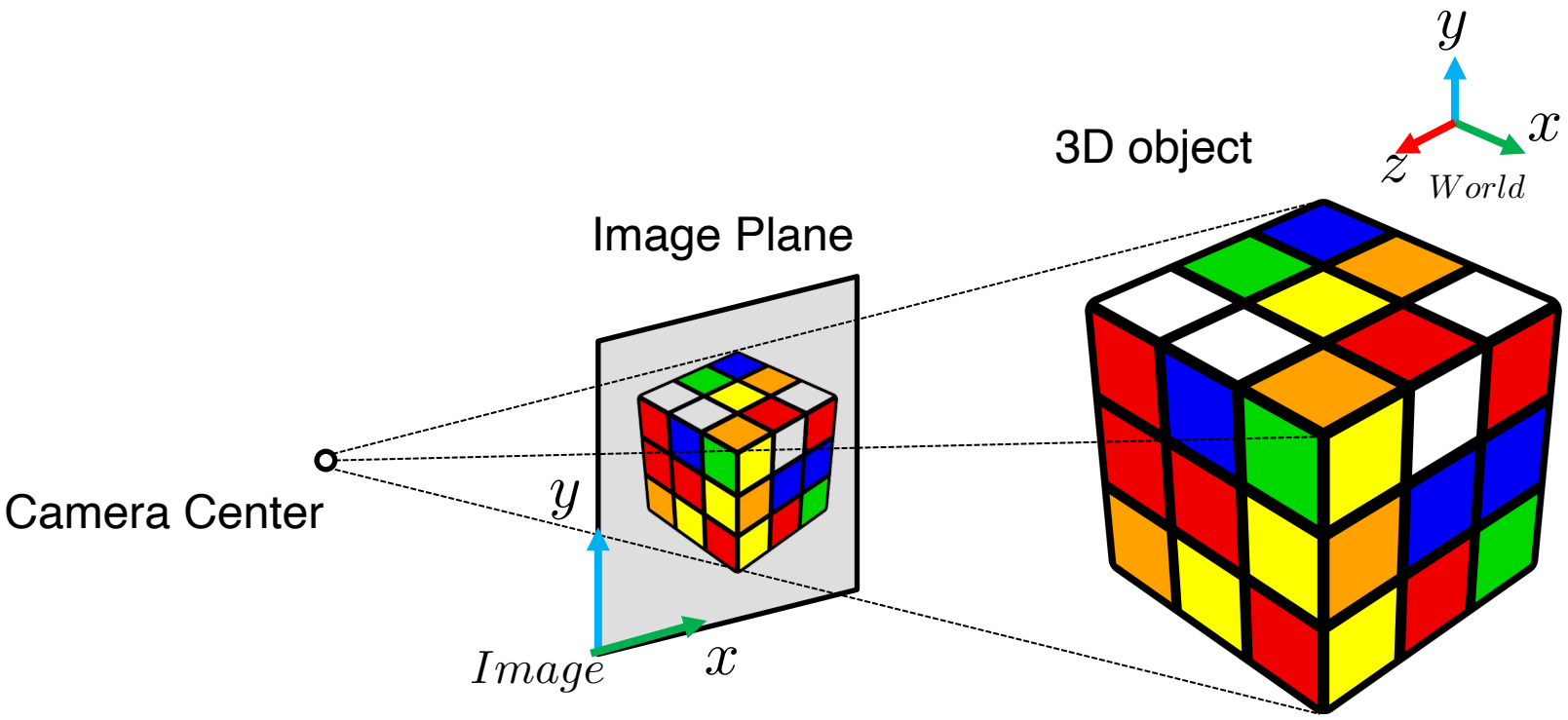
Midlenhall et al. Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV 2

Neural Radiance Field

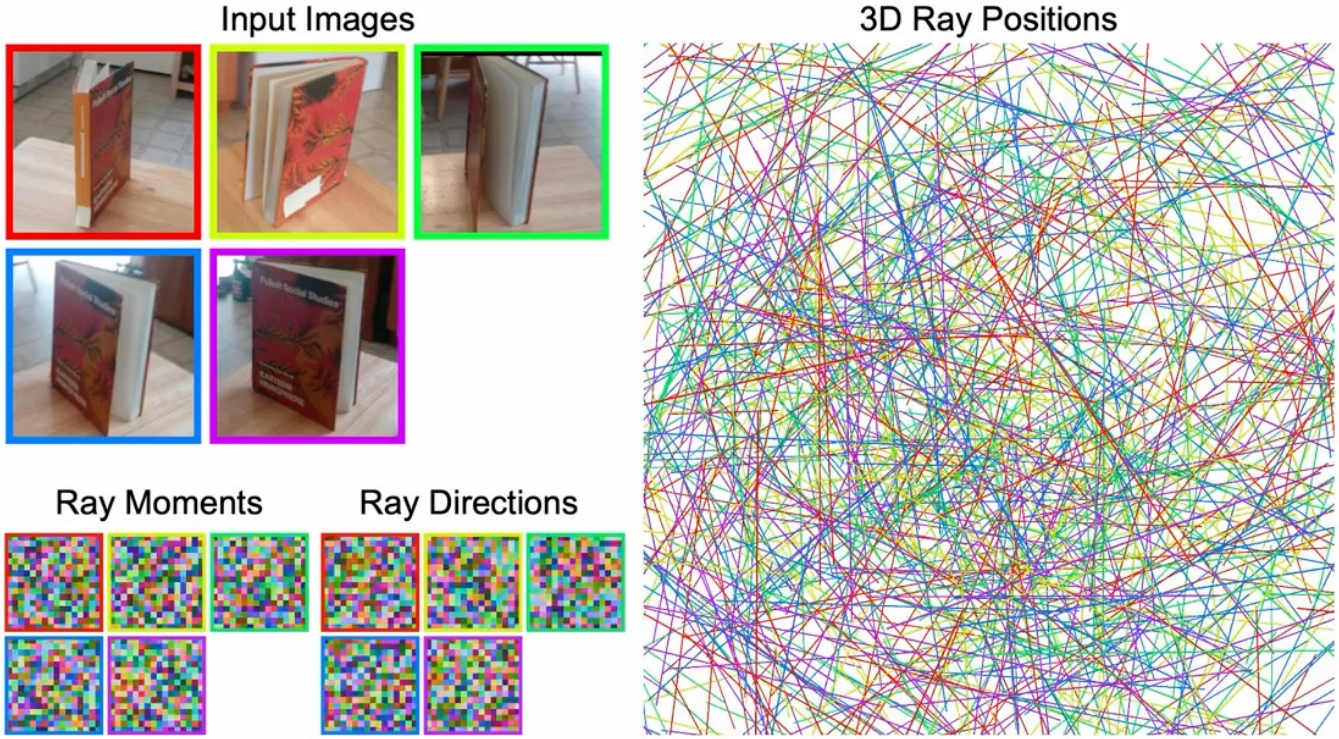
Zip-NeRF: Anti-Aliased Grid-Based Neural Radiance Fields

What topics we will cover in 3D vision?

Fundamentals

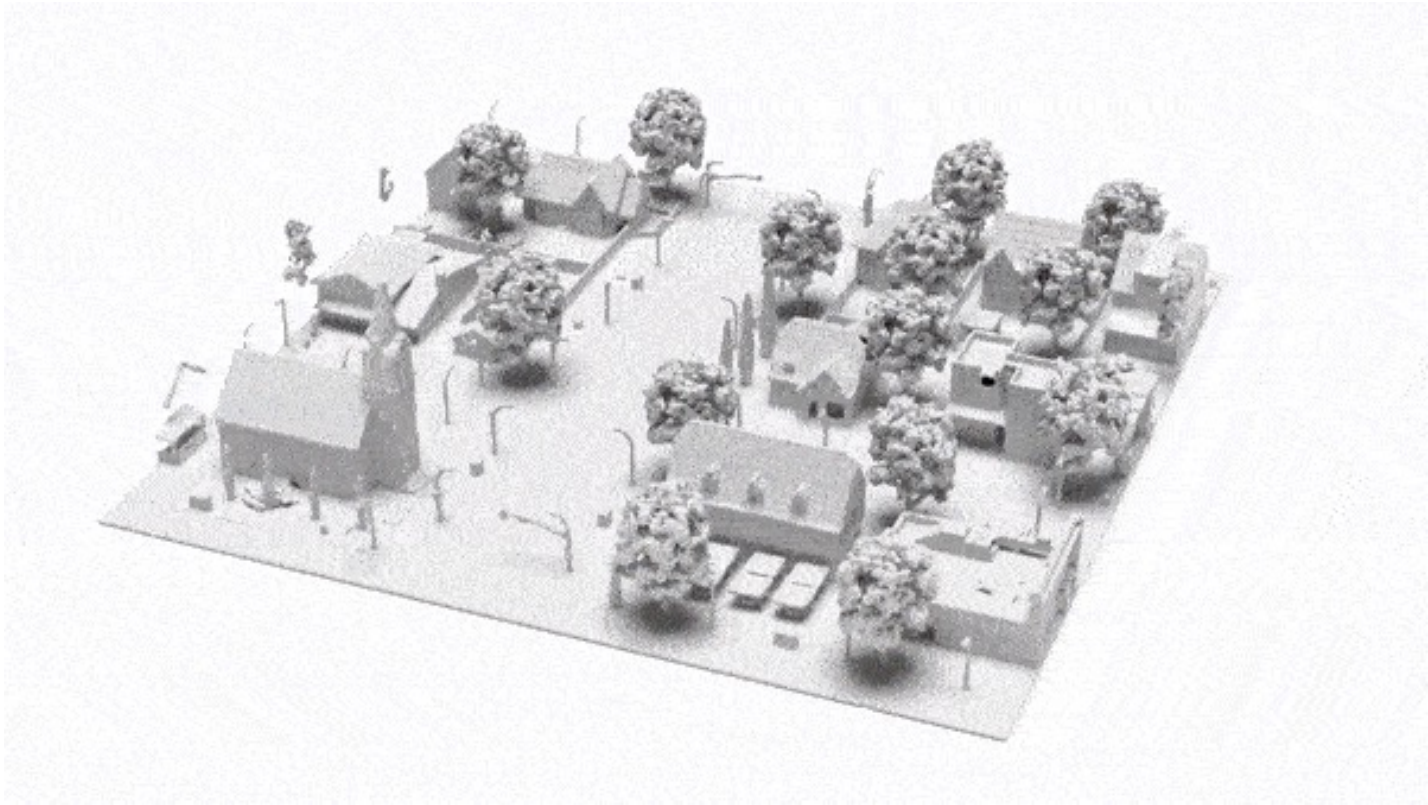


3D modeling from multi-views



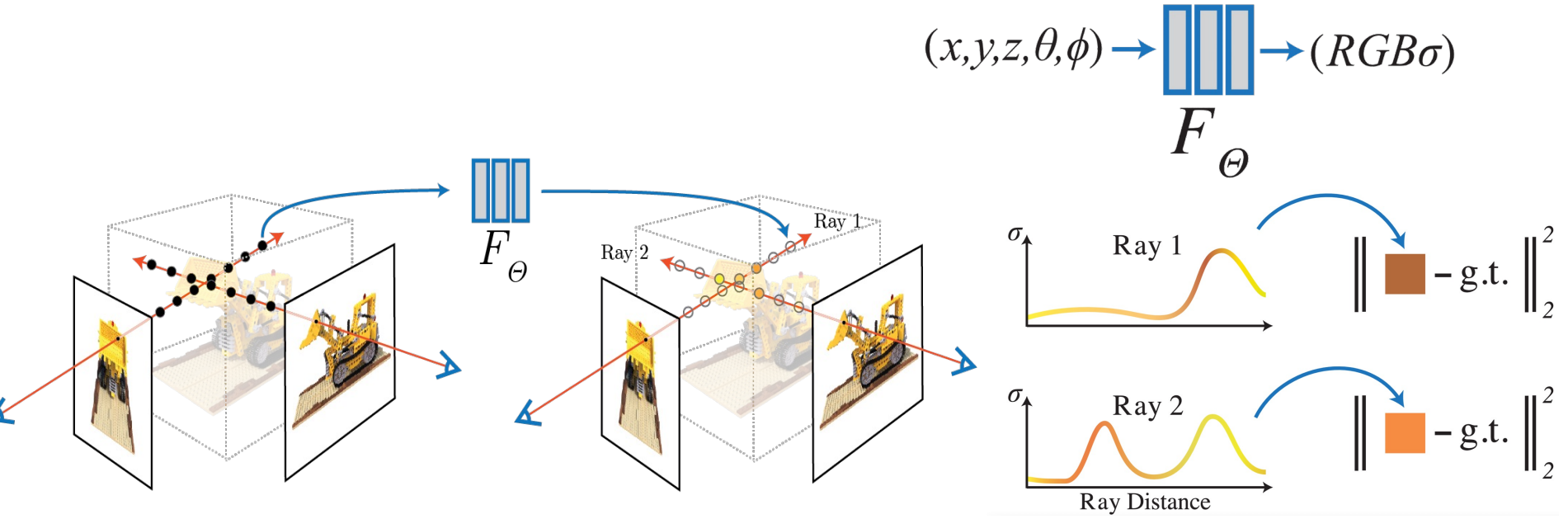
Ray Diffusion

3D representations



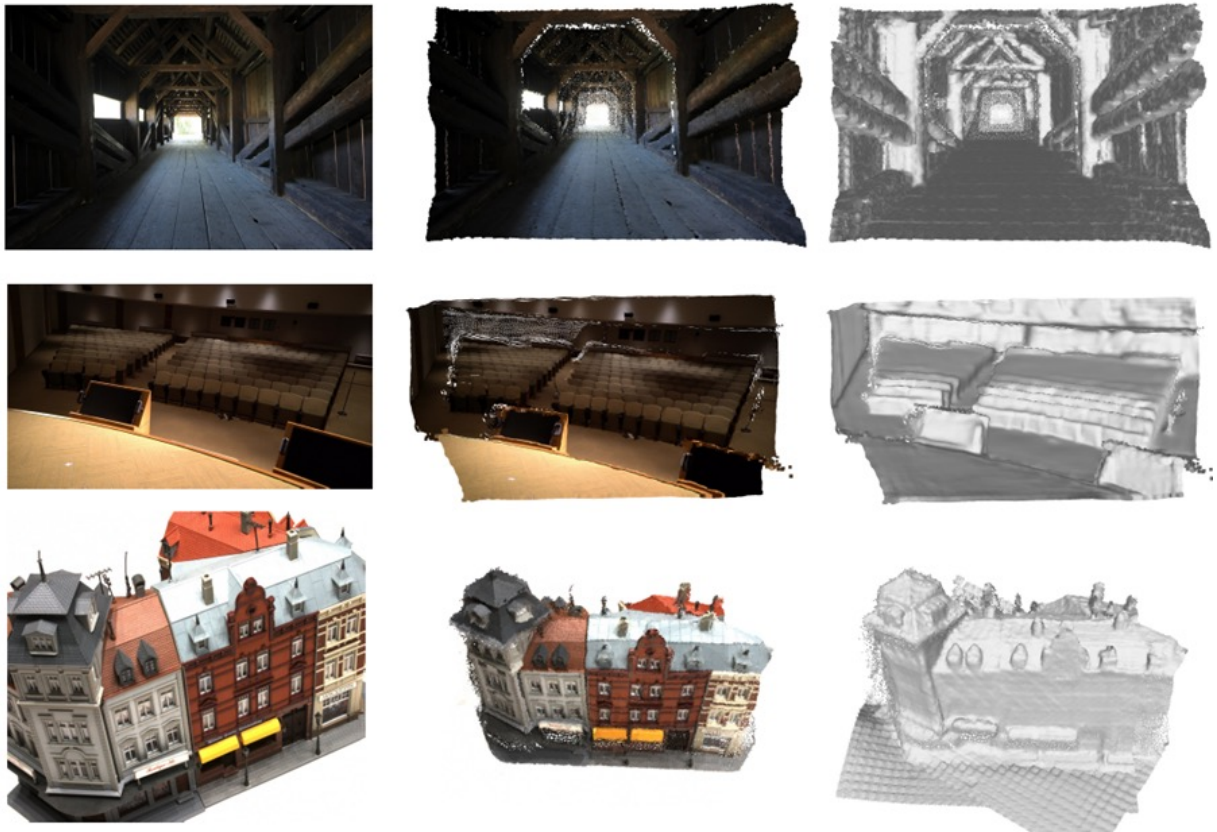
fVDB

Differentiable and Neural Rendering



Midlenhall et al. Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV 2020

Monocular and Few-View 3D



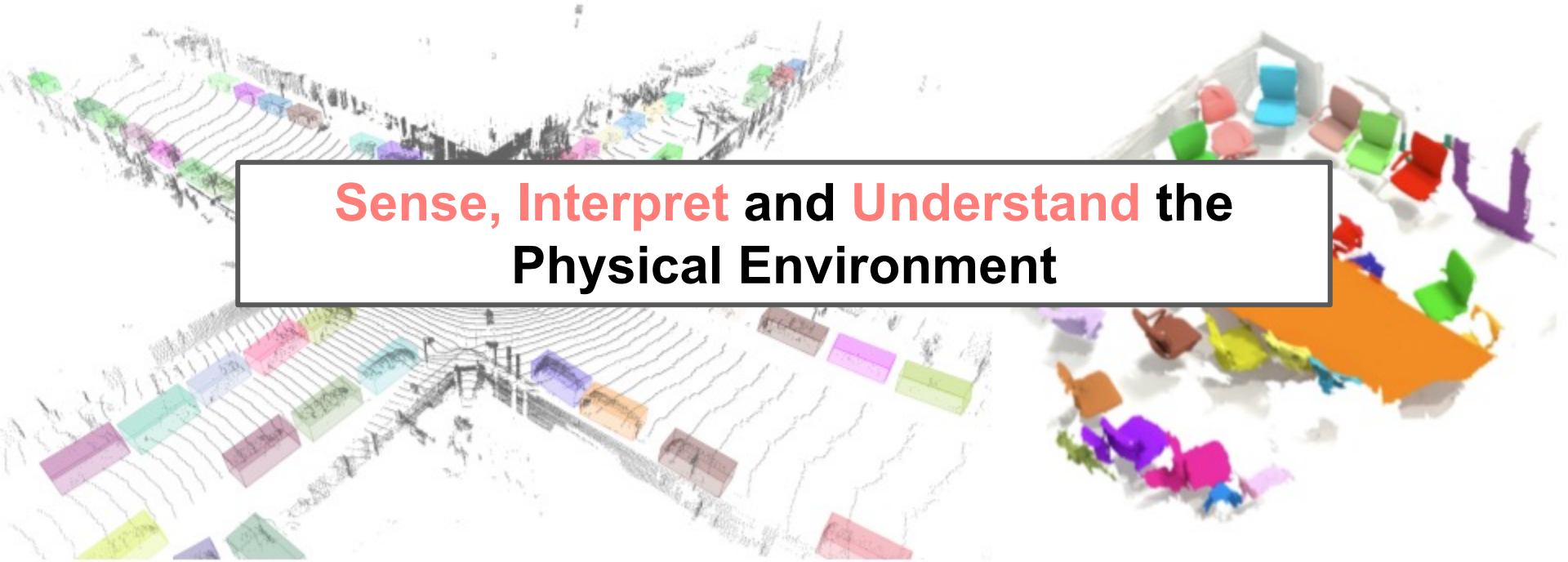
Dust3R

3D Content Creation

Dynamic 3D (4D)

Sapiens

Machine Perception



Sense, Interpret and Understand the Physical Environment

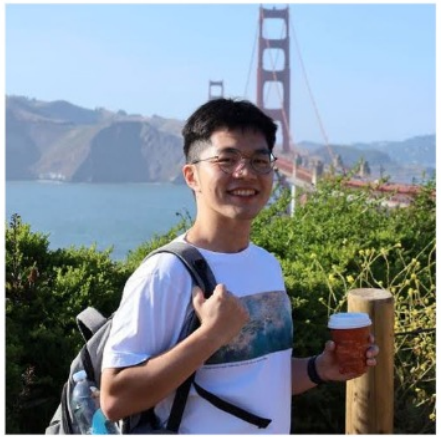
Image credit: Google, Waymo Open Dataset, ScanNet

Logistics

CS498 Staff and Office Hours

Shenlong Wang

Instructor



Zhi-Hao Lin

TA

Schedule

- **Time & Location**: Tue/Thurs 2:00pm-3:15pm (Siebel 0216)
- **Office Hour:** Thurs 3:15pm - 4:30pm (Siebel 4124)
- **Website**: https://courses.grainger.illinois.edu/cs598shw/fa2024/index.html
- **Syllabus**: https://shorturl.at/ySWr7
- **Schedule**: https://shorturl.at/wOh6k
- **Slack**: https://shorturl.at/jV1NL

Prerequisites

- Graduate-level understanding of computer vision (equivalent of **CS543** or **CS445+CS444** combined) , including camera models, image filters, two-view geometry, feature detection and matching, and recognition.
- Graduate-level understanding of machine learning (equivalent of **CS446**), including (stochastic) gradient descent, loss functions, optimization, neural network, generative models (GANs, diffusion, etc).
- You should be ***engaged in or interested in research in 3D vision***.

The class is NOT

- Introductory computer vision class: We might briefly review some of the topics but will not introduce them.
- Assignment-based: only mini-quizzes. Gradings are heavy on role-playing and final projects.
- All lectures: Everyone is expected to lead discussions every few lectures.
- Low-effort / Just coming to class: This course requires significant engagement and preparation during and before class.
- Remote / Hybrid: It’s highly participatory and focuses on getting to know your peers and engaging in discussion.

Student Deliverables

1. 35%: Role-play Discussion
    1. 25%: Presentations (5% each, tentative)
    2. 10%: Hacker's Deliverables (code, demo, presentation)
2. 10%: Research Topic Survey (4 pages, assigned group of 3-4)
3. 40%: Final Project (self-formed group)
    1. 10% Proposal (3-4 pages)
    2. 15% Poster & Demo Presentation
    3. 15% Final Report (6-8 pages)
4. 10%: Quiz (2.5% each, mini-summary or illustration)
5. 5%: Participation in class and on slack

Grading Policy

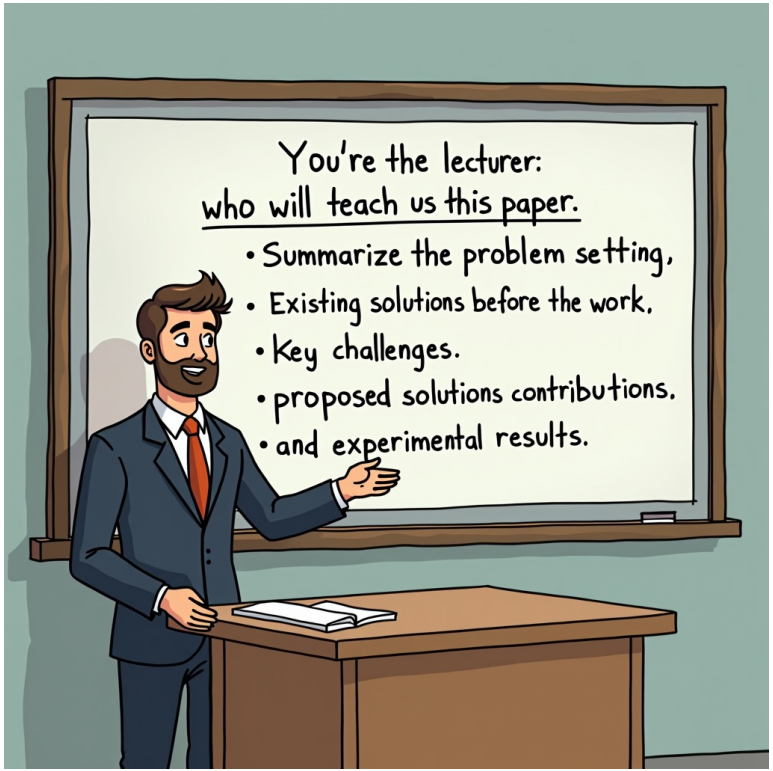
Grading policy: Each deliverable will be rated as “Satisfactory” (full credit), “Needs Improvement” (3/4 credit), or “Unsatisfactory” (1/2 credit). Project deliverables, quiz and topic survey can be resubmitted once without penalty if rated as below “Satisfactory”. I will provide feedback below “Satisfactory” to help you understand how to further improve. Truly exceptional submissions that exceeds expectations might get a “Exceptional” rate – it is *not normal* to get “Exceptional” rating.

Bonus point (capped at 5% in total per person) will be given for “Exceptional” deliverables (1% each), peer-rated best poster/demo award (2%), most engaged / helpful student in class and slack (2%).

Role-Playing Discussion

Lecturer

You’re the lecturer who will teach us this paper. Summarize the problem setting, existing solutions before the work, key challenges, proposed solutions and contributions, evaluation, and experimental results.



Generated by Dall-E

Role-Playing Discussion

Archeologist

You’re an archeologist who must determine where this paper sits in the context of previous and subsequent work. Find and report on one older paper cited within the current paper that substantially influenced it, and one newer paper that cites this current paper. Discuss the relationship between the papers.

Generated by Dall-E

Role-Playing Discussion

Private Investigator

You’re a private investigator. Find out background information on one of the paper's authors. Where have they worked? What did they study? What contexts might motivate them to work on this problem? What previous projects might have led to their work on this one?

Generated by Dall-E

Role-Playing Discussion

Hacker

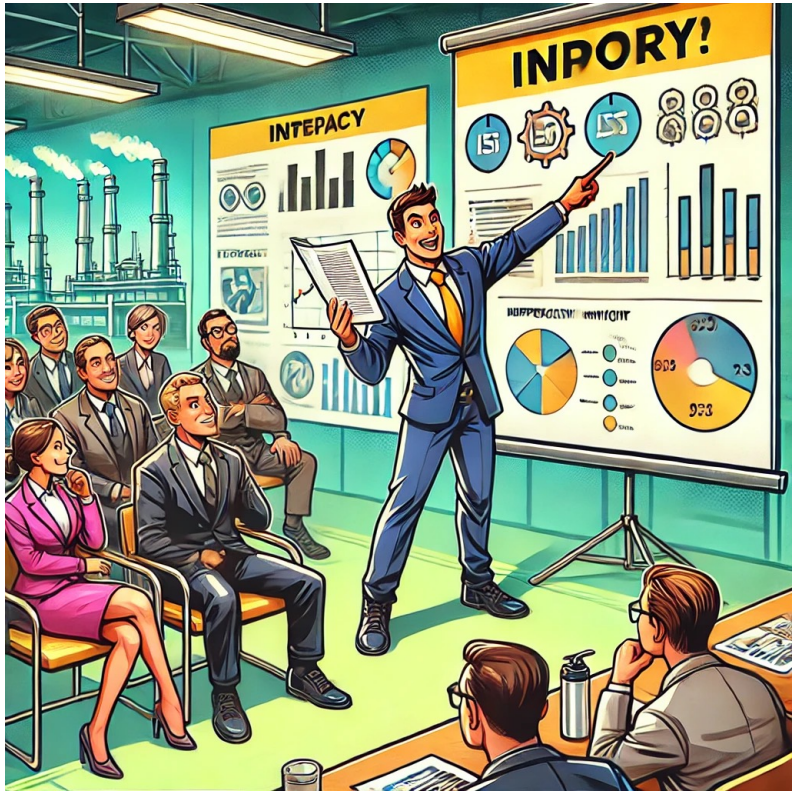
You’re a hacker who needs a demo of this paper ASAP. Implement a small part or simplified version of the paper on a small dataset or toy problem. Prepare to share the core code of the algorithm with the class and demo your implementation. Do not simply download and run an existing implementation.

Generated by Dall-E

Role-Playing Discussion

Product Manager

You work at a company or organization developing an application or product of your choice (that has not already been suggested in a prior session). Bring a convincing pitch for why you should be paid to implement the method in the paper, and discuss at least one positive and negative impact of this application.



Generated by Dall-E

Role-Playing Discussion

Critic

You’re a critic who challenges ideas and decisions with a skeptical eye, aiming to find weaknesses, question assumptions, and ensure everything is thoroughly tested. You might act as a negative reviewer, a decision-maker who rejects the product manager's request; a hacker on the dark side who stress-tests the system; or an advocate for a different approach.

Generated by Dall-E

Role-Playing Discussion

Graduate Student

You’re a graduate student working on a new project in this area. Propose two imaginary follow-up project ideas, not just based on the current paper, but possible only due to its existence and success.

Generated by Dall-E

Mini-Quiz

- We will have four ***small post-class quizzes*** after lecture-based classes.
- Each quiz will consist of a single question that summarizes topics covered in class. Answering it will require some post-class investigation, with a lightweight deliverable (such as writing a 1-2 paragraph summary or creating a graph illustration).
- Due within one week after the class. For example, a quiz given on Thursday, August 29, will be due by 1:59 pm on Thursday, Sept 5.
- After answering all the four quizzes, you will get a one-page cheat sheet on 3DV!

Final Project

- Most important: have **FUN**!
- Need to be different from existing works, but not necessarily need to solve AI ;-)
- Do something cool/fun/useful, demonstrate techniques you learned from the class.
- Artsy stuff, cool video / interactive demos, non-traditional problems are particularly encouraged! Do something you would love to put on your website.
- Start prototyping **EARLY**, exploit open-source tools.
- Doesn’t have to be your next CVPR / ICCV (I hope some groups will make it there).

Topic Survey

- You will select a topic to explore with a small group (3-5 members, instructor assigned), perform a literature review, and write a survey (4 pages)
- The survey includes taxonomy of key design decisions and techniques, a summary of evaluations, an assessment of current capabilities and gaps, and the identification of new research ideas.
- Submitted to the class Github repo and shared with everyone.

Slack

- **We will use Slack for communication.**
- Important announcements will be sent on both Slack and your email account.
- One of the course staff will monitor Slack during 10pm –11pm every day and address all the questions posted by the students.

Academic Integrity

- All work you submit should be your own – do not copy any text from any online reviews or papers – Cite sources diligently
- If your research project builds on prior/ongoing work, discuss with professor first
- Violations will be penalized through official channels

Teamwork

- For teamwork (survey and course project), take notes on each member's contributions in the last of the write-ups.
- Be a friendly and reliable team member.
- Don't be a hitchhiker and know how to handle these situations if they arise in your group.

https://www.cs.cornell.edu/courses/cs3110/2018fa/teams/hitchhikers.html

Todo List

- Join Slack
- Check course website and syllabus
- Start forming final project group
- Enjoy the rest of the day