
mit
media
lab

Learning From Visual Cues

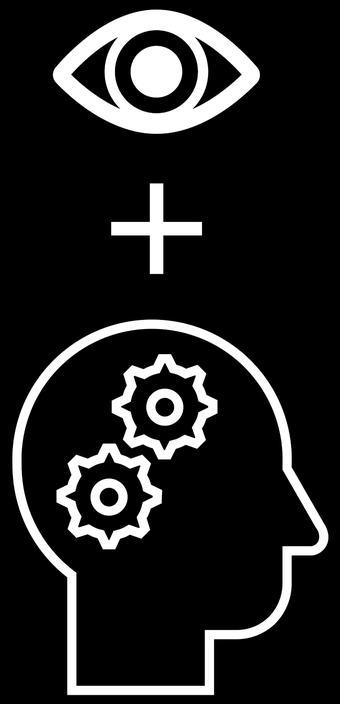
How many babies are in the scene?

How many babies are in the scene?

How many people are in this scene?

We exploit Visual Cues

Visual Cues are ubiquitous



Model Cues

Process Cues

- Exploit Cues
- Ignore Cues

Kushagra Tiwary
Camera Culture,
MIT Media Lab, MIT

AI must process visual cues

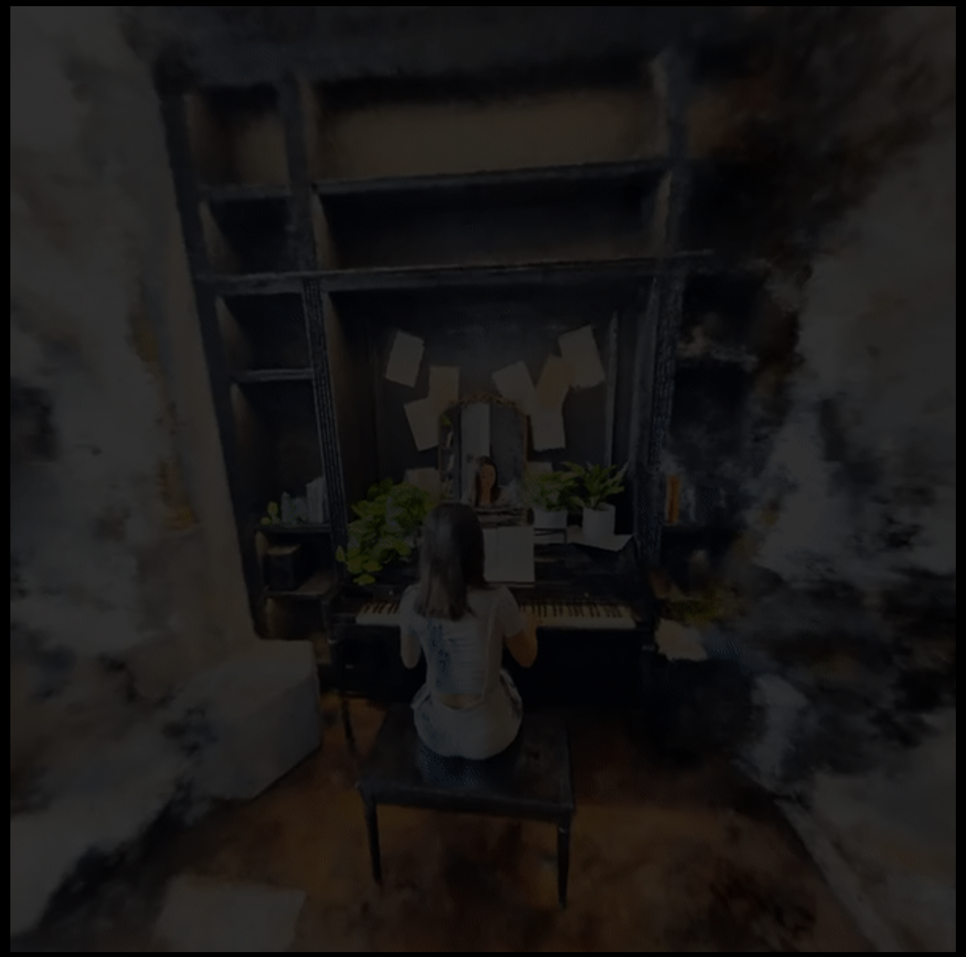
Machines **must** process complex **visual cues** to **interact** with the **physical world**

Top: 3D reconstruction for a room with mirrors

Left: Robot fails to pick up a transparent ball

AI must process visual cues

Machines **must** process complex **visual cues** to **interact** with the **physical world**

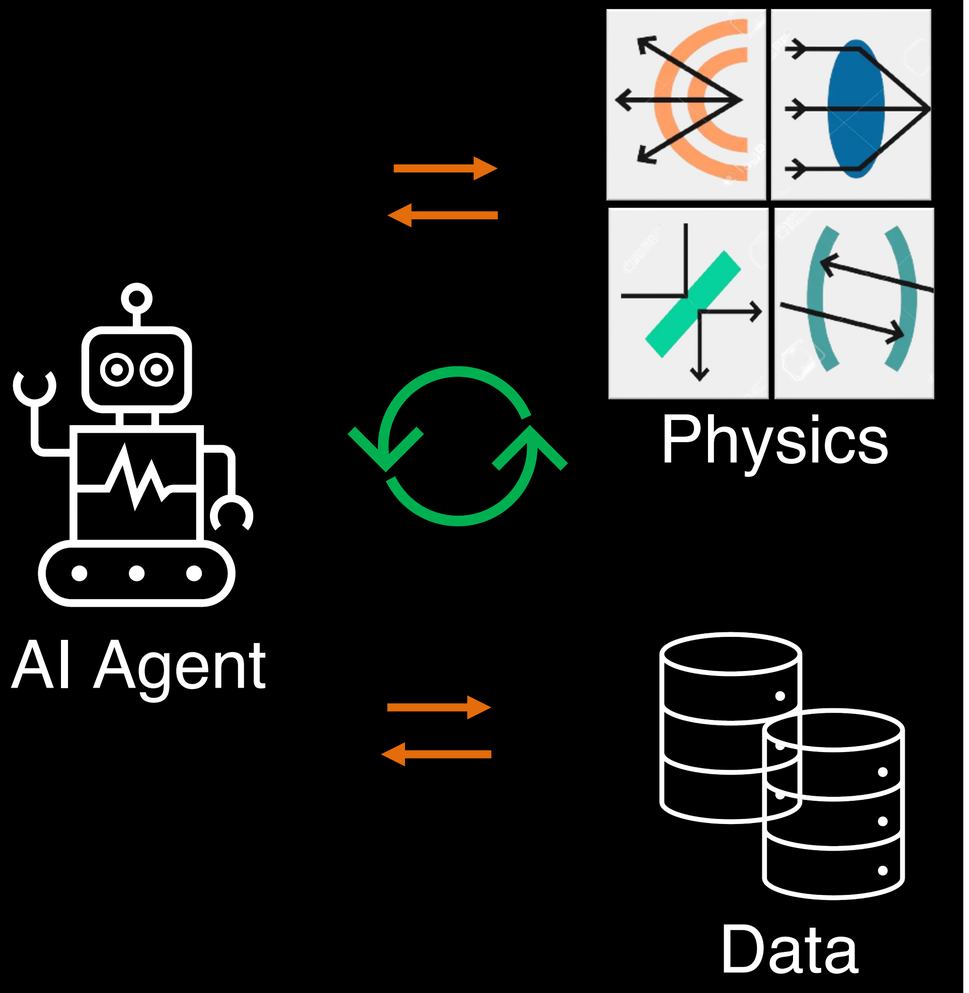


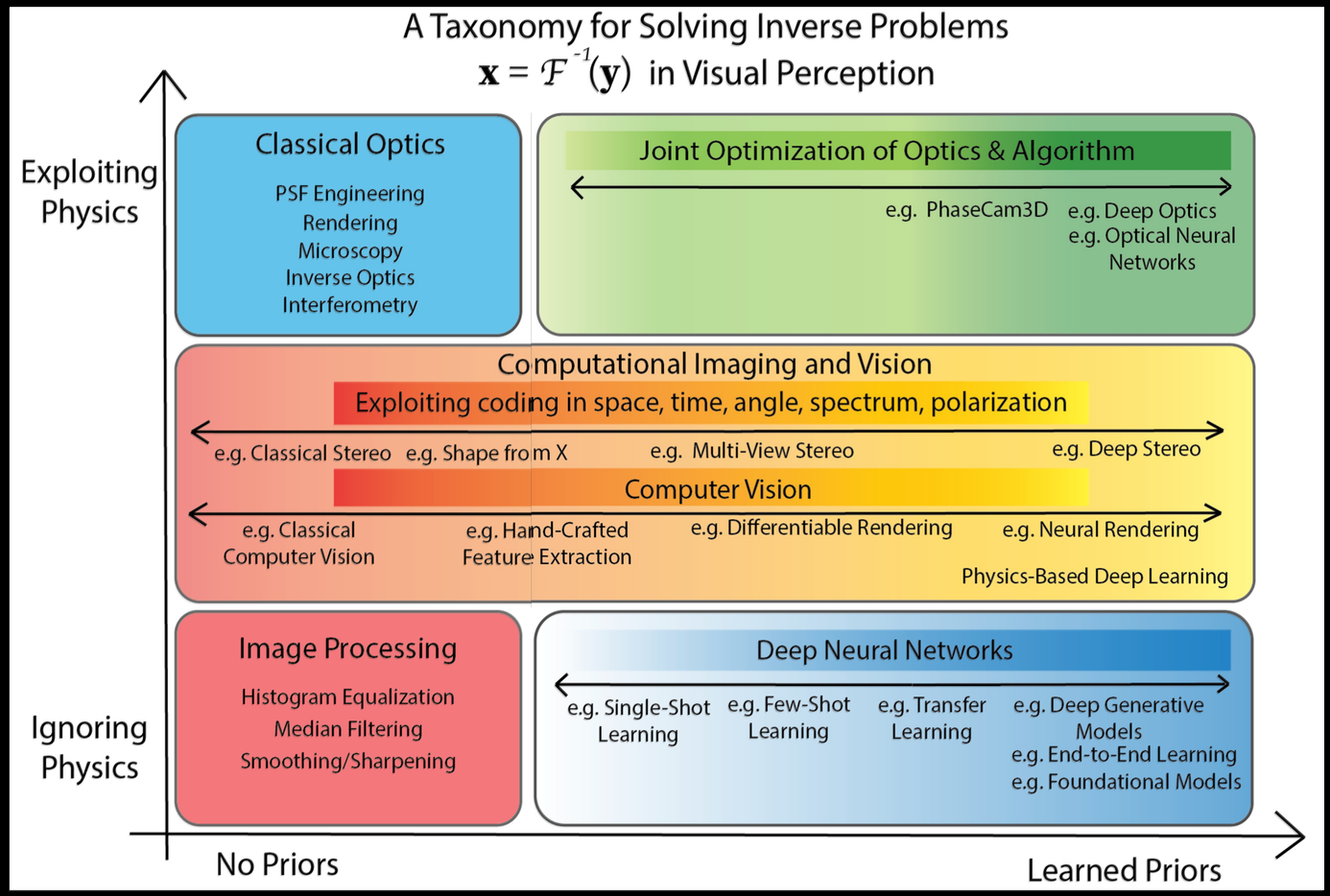
Top: 3D reconstruction for a room with mirrors

Left: Robot fails to pick up a transparent ball

Kushagra Tiwary
Camera Culture,
MIT Media Lab, MIT

Exploiting Physics with Machine Learning





Top: Perspective and Review on *designing* future Perception Systems that exploit physics and machine learning[1]

[1]Klinghoffer*, Somasundaram*, **Tiwary***, Raskar; Physics vs. Learned Priors: Rethinking Camera and Algorithm Design for Task-Specific Imaging, *ICCP 2022*

Timeline

Physics vs. Learned Priors: Rethinking Camera and Algorithm Design for Task-Specific Imaging, *ICCP 2022*

Perspective Review

March 2022

April 2022

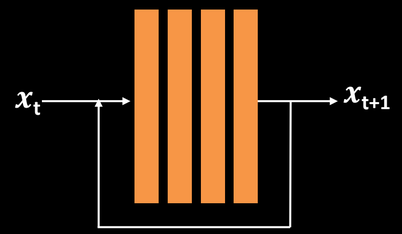
Physical Cue: Shadows

Towards Learning Neural Representations through Shadows, *ECCV 2022*

Nov 2022

Physical Cues: Reflections, Polarization

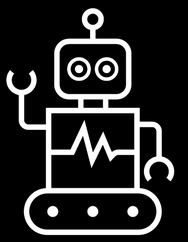
Objects as Radiance Field Cameras, *Under Review 2022*



Physical Cues: Shadows, Reflections, Polarization

Scenes as Cameras, *Proposed Work*

March 2023



Direct Applications of Cues for Robots

April 2023

PhD? ☺

Learning Neural Representations from Shadows

Problem: What information can we recover from shadows in images? [1]

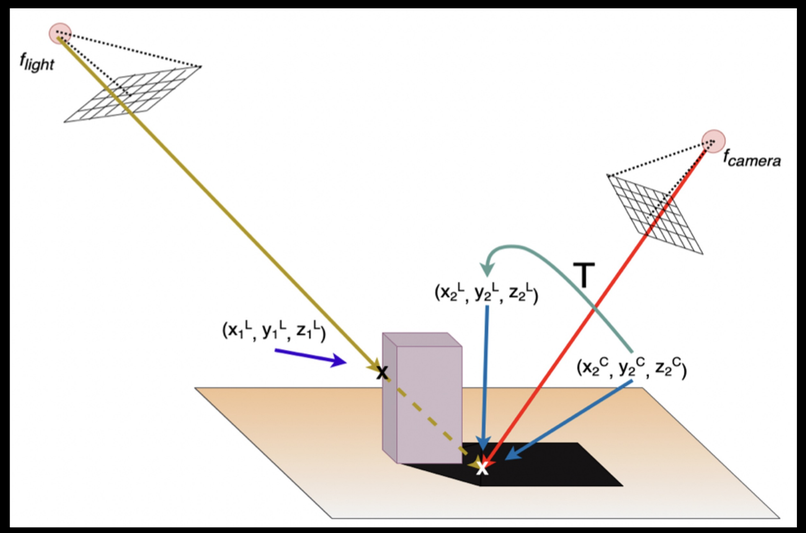
Top: Cast shadows present information otherwise hidden from direct line-of-sight

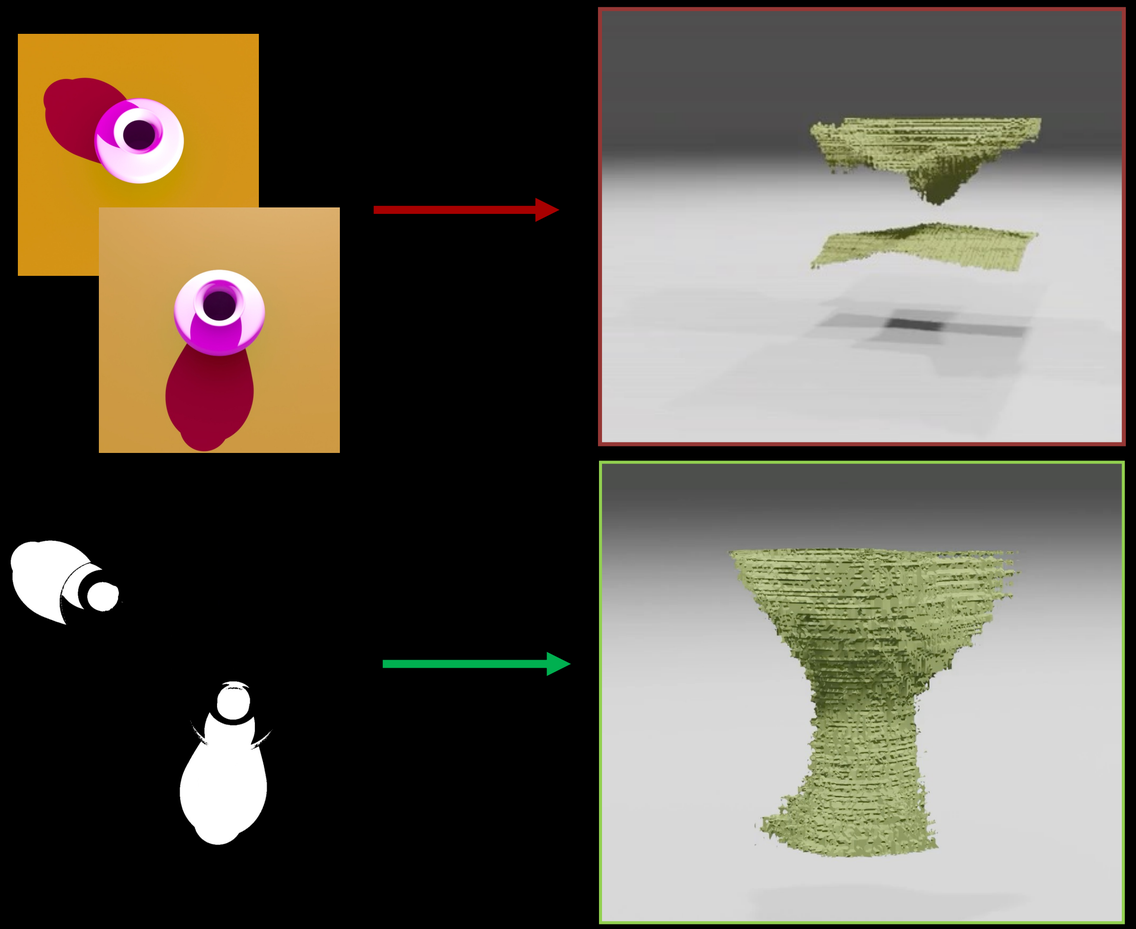
Tiwary et al; Towards Learning Neural Representations from Shadows, *ECCV 2022*

Kushagra Tiwary
Camera Culture,
MIT Media Lab, MIT

Learning Neural Representations from Shadows

Key Insight: Differentiable forward-models that can learn to render shadows[1]





Top: We can recover hidden information from the scene by ***only*** exploiting shadows casted in the scene

Left: Differentiable Shadow Mapping helps ML learn from data

[1]Tiwary, et. al; Towards Learning Neural Representations from Shadows, *ECCV 2022*

Kushagra Tiwary
Camera Culture,
MIT Media Lab, MIT

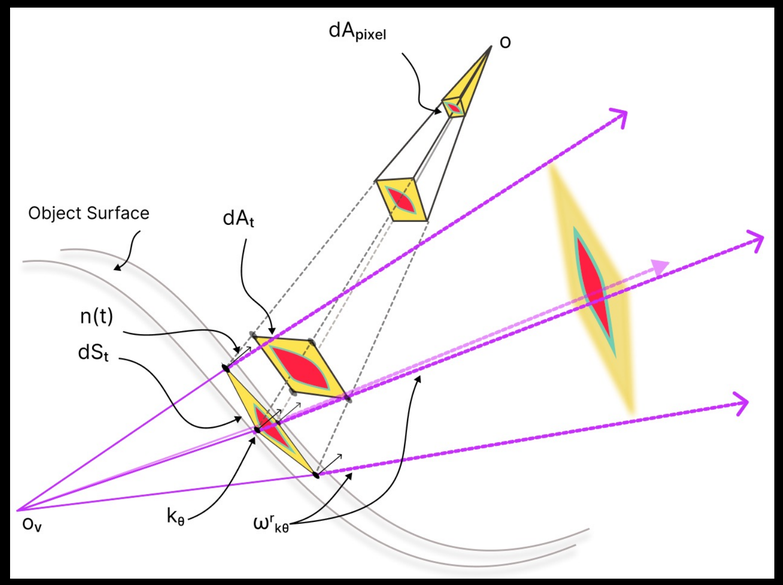
Objects as Radiance Field Cameras

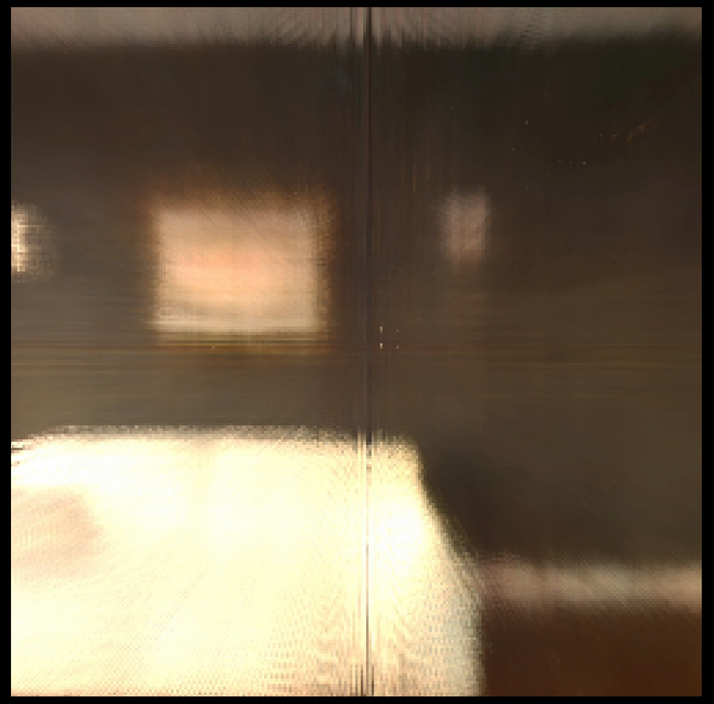
Problem: What information can we recover from reflections on everyday objects in images?

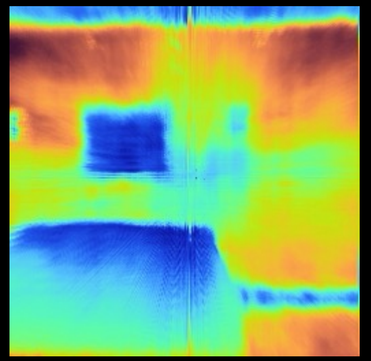
Top: Objects encode information about their environment by reflecting light

Objects as Radiance Field Cameras

Key Insight: Modeling reflections as a projection of the 5D radiance field of the environment onto the object surface[1]





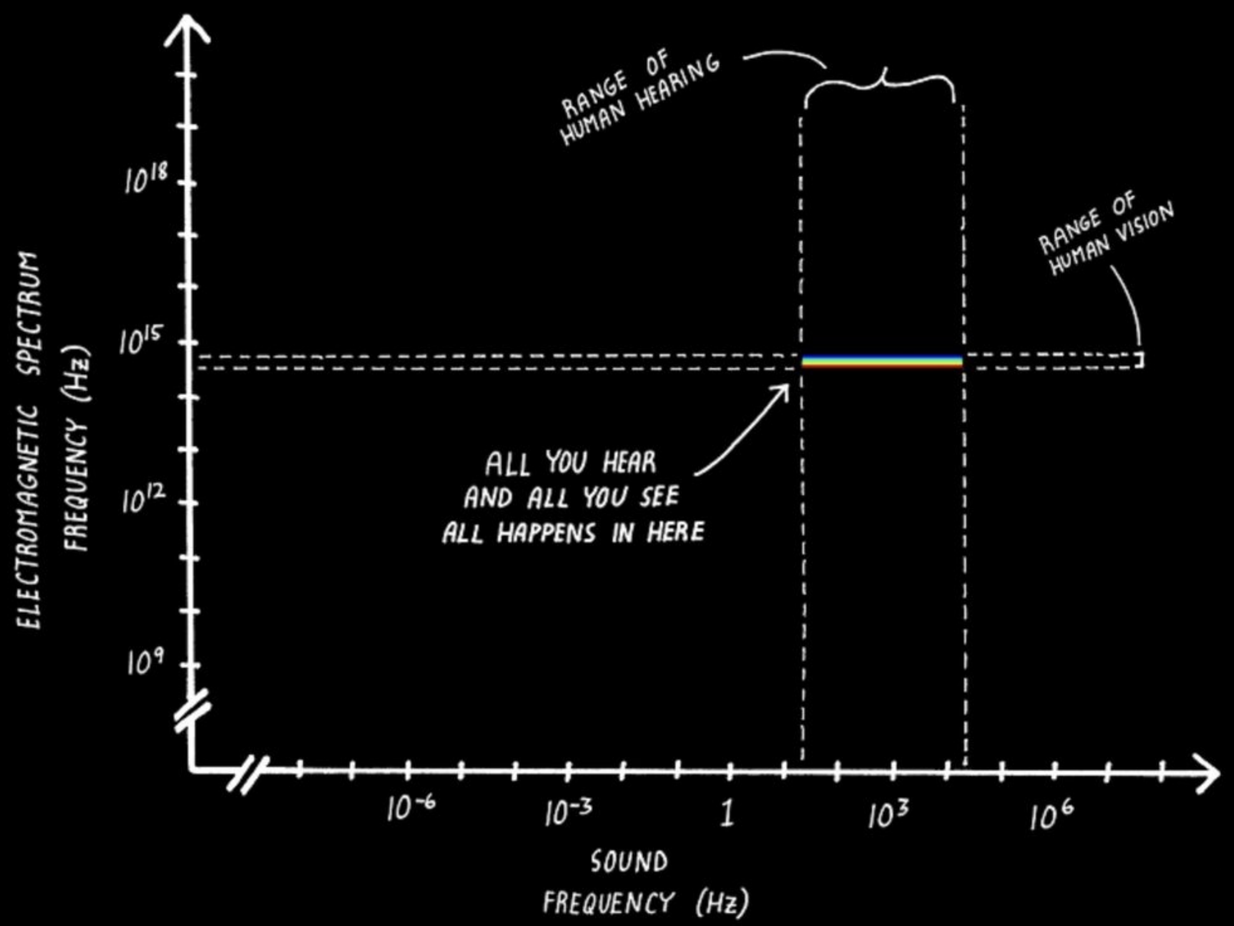


Top: By exploiting reflections we can see the world from the object's perspective

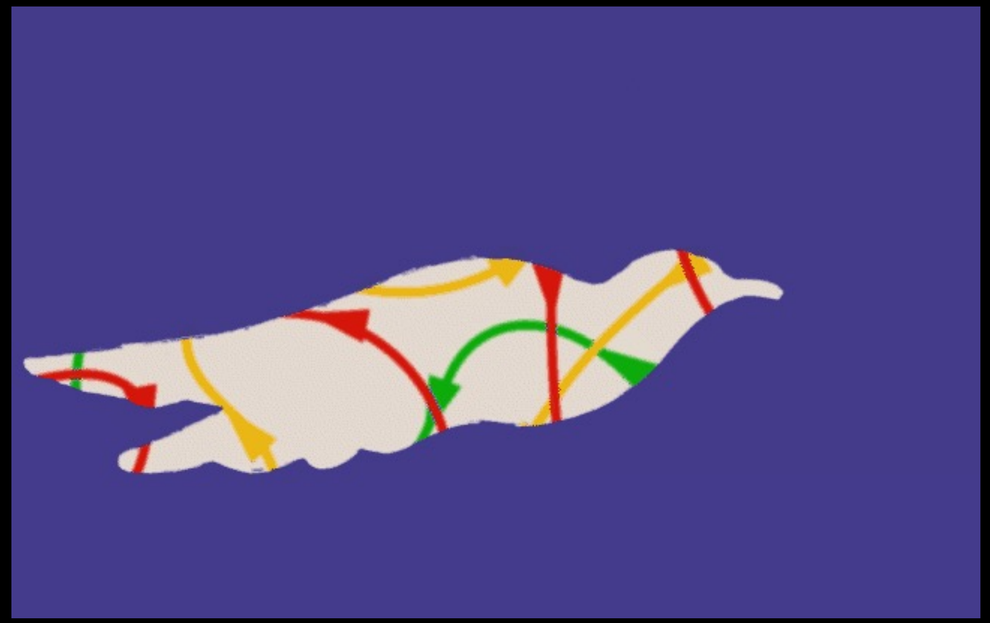
Left: Imaging through Virtual Sensors on the object's surface

[1]Tiwary, et. al.; ORCa: Glossy Objects as Radiance Field Cameras, *Under Review 2022*

Why limit machines to Human-Like Cues?

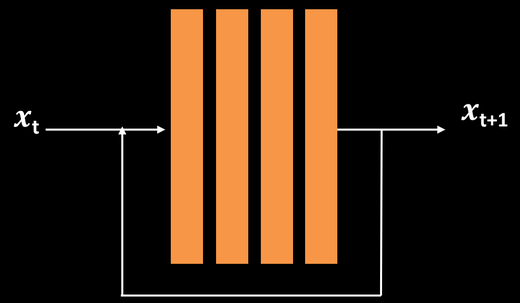


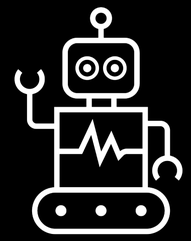
Top: Going beyond the visual spectrum to sample the world



Top: Birds use Polarization as a cue to navigate

Timeline





Physics vs. Learned Priors: Rethinking Camera and Algorithm Design for Task-Specific Imaging, *ICCP 2022*

Perspective Review

Physical Cues: Shadows, Reflections, Polarization

Scenes as Cameras, *Proposed Work*

Applying Cues with downstream Robotics

March 2022

April 2022

Nov 2022

March 2023

April 2023

Writing

Physical Cue: Shadows

Towards Learning Neural Representations through Shadows, *ECCV 2022*

Physical Cues: Reflections, Polarization

Objects as Radiance Field Cameras, *Under Review 2022*