Stealing from Nature

Manifolds and Models

Jack Naylor

PhD Candidate
Australian Centre for Field Robotics

MUGS Aca-dustry Seminar
Friday 6/5 4pm

THE UNIVERSITY OF SYDNEY

ACFR AUSTRALIAN CENTRE FOR FIELD ROBOTICS

A little about me...

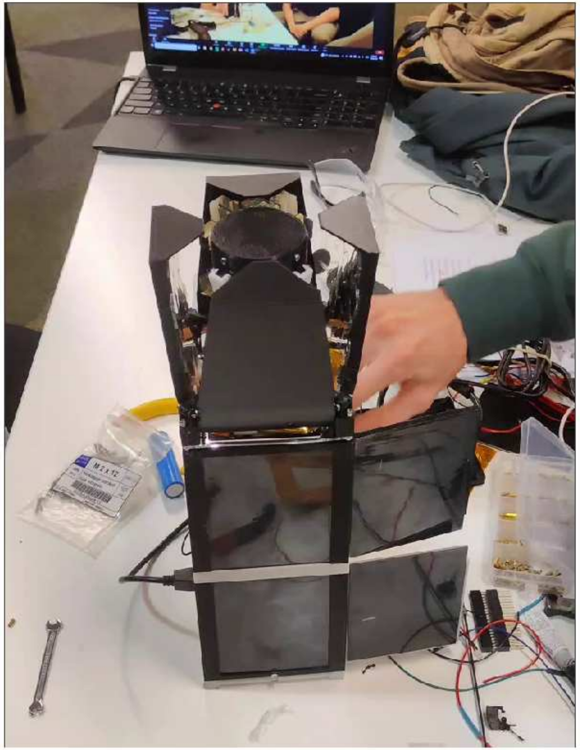
PhD Candidate @ ACFR working on: *"Simultaneous Localisation and Mapping Through Neural Radiance Fields"*

BE (Mechanical) (Hons. I)/BSc (Advanced)
Majored in Space Engineering & Physics

UG Thesis @ Nearmap

Things I Work On!

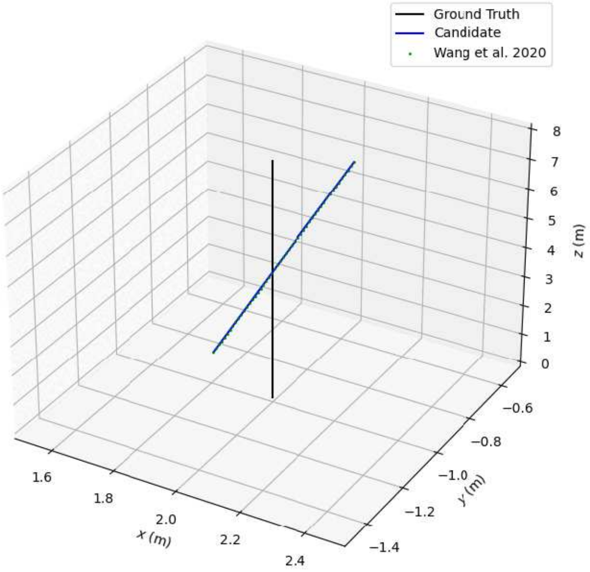
- Robotics
- Remote Sensing
- Embedded AI
- Perception & Sensing



My UG Thesis

SA.Adelaide.302 – Flagpole at RAAF Edinburgh

Comparison to State of The Art



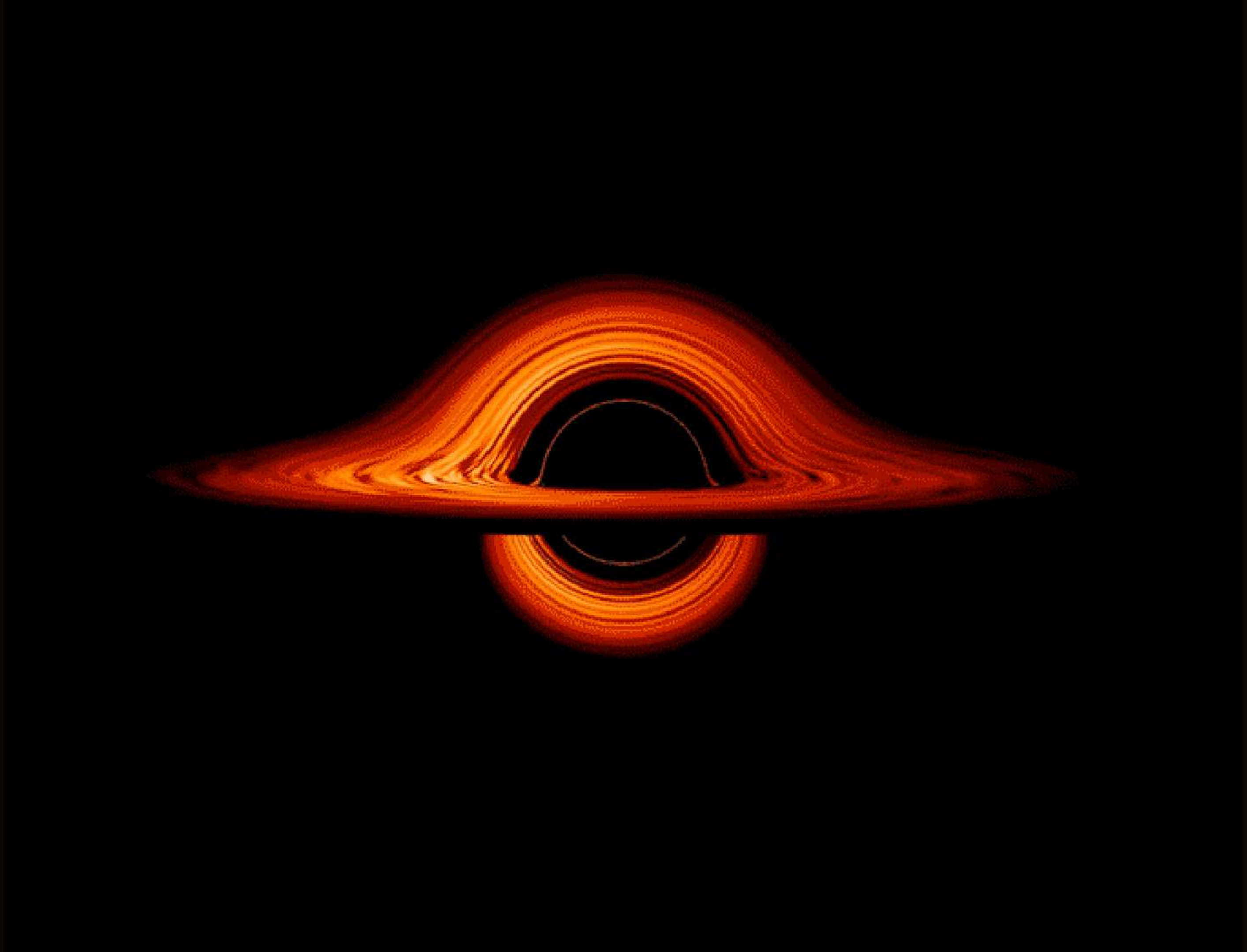
- Benchmark on simulated imagery from 4.5km altitude
- Candidate features 8m long, single pixel wide
- 60 observations

	Ours^	Ours*	Wang et al. (2020)*
Time (s)	**1.52**	**43.25**	387.42
Iterations	**436**	**2572**	32486
Error† (m)	**0.005**	**0.013**	0.096

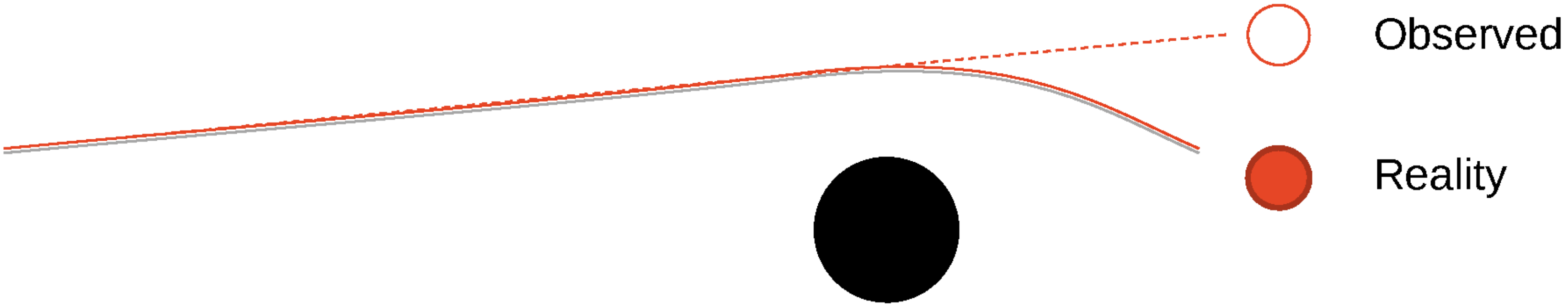
^Implementation in C++ (Ceres Solver)
*Implementation in Python
†Root mean squared error (RMSE) in xy

All models are wrong, but some models are useful.

George Box

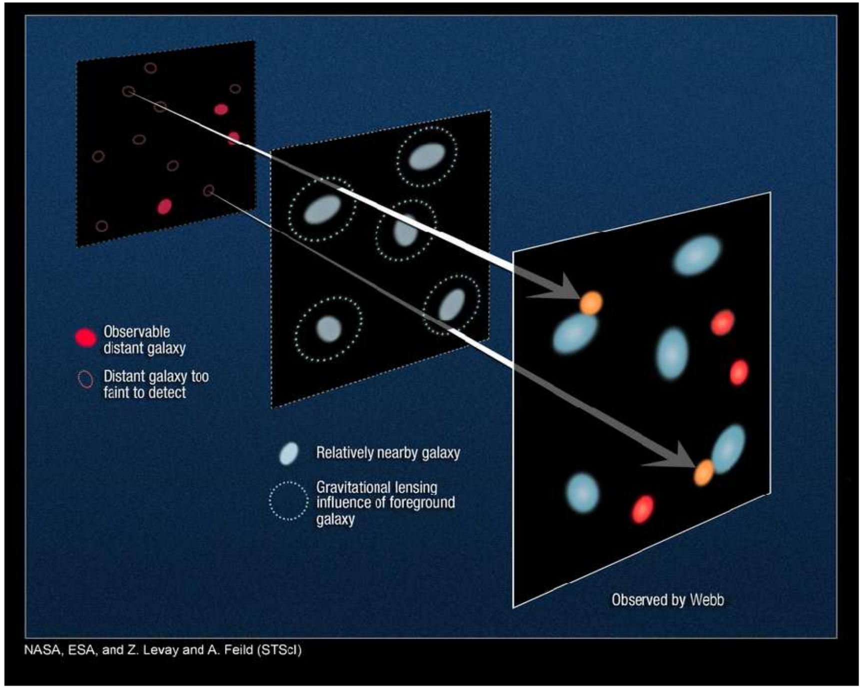


Nature is Brutal



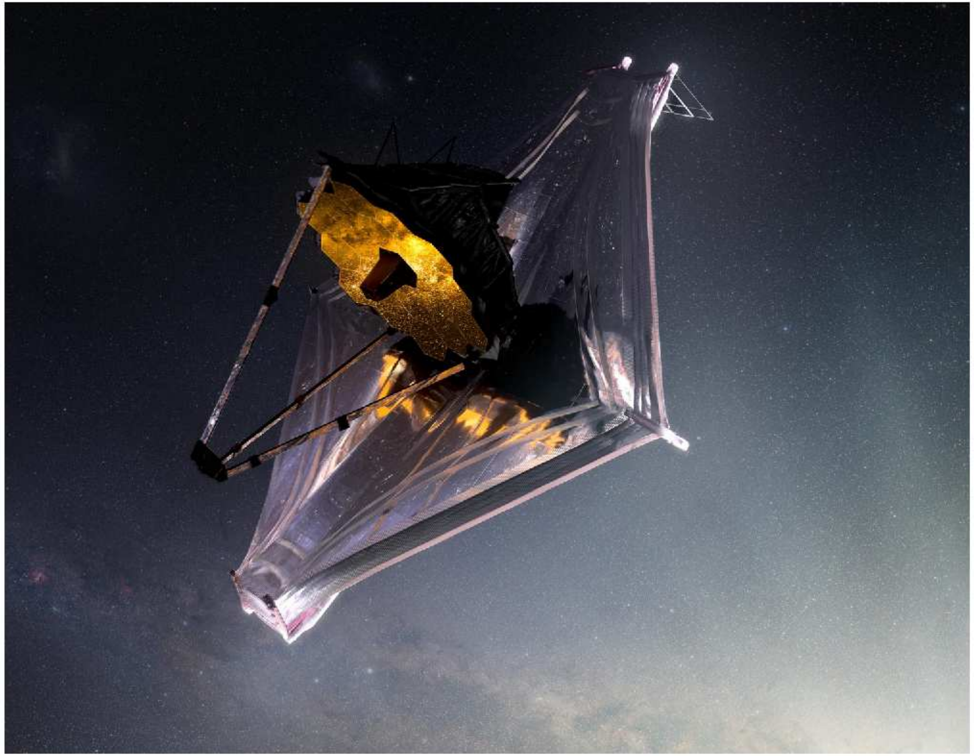
Light bends in gravity!

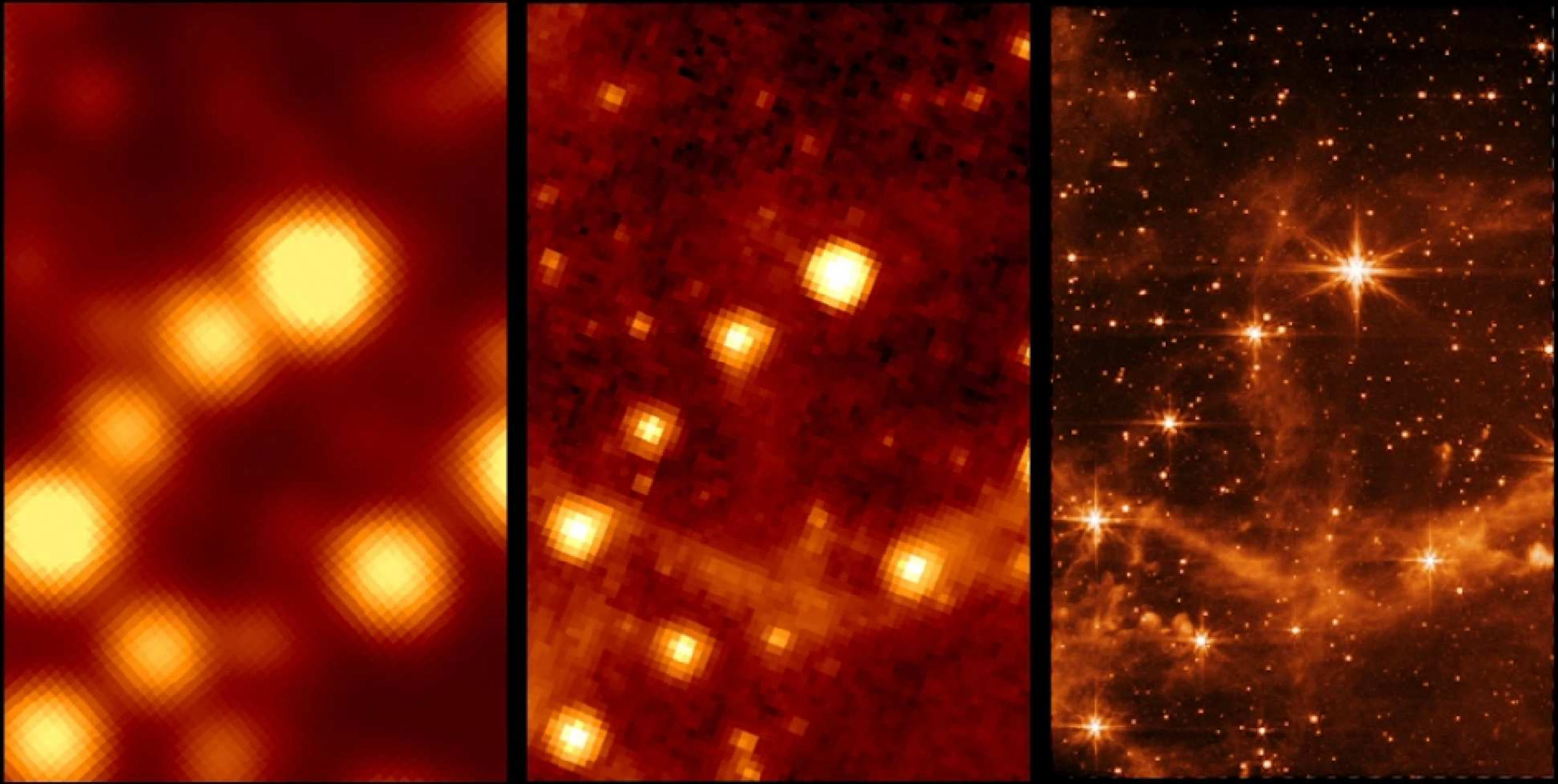
But we are cunning!



- Lensing? Like a camera?
- Use physics to form an enormous camera!
- Send a telescope to take nice pictures!

JWST



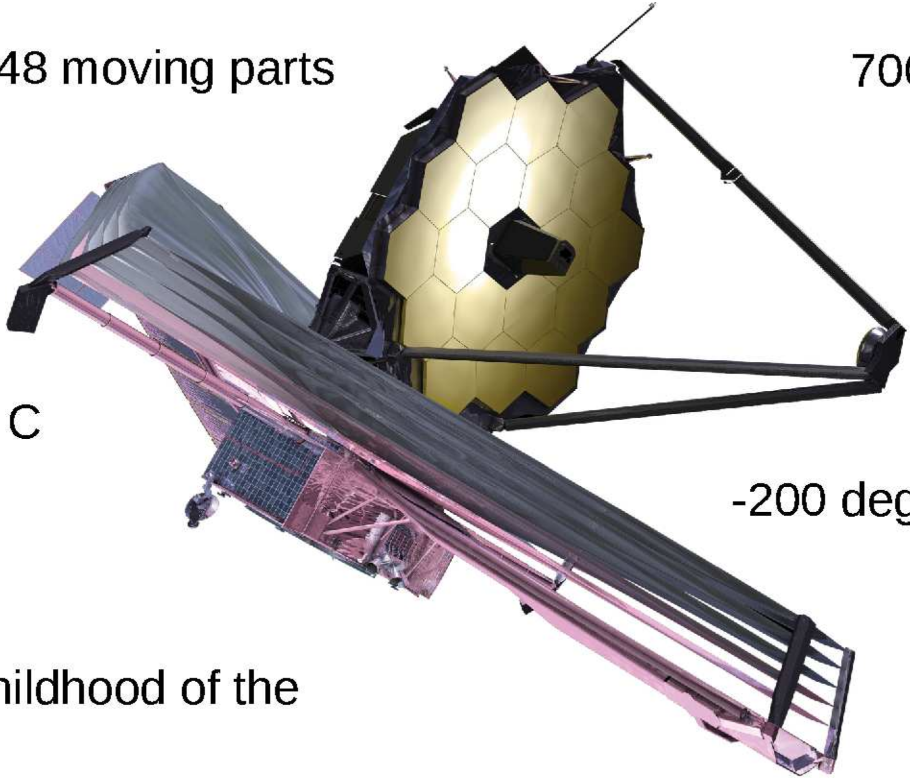


WISE W2 4.6 μm   Spitzer/IRAC 8.6 μm   JWST/MIRI 7.7 μm

A space telescope? So what?

548 moving parts

7000 parts

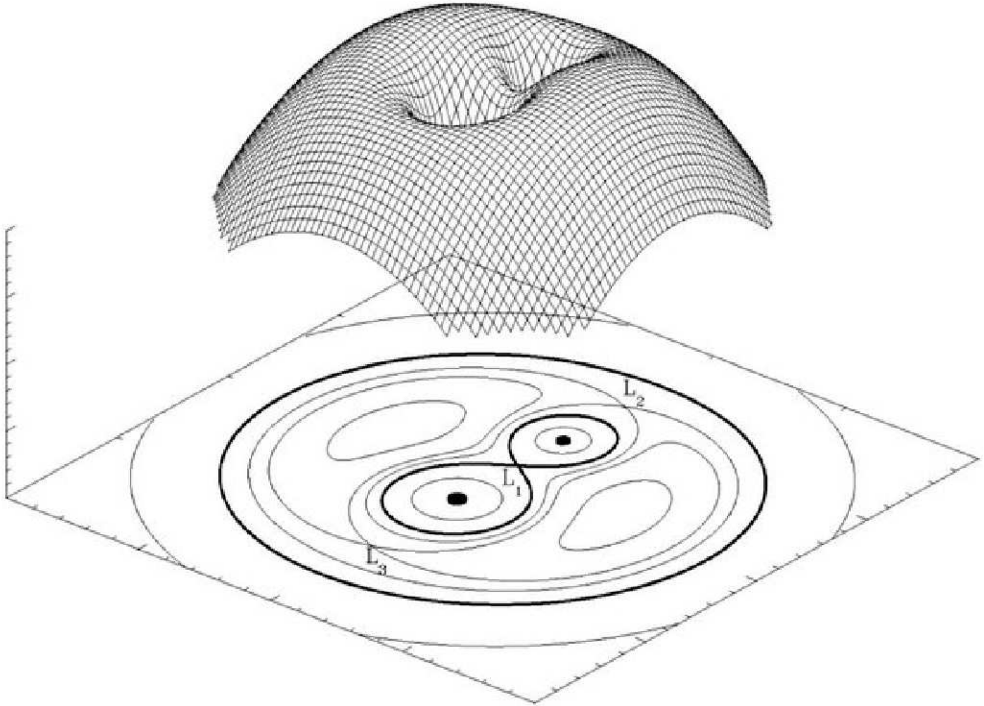


200 deg C

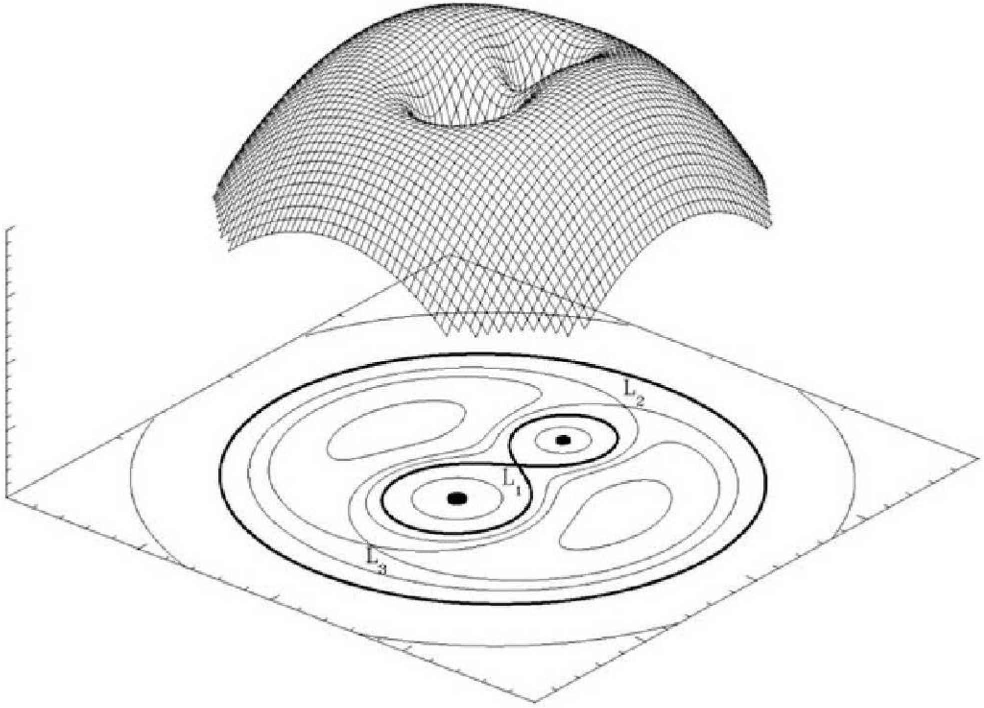
-200 deg C

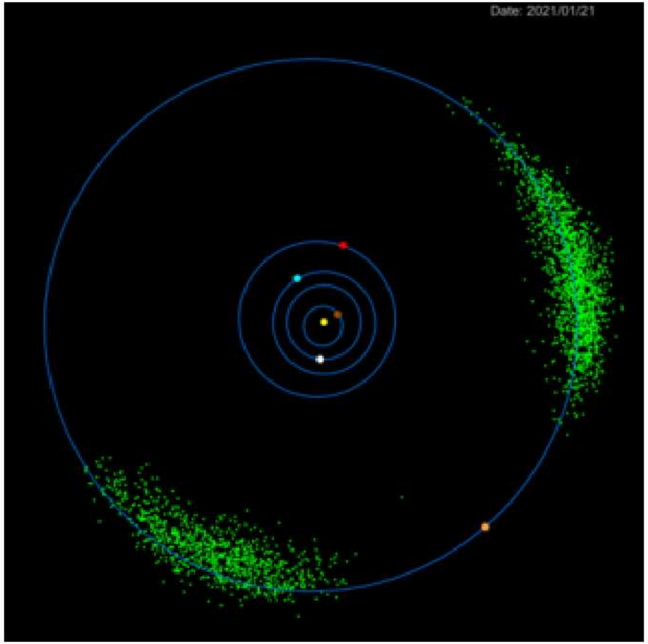
A lens to the childhood of the universe…

Exploiting Gravity as a Space Engineer



Exploiting Gravity as a Space Engineer





Nature beat us to it!

Nature

Luckily for us: we live in a continuous world.

Things are smooth, differentiable and explainable by

$$r: \quad \rho\left(\frac{\partial u_r}{\partial t} + u_r\frac{\partial u_r}{\partial r} + \frac{u_\phi}{r\sin(\theta)}\frac{\partial u_r}{\partial \phi} + \frac{u_\theta}{r}\frac{\partial u_r}{\partial \theta} - \frac{u_\phi^2 - u_\theta^2}{r}\right) = -\frac{\partial p}{\partial r} + \rho g_r +$$

$$\mu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_r}{\partial r}\right) + \frac{1}{r^2\sin(\theta)^2}\frac{\partial^2 u_r}{\partial \phi^2} + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial u_r}{\partial \theta}\right) - 2\frac{u_r + \frac{\partial u_\theta}{\partial \theta} + u_\theta\cot(\theta)}{r^2} - \frac{2}{r^2\sin(\theta)}\frac{\partial u_\phi}{\partial \phi}\right]$$

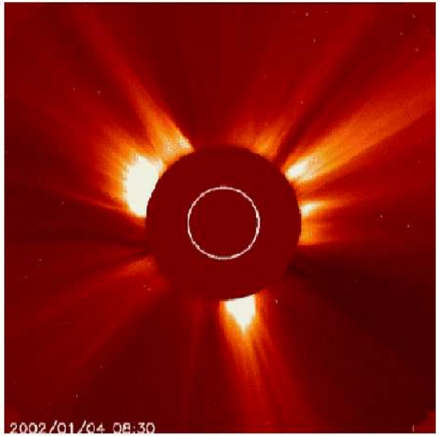
$$\phi: \quad \rho\left(\frac{\partial u_\phi}{\partial t} + u_r\frac{\partial u_\phi}{\partial r} + \frac{u_\phi}{r\sin(\theta)}\frac{\partial u_\phi}{\partial \phi} + \frac{u_\theta}{r}\frac{\partial u_\phi}{\partial \theta} + \frac{u_r u_\phi + u_\phi u_\theta\cot(\theta)}{r}\right) = -\frac{1}{r\sin(\theta)}\frac{\partial p}{\partial \phi} + \rho g_\phi +$$

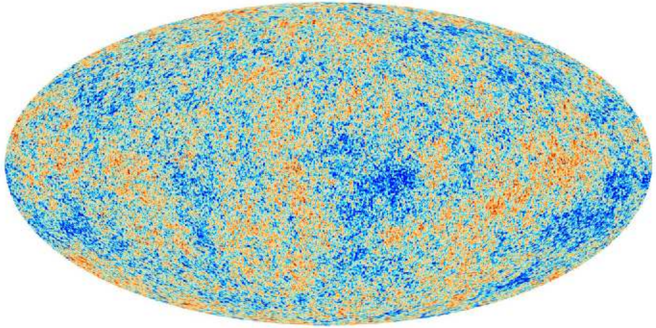
$$\mu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\phi}{\partial r}\right) + \frac{1}{r^2\sin(\theta)^2}\frac{\partial^2 u_\phi}{\partial \phi^2} + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial u_\phi}{\partial \theta}\right) + \frac{2\sin(\theta)\frac{\partial u_r}{\partial \phi} + 2\cos(\theta)\frac{\partial u_\theta}{\partial \phi} - u_\phi}{r^2\sin(\theta)^2}\right]$$

$$\theta: \quad \rho\left(\frac{\partial u_\theta}{\partial t} + u_r\frac{\partial u_\theta}{\partial r} + \frac{u_\phi}{r\sin(\theta)}\frac{\partial u_\theta}{\partial \phi} + \frac{u_\theta}{r}\frac{\partial u_\theta}{\partial \theta} + \frac{u_r u_\theta - u_\phi^2\cot(\theta)}{r}\right) = -\frac{1}{r}\frac{\partial p}{\partial \theta} + \rho g_\theta +$$

$$\mu\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u_\theta}{\partial r}\right) - \frac{1}{r^2\sin(\theta)^2}\frac{\partial^2 u_\theta}{\partial \phi^2} + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial u_\theta}{\partial \theta}\right) - \frac{2}{r^2}\frac{\partial u_r}{\partial \theta} - \frac{u_\theta + 2\cos(\theta)\frac{\partial u_\phi}{\partial \phi}}{r^2\sin(\theta)^2}\right].$$

Nature is smooth and continuous!





**Soun
d**

**Light
+
Fluids**

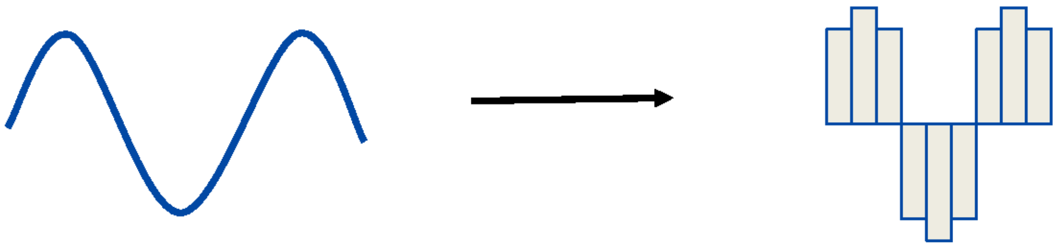
**Even
the
CMBR**

Everything is smooth, but not everything is solveable…

Everything is smooth, but not everything is solveable…

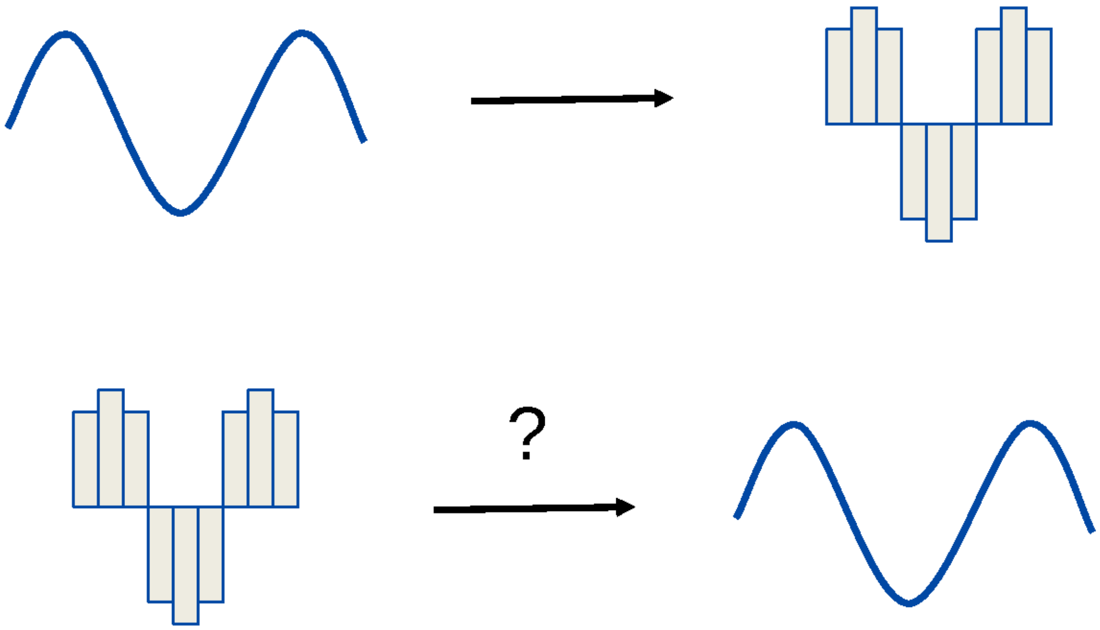
Do what everyone else does!
Throw a neural network at it!

Implicit Neural Representations

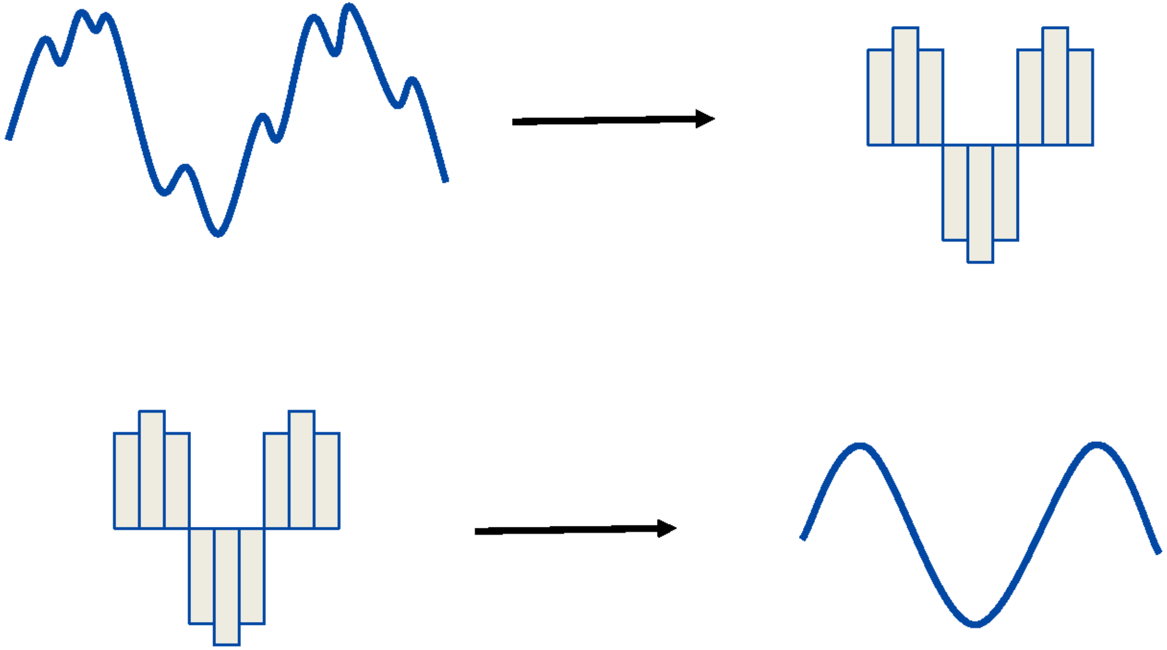


Measure Continuous Functions as Discrete Samples

Implicit Neural Representations



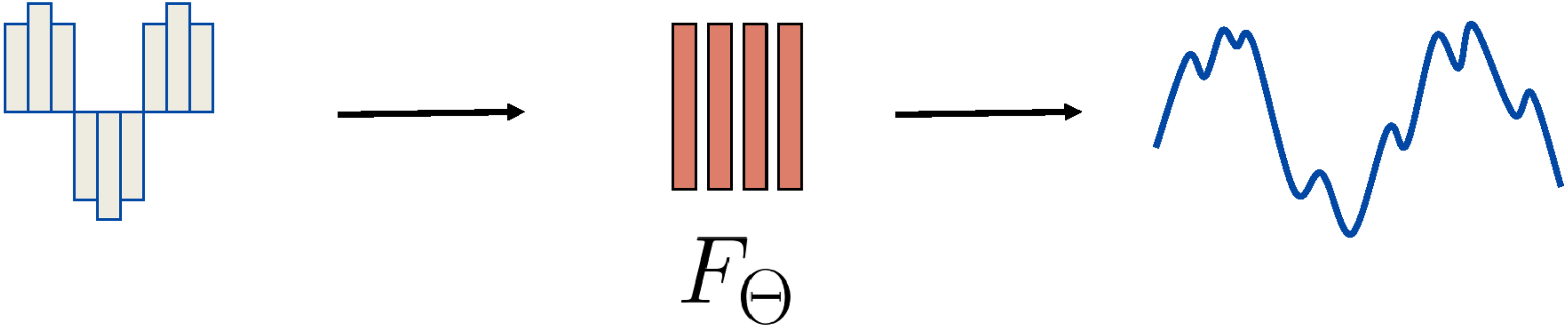
Cannot Always Reconstruct Difficult Continuous Functions from Discrete



No!

Not Unique if Undersamp led!

Implicit Neural Representations

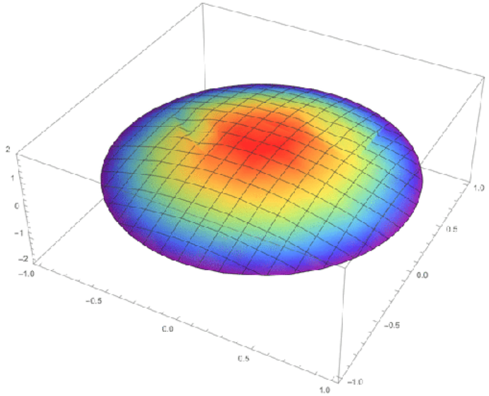


Learn an implicit function
to approximate the
continuous signal

Common Discretised Signals

Pixels are a discrete space

Video has pixels and a framerate (temporally discrete)



Meshes, pointclouds and PDE's all have discrete domains

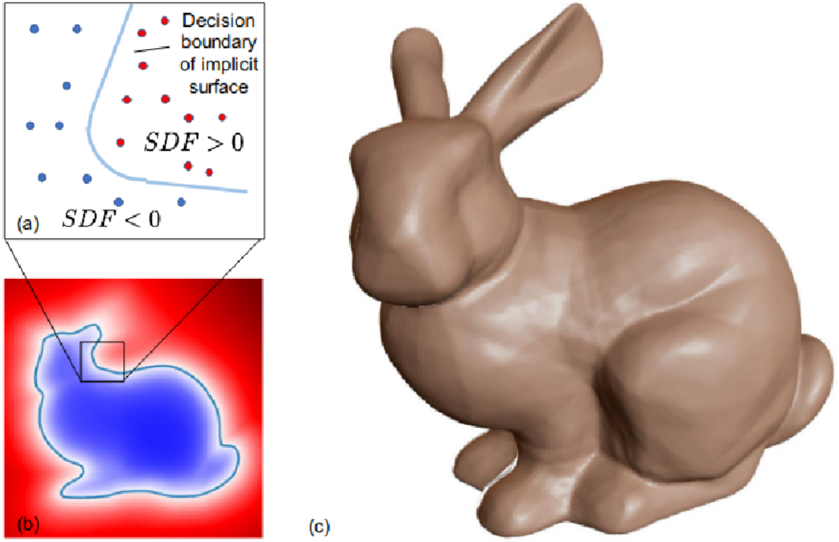
NFSA, NASA and Purdue

An example: DeepSDF

- A simple case: learn where a surface is.

- Discretise 3D space, sample points and say whether inside, or outside the bunny.

- Learn a continuous, smooth surface which separates physical regions.



Park et al. 2019

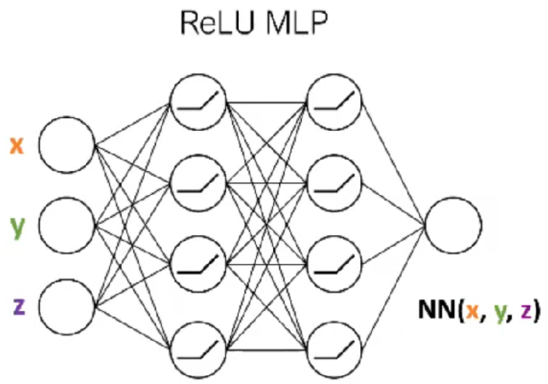
What sort of network do we need?

$$F(\mathbf{x}, \Phi, \nabla_{\mathbf{x}}\Phi, \nabla^2_{\mathbf{x}}\Phi, \ldots) = 0,\ \Phi : \mathbf{x} \mapsto \Phi(\mathbf{x})$$

Approximate *some* function

Using *some* nonlinear activation function

An MLP works as a function approximator, and by Cybenko's theorem: there exists an MLP of sufficient dimension which can approximate our function well enough.



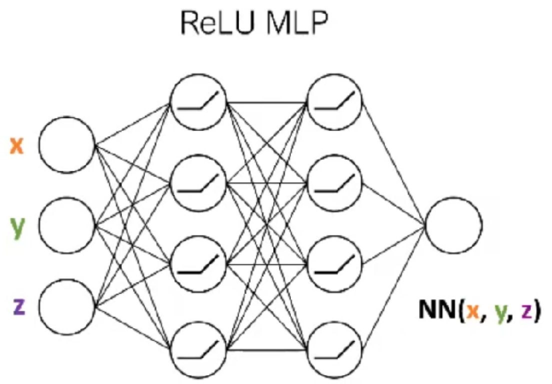
ReLU? Step? Leaky ReLU?

What sort of network do we need?

$$F(\mathbf{x}, \Phi, \nabla_{\mathbf{x}}\Phi, \nabla^2_{\mathbf{x}}\Phi, \ldots) = 0, \; \Phi : \mathbf{x} \mapsto \Phi(\mathbf{x})$$

Approximate *some* function

An MLP works as a function approximator, and by Cybenko's theorem: there exists an MLP of sufficient dimension which can approximate our function well enough.



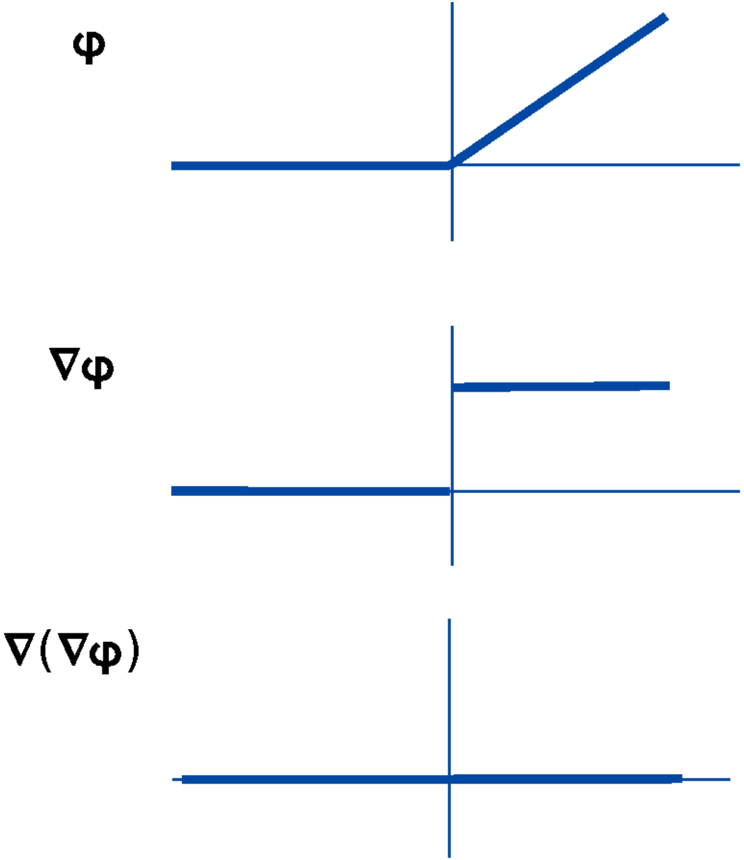
Using *some* nonlinear activation function

ReLU? Step? Leaky ReLU?

φ and ∇φ and ∇(∇φ) and …

Differentiability of continuous functions is key!

A ReLU's 2nd derivative is 0 - similar for many other nonlinear activation functions!

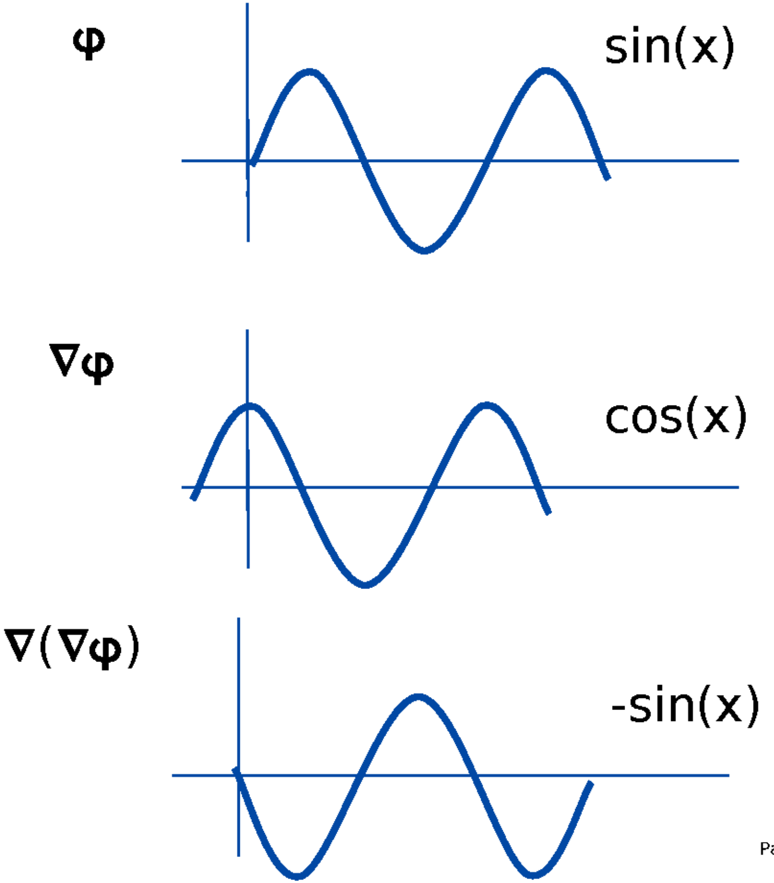


φ and ∇φ and ∇(∇φ) and ...

Differentiability of continuous functions is key!

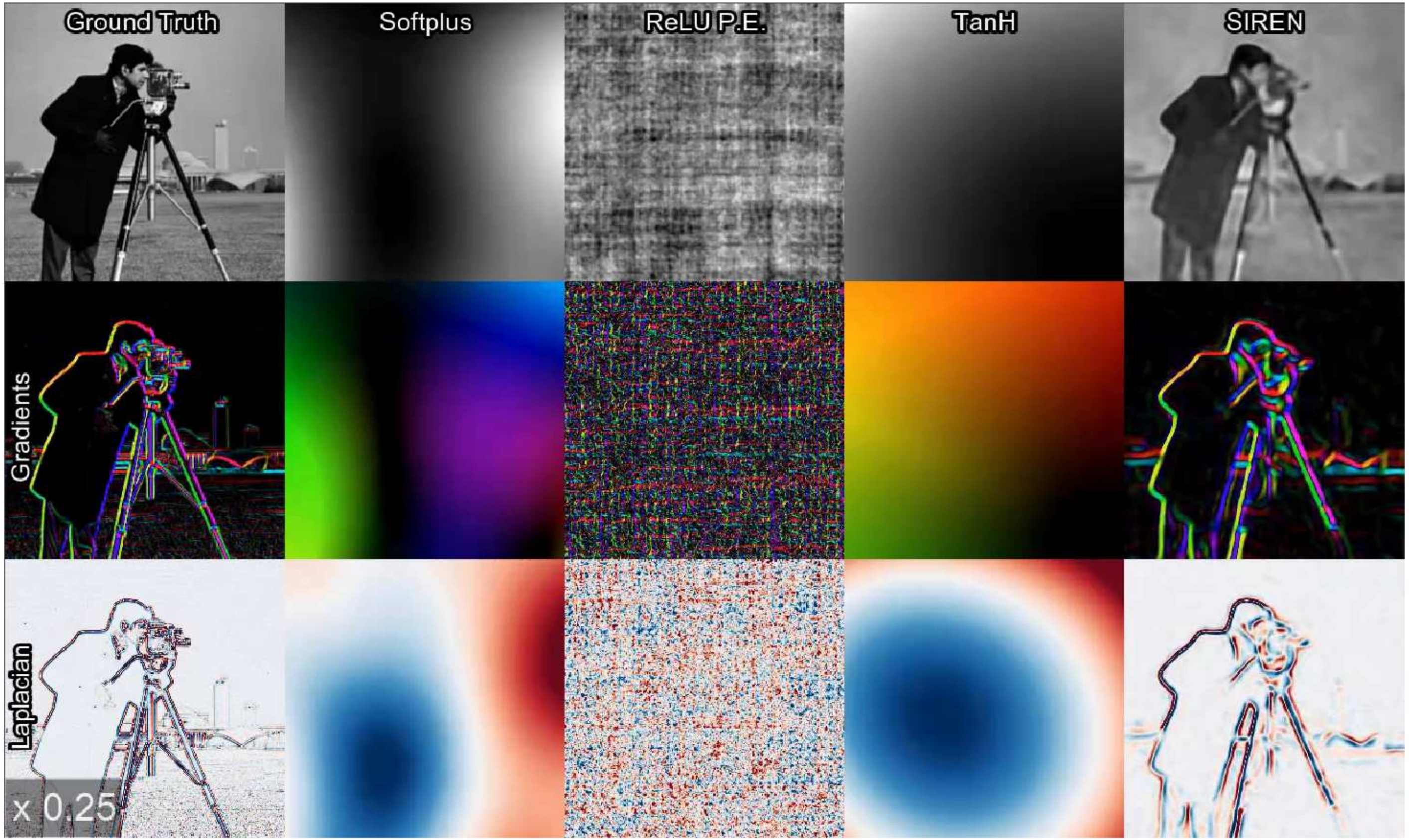
Sine functions are continuously differentiable!

We can model information of higher orders! Higher frequencies!



SIREN

Sitzmann et al. 2020



Sitzmann et al. 2020

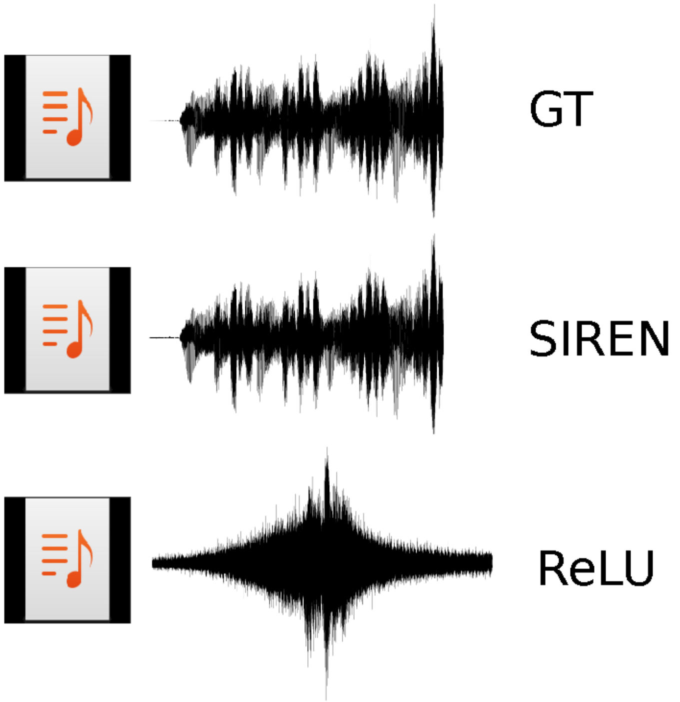
Less pictures, more physics!

Why does SIREN work?

- Underlying smoothness to derivatives

- Derivative of a SIREN is a SIREN i.e. decision making with derivatives.

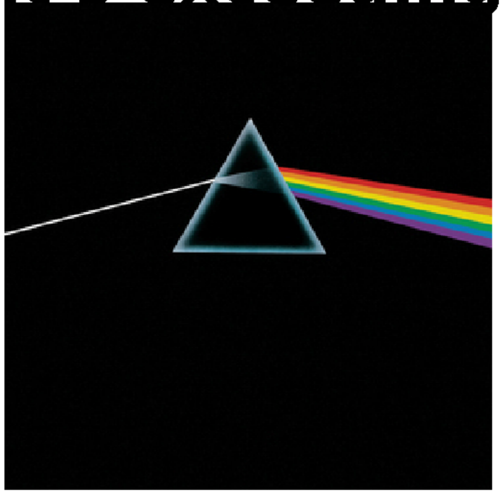
- Pseudo-Fourier decomposition



Let's go back to light… that's smooth right?

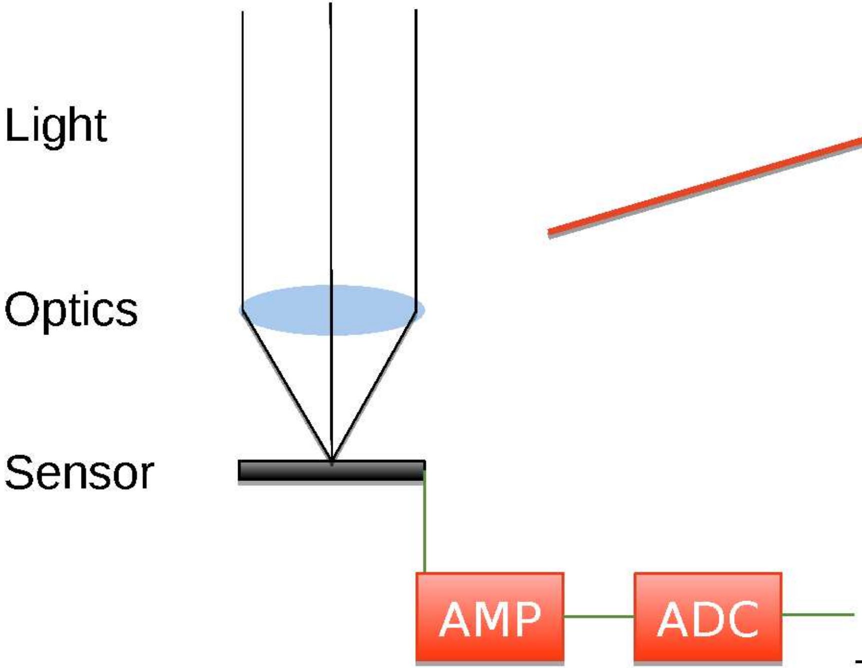
– You're right, it's smooth.

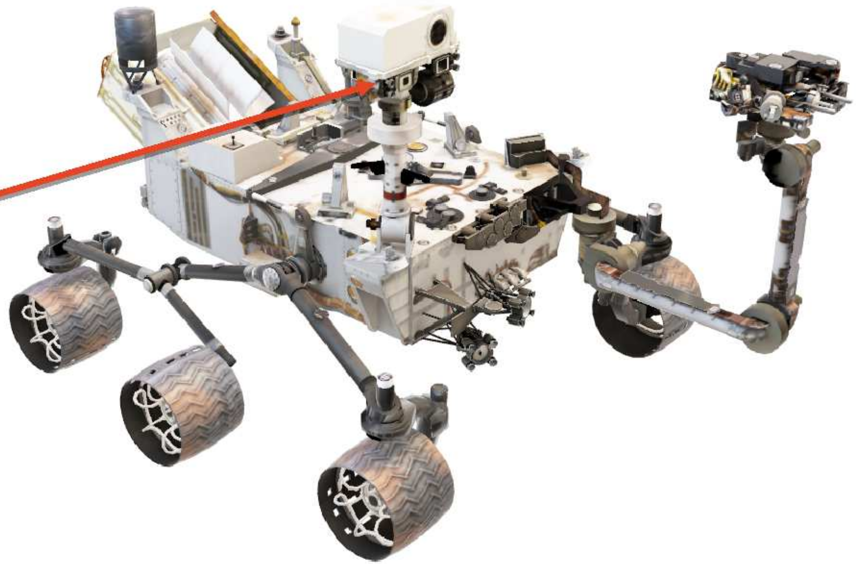
– But it's exceedingly com



$$\nabla \cdot \mathbf{D} = \rho$$
$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

Cameras

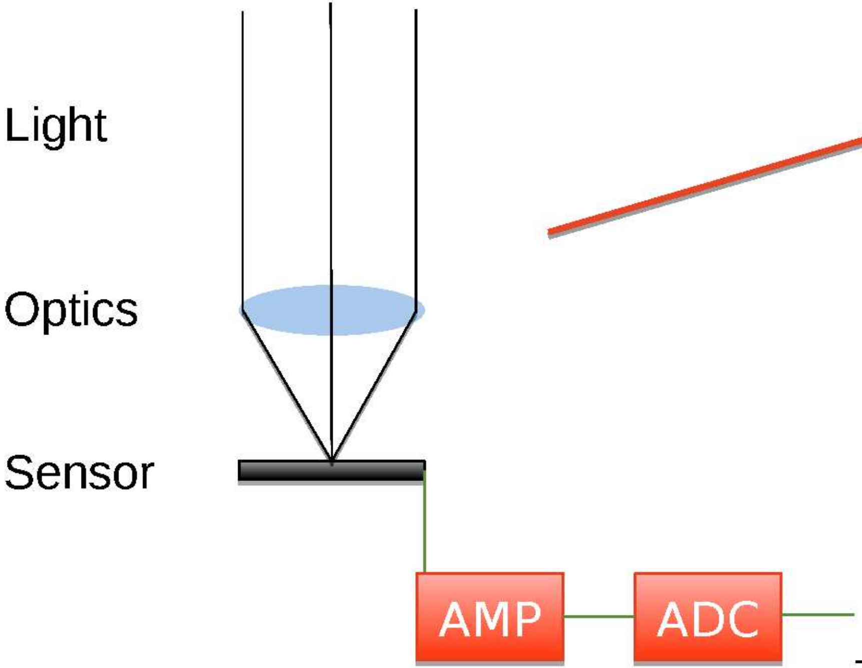


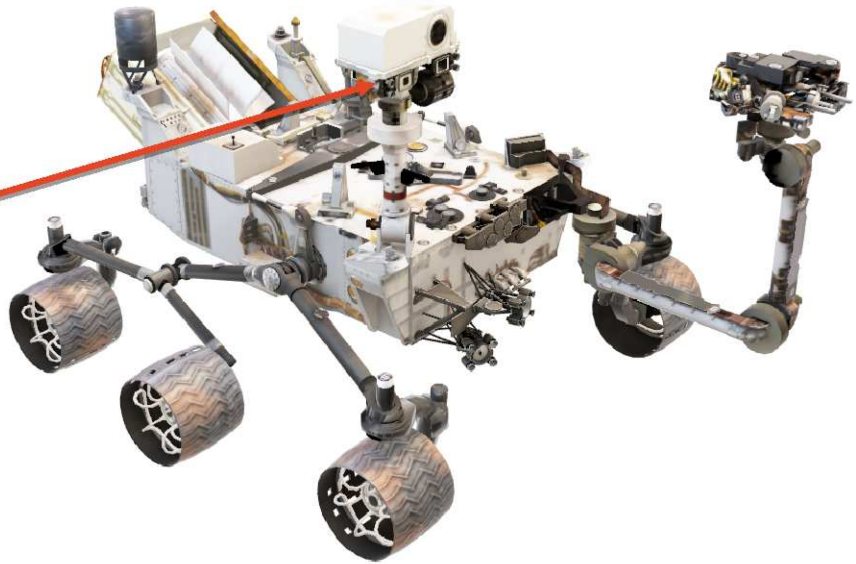


75	20	75	20
25	92	25	92
72	22	72	22

Image
(Bayer Filter - Colour)

Cameras

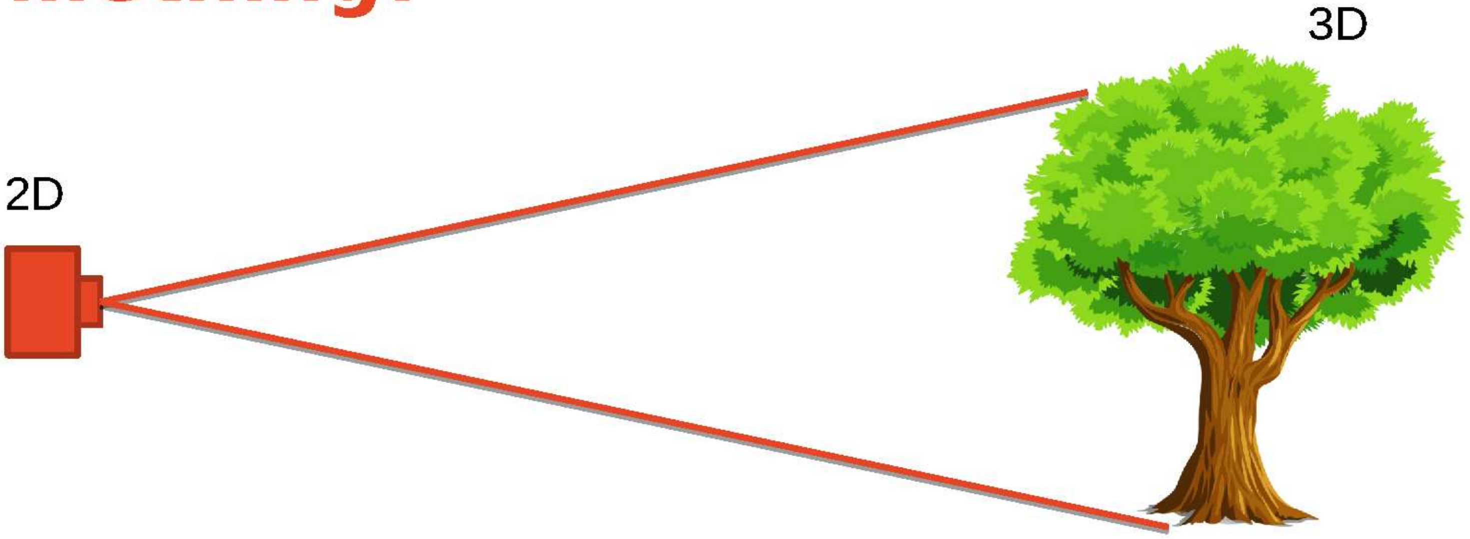




80	76	84	71
83	75	88	92
72	92	72	70

Image (Mono)

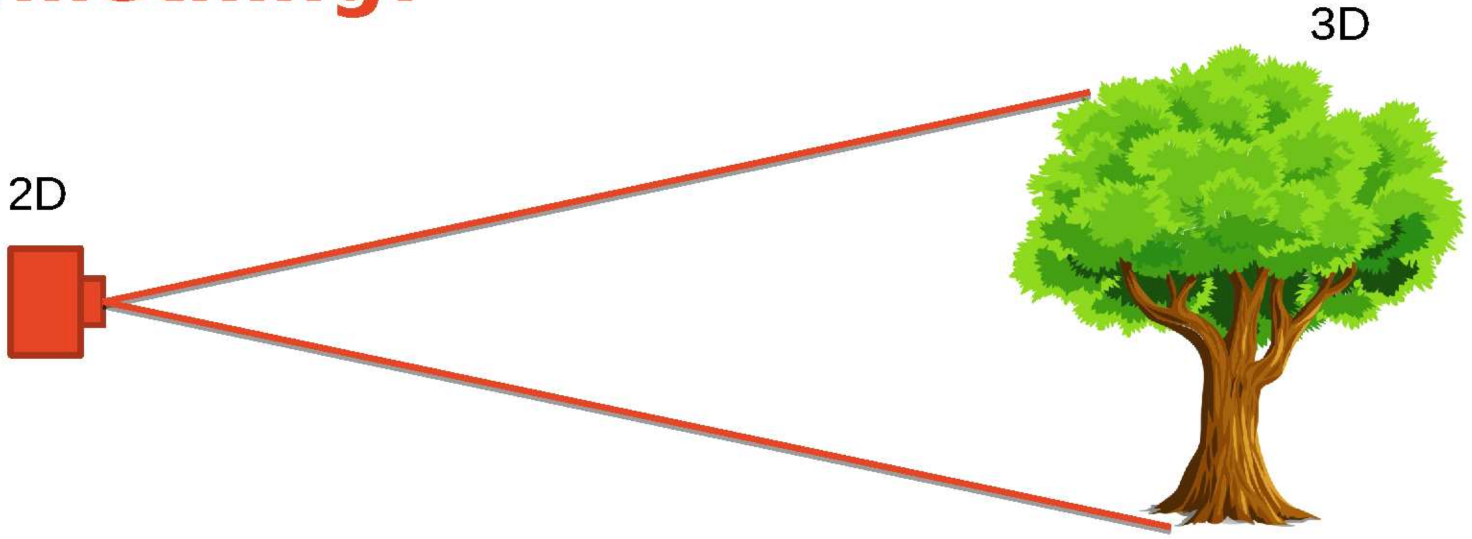
What happens when we image something?



$$P(\theta,\phi,\lambda,t,p,V_x,V_y,V_z)$$

How many "D" is light?

What happens when we image something?



$$P(\theta, \phi, \lambda, t, p, V_x, V_y, V_z)$$

How many “D” is light?

8D!

What is the most we can ask the camera?

$$P\left(\theta,\phi,\lambda,t,p,V_x,V_y,V_z\right)$$

Angles Passing Through Aperture

Wavelength of Light

Time

Polarisation

Aperture Pose

What is the most we can ask the camera?

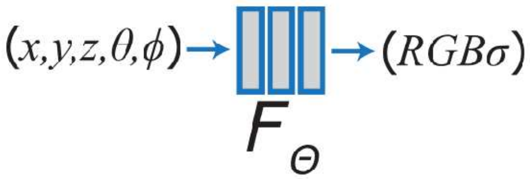
$$P(\theta, \phi, \lambda_{vis}, t, p, V_x, V_y, V_z)$$

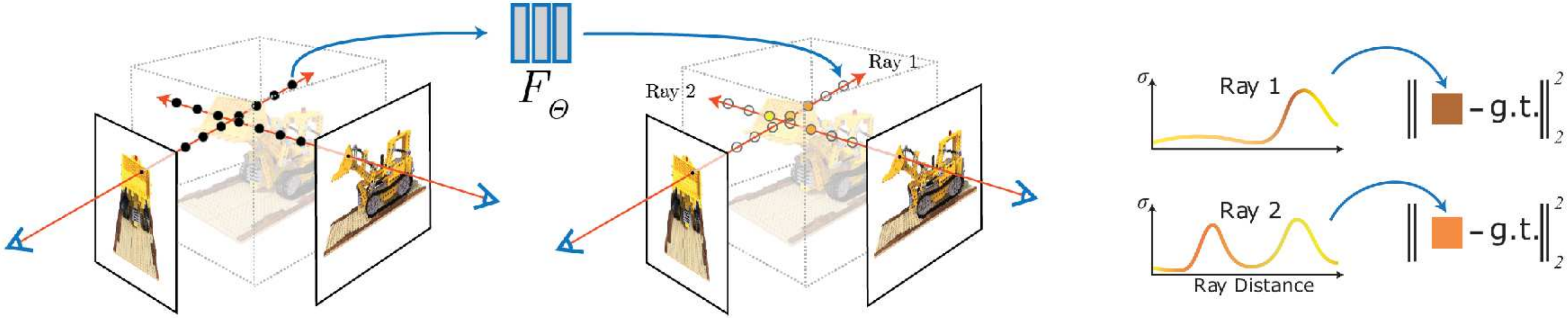
Regular Camera
(And t just implies videos)

Other cameras can give different, and sometimes greater range of the plenoptic function.

"It's NeRF or nothing."

- Learn a 5D representation of light in a NN.
- Spatially varying density, spatially/view varying colour.
- Novel view synthesis.

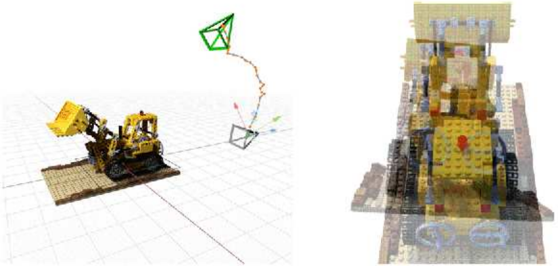




Mildenhall et al. 2020

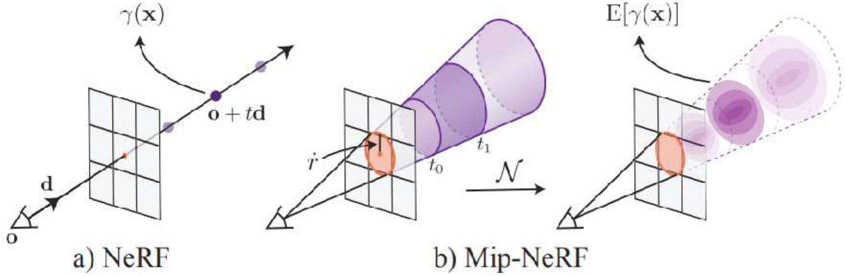
NeRF Variants

The inverse problem!



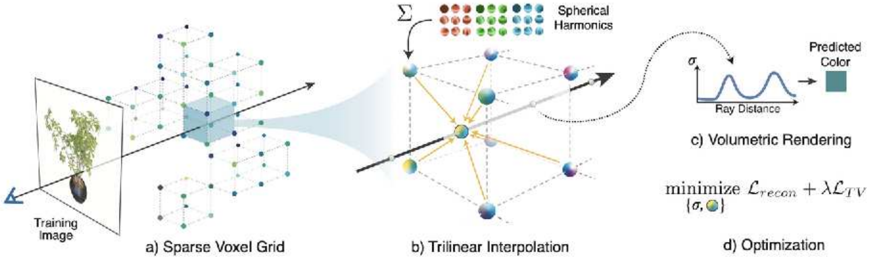
Yen-Chen et al. 2021

Pixels aren't rays!



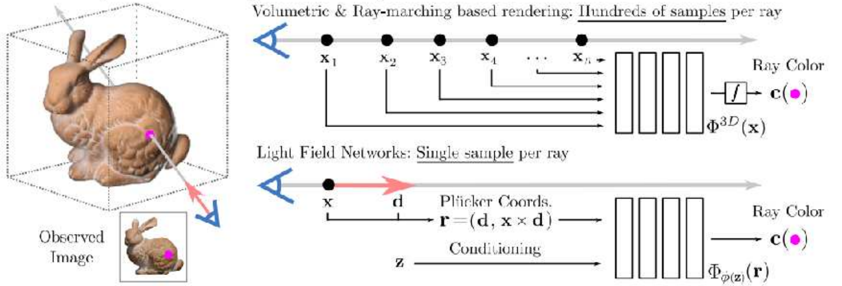
Barron et al. 2020

Get rid of the NN!



Yu et al. 2022

New way of thinking about rays!

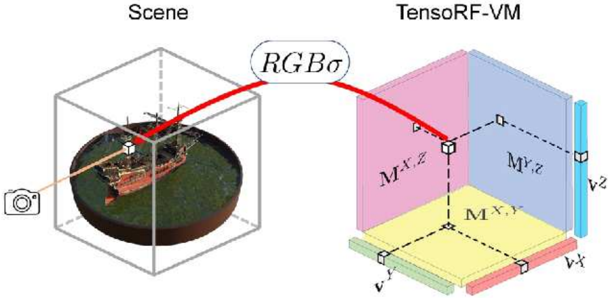


Sitzmann et al. 2021

+ many other variants!

+ many other variants!

My personal favourite: consider NeRFs as a bunch of tensors and do linear algebra…



Chen et al. 2022

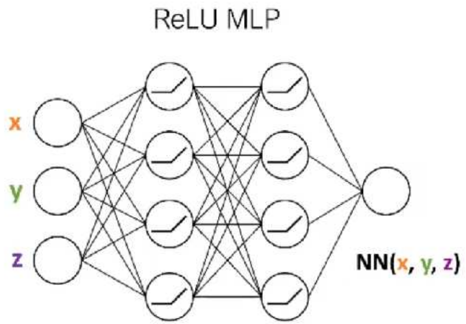
Tancik et al. 2021

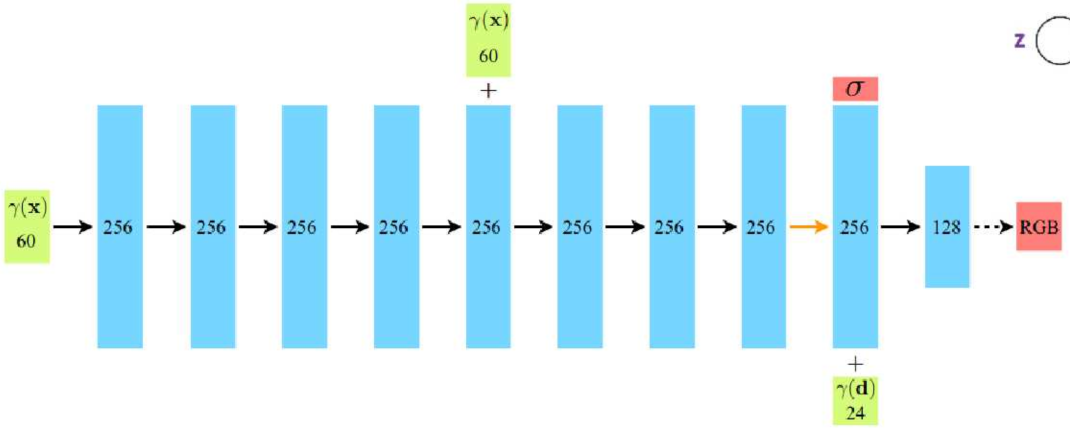
Different light fields!

Depth for free!
(We learn a volume!)

Mildenhall et al. 2020

The Key to NeRF

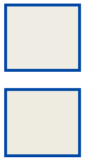


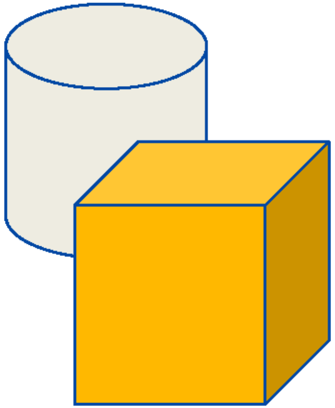


$$\gamma(\mathbf{x}) = \left(\sin(2^0\pi\mathbf{x}), \cos(2^0\pi\mathbf{x}), \ldots, \sin(2^{L-1}\pi\mathbf{x}), \cos(2^{L-1}\pi\mathbf{x})\right)$$

Mildenhall et al. 2020

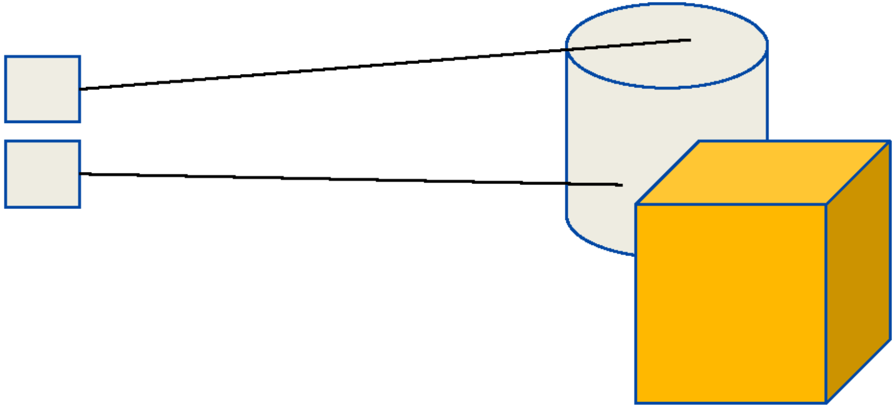
Compare the pair?





Compare the pair?

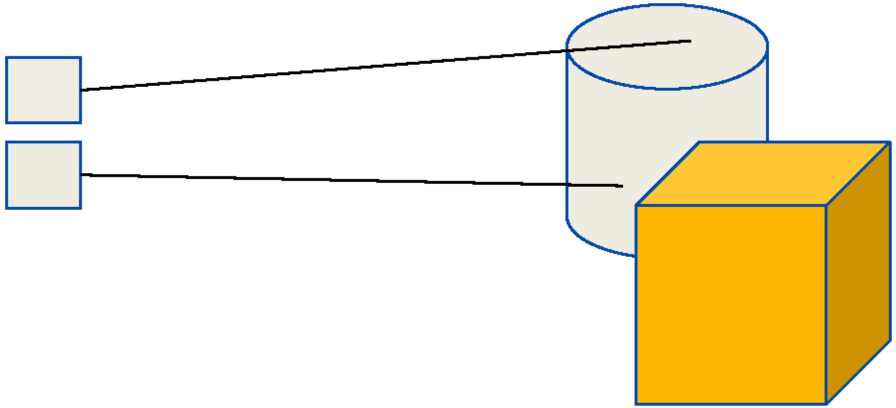
Virtually
the same
to us!



Compare the pair?

Virtually
the same
to a
compute
r!

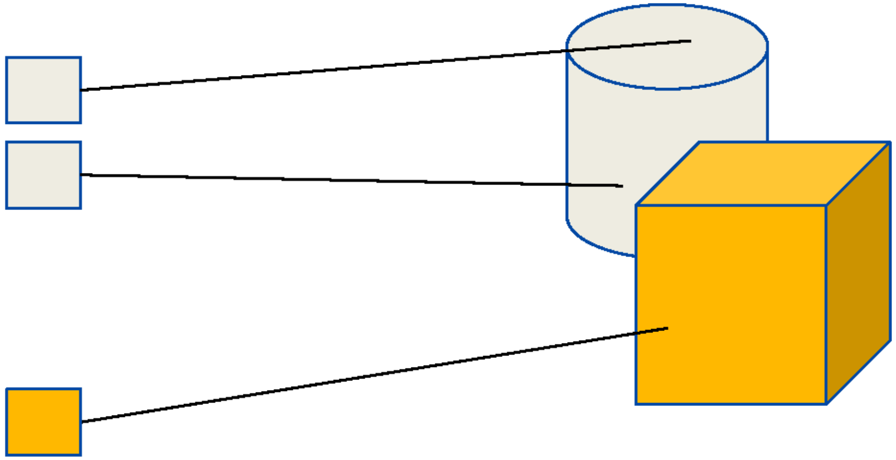
0.6
2
0.6
3



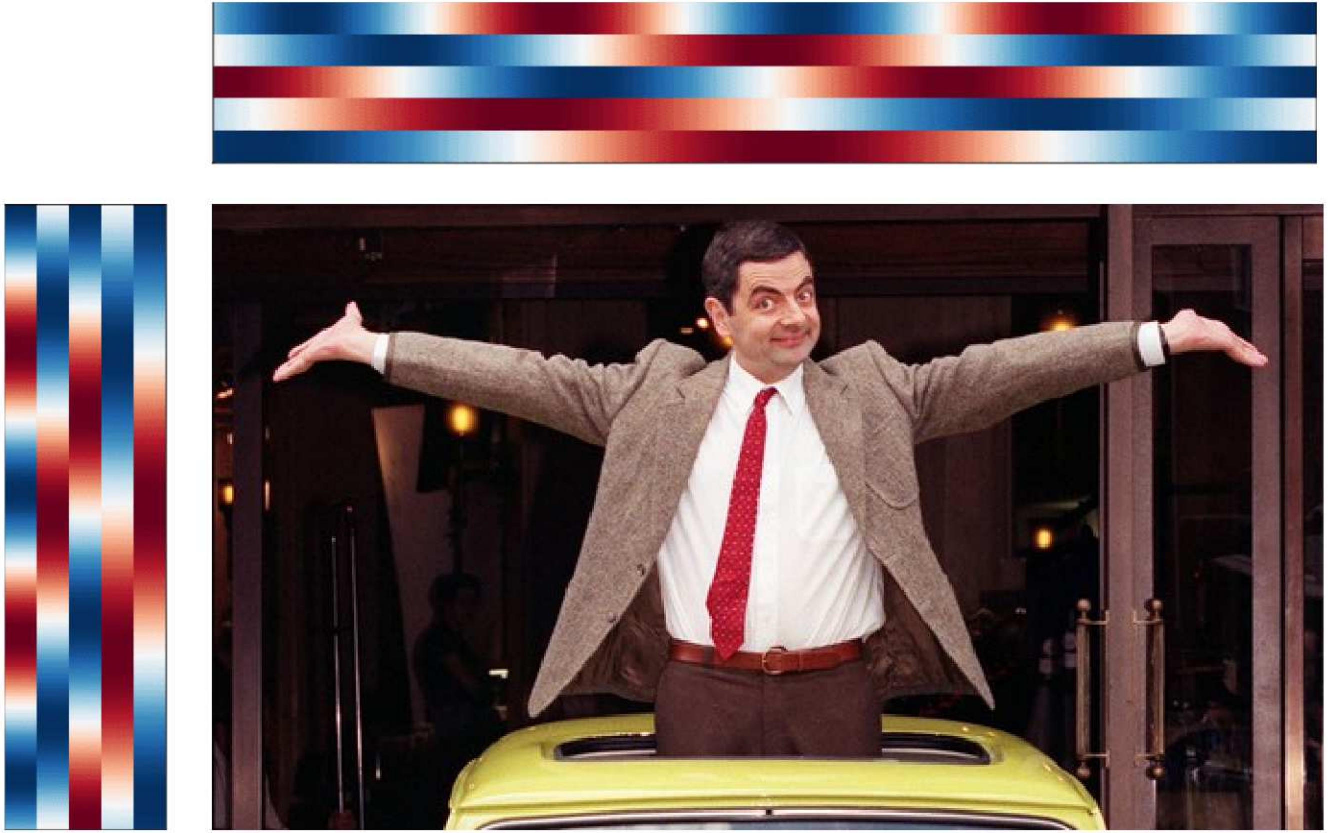
Compare the pair?

0.6
2

0.6
3

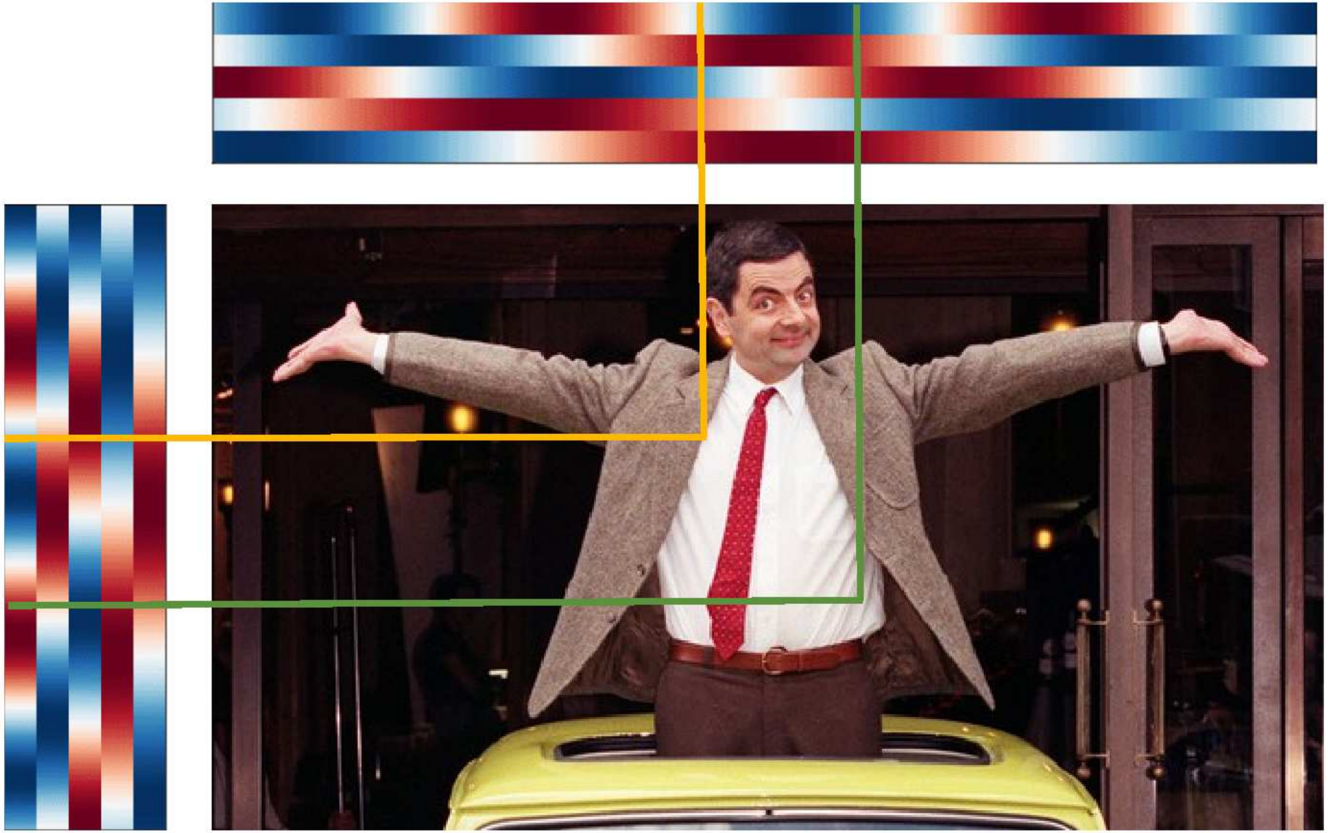


0.8
6



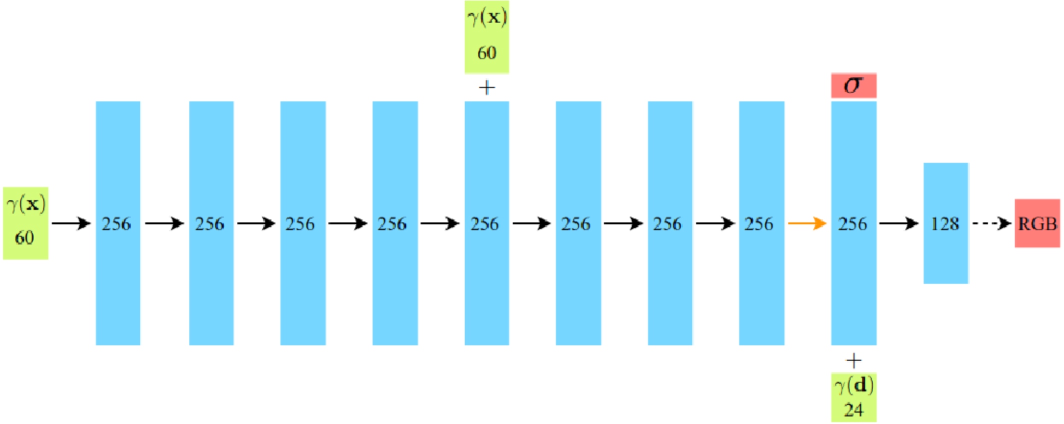
Sinusoids of different frequency

$$\gamma(\mathbf{x}) = \left(\sin(2^0\pi\mathbf{x}), \cos(2^0\pi\mathbf{x}), \ldots, \sin(2^{L-1}\pi\mathbf{x}), \cos(2^{L-1}\pi\mathbf{x})\right)$$



Physics

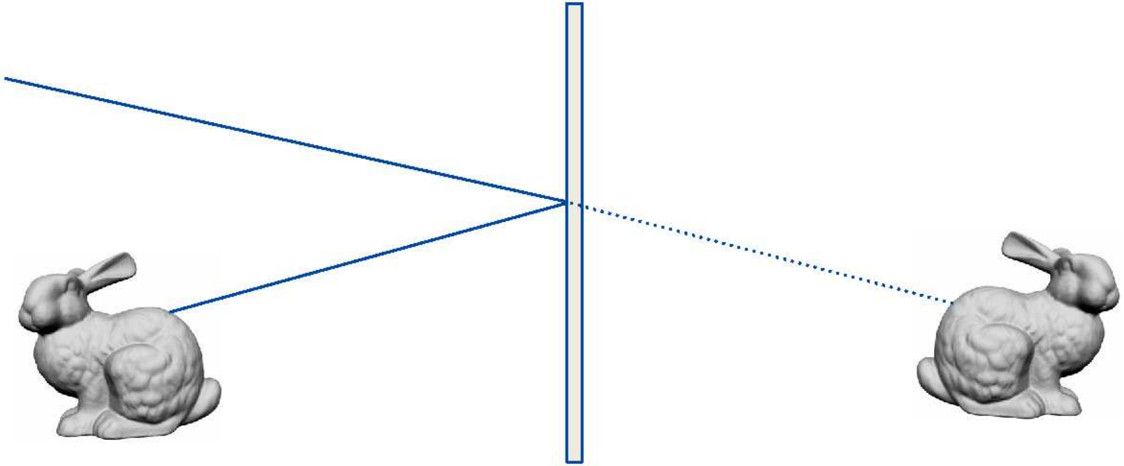
- Neural networks know nothing about the real world!
- Fortunately, we've got a few hundred years of understanding physics and nice maths!
- What is NeRF modelling?



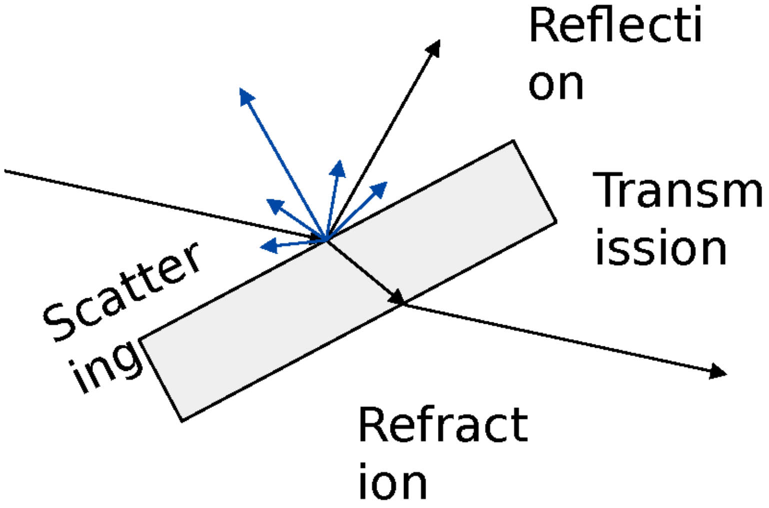
House of Mirrors

Mirrored Worlds

A NeRF is making sense of it, the best way it knows!

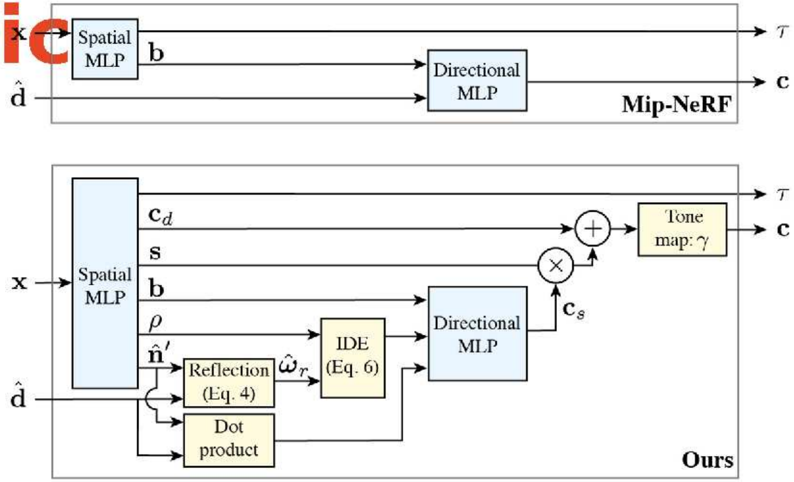


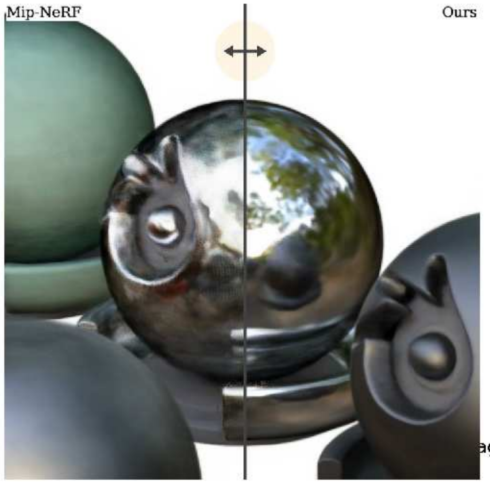
How Light Usually Works



Learning under physic

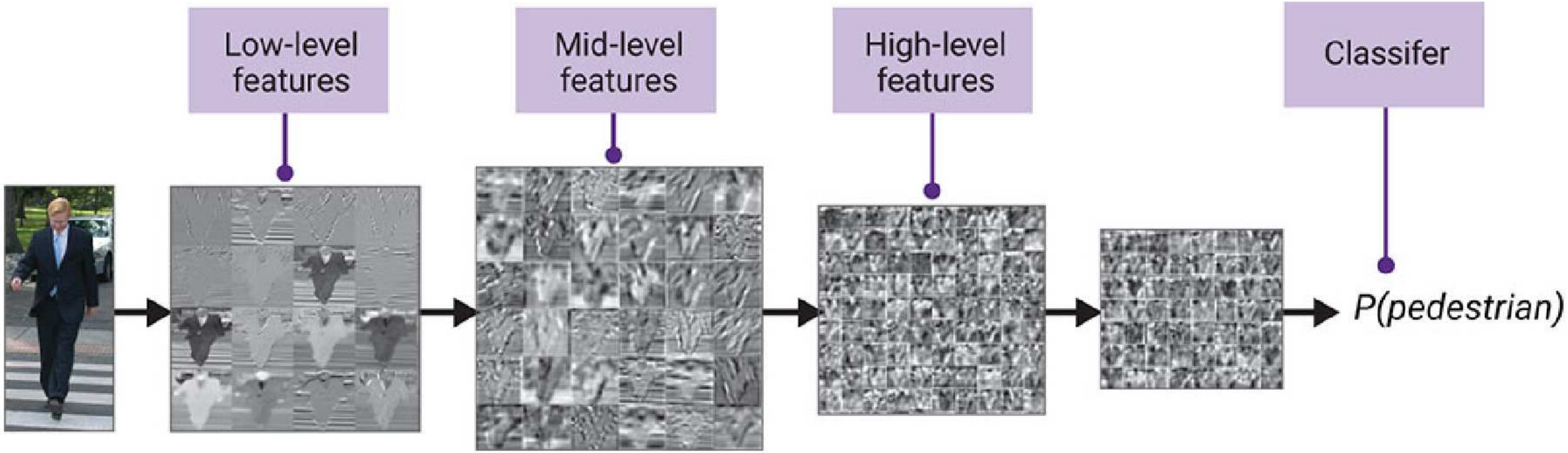
- **Give the network some physics!**
- **Reflections change ray direction**





Verbin et al. 2022

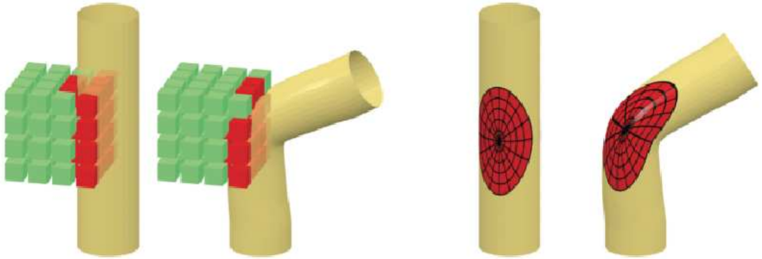
Traditional Deep Learning

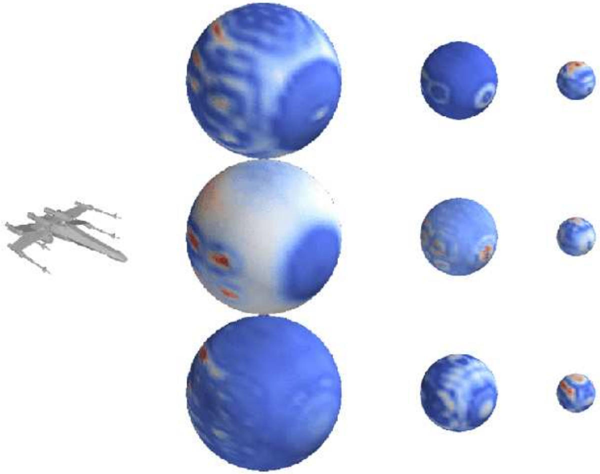


The problem?

Learning under Geometry?

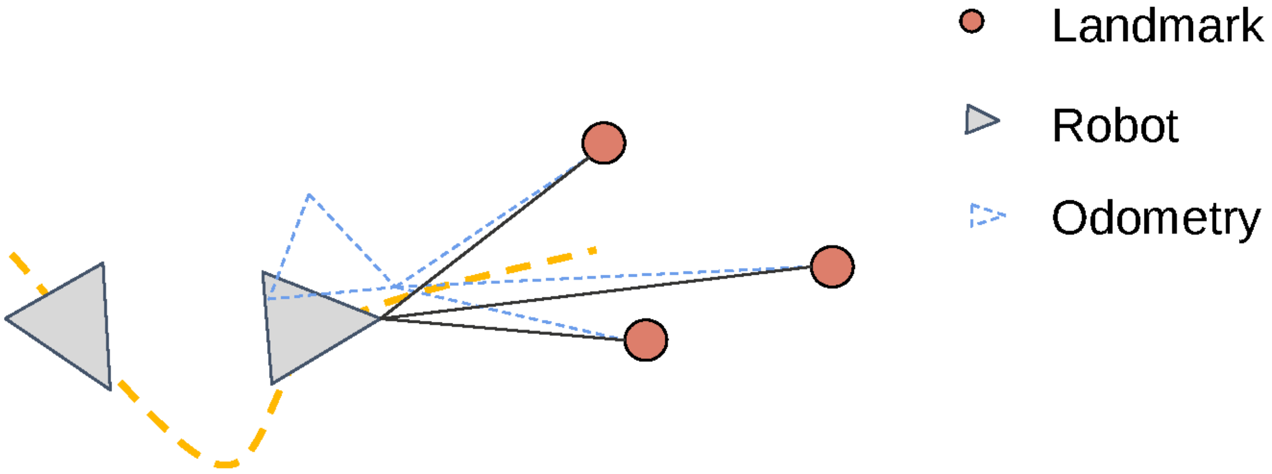
Masci et al. 2016

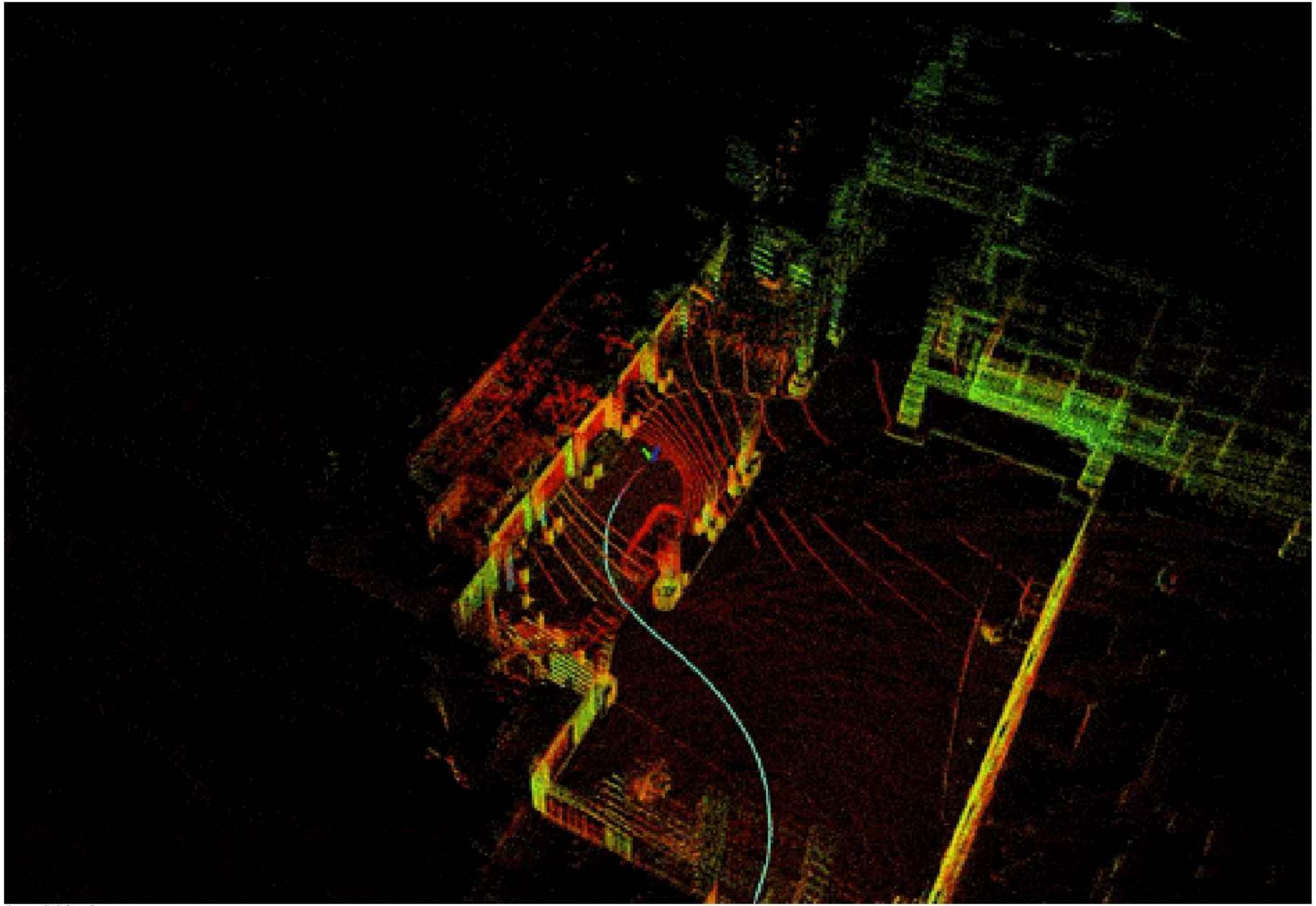




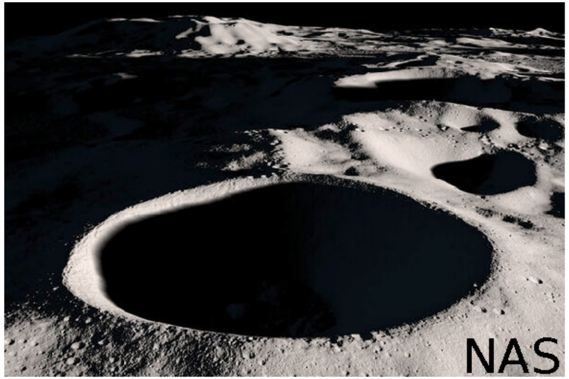
Cohen et al. 2018

SLAM



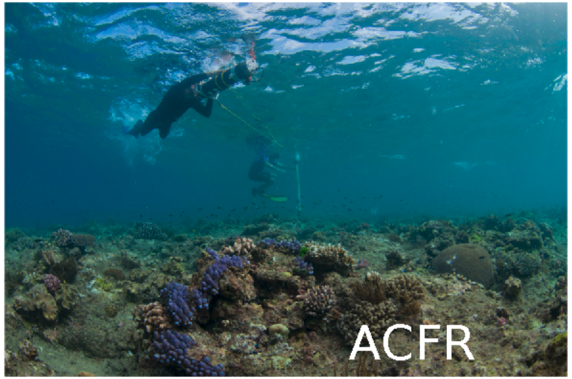


Challenges



Traditional sensors fail: e.g. high dynamic range

Also reflection, backscatter, transparency



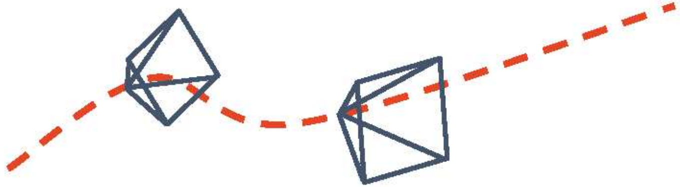
Mapping under scattering and attenuating media

Changing scene conditions

NeRFs for Robotics: Our Approach

Graphics: Fixed Poses

- Ray-based localisation framework leveraging unique spatial and visual representation in NeRF
- Incremental mapping with online updates



Robotics: Smooth, controllable capture

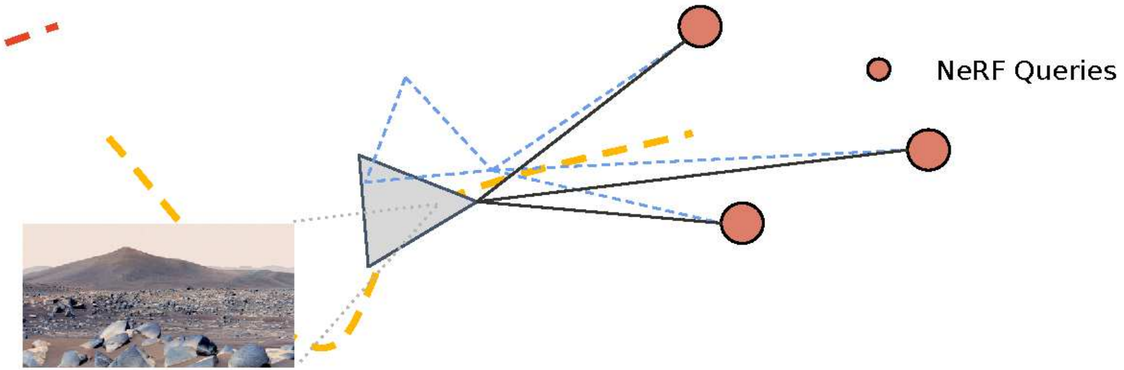
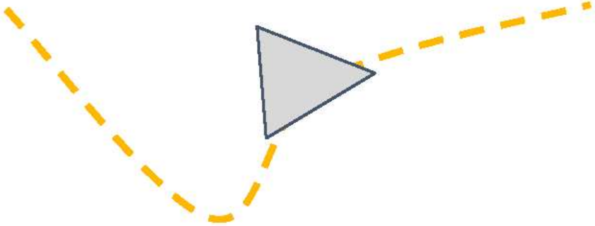
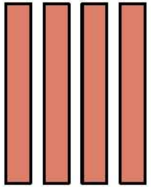


Image from Sensor

Simultaneously Derive

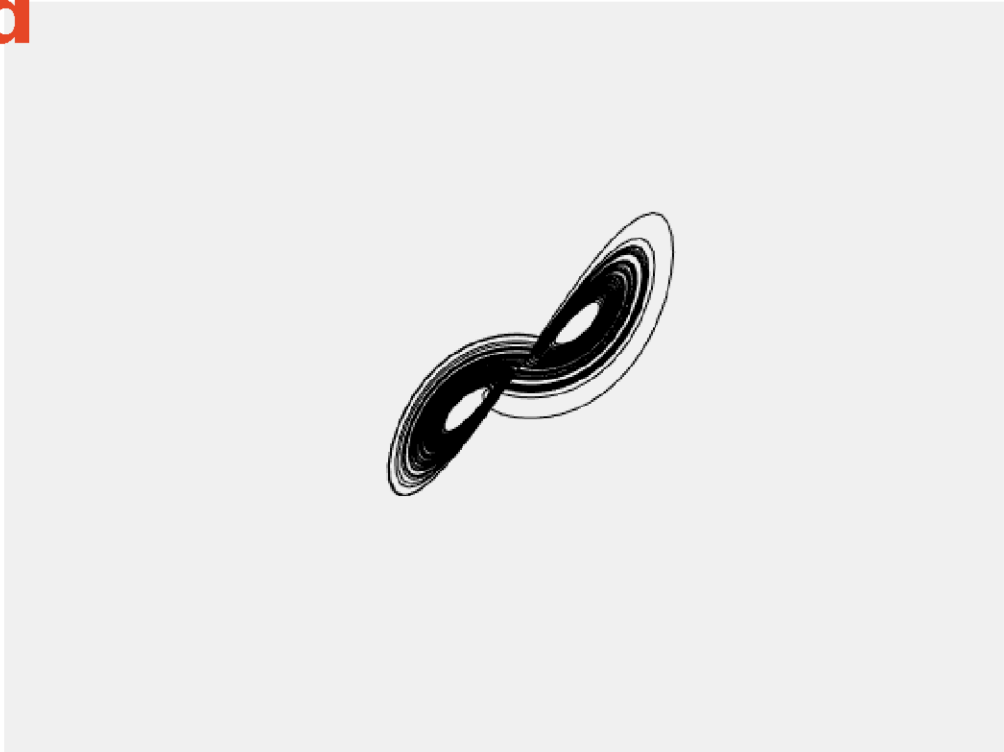


+



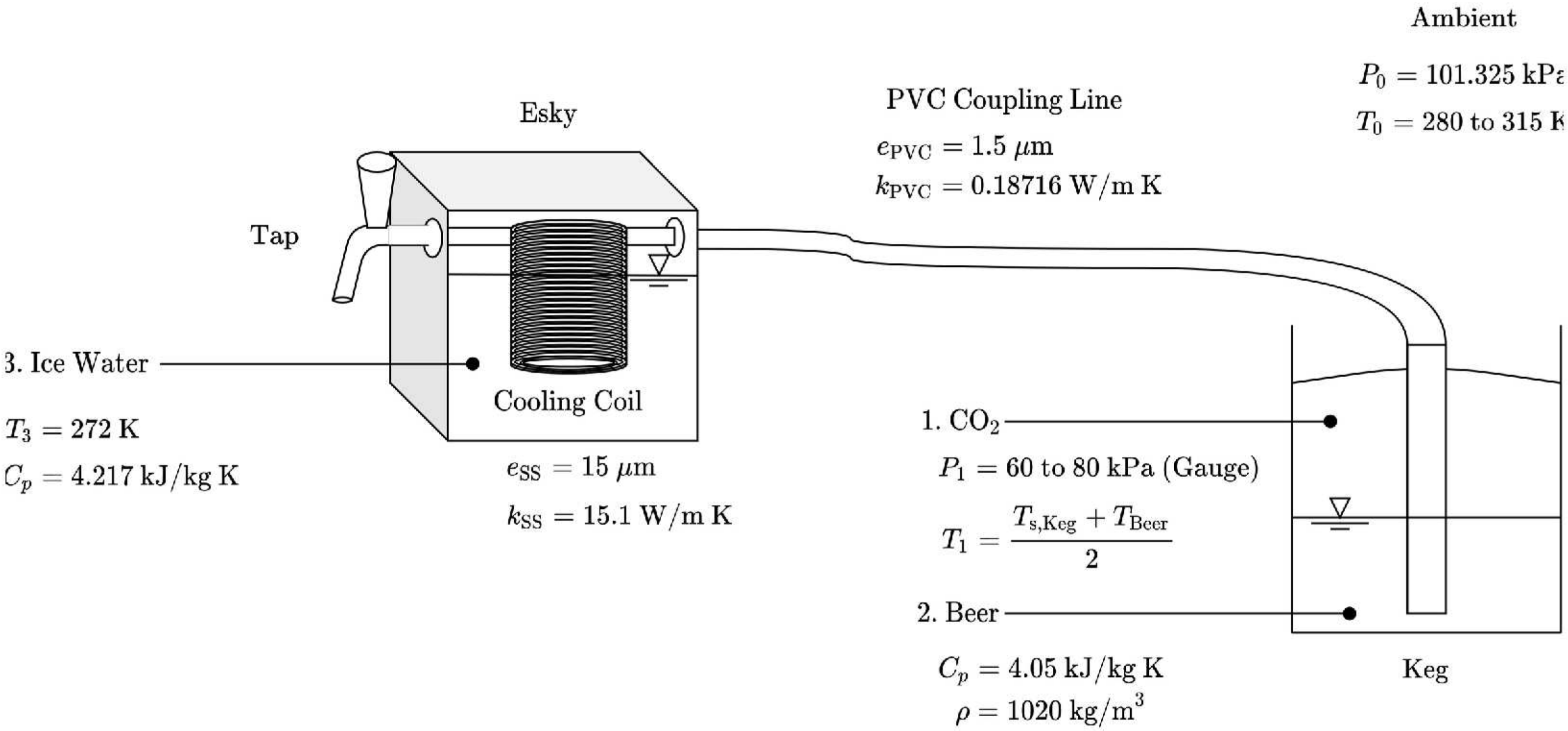
F_Θ

Nature Doesn’t Always Work as Intended



Throwing neural networks at things is cheating.

We study physics/maths/engineering to understand the universe.


Ambient
$P_0 = 101.325$ kP
$T_0 = 280$ to 315
PVC Coupling Line
$e_{\text{PVC}} = 1.5\ \mu\text{m}$
$k_{\text{PVC}} = 0.18716\ \text{W/m K}$
Esky
Tap
3. Ice Water
Cooling Coil
$T_3 = 272\ \text{K}$
$C_p = 4.217\ \text{kJ/kg K}$
$e_{\text{SS}} = 15\ \mu\text{m}$
$k_{\text{SS}} = 15.1\ \text{W/m K}$
1. CO_2
$P_1 = 60$ to 80 kPa (Gauge)
$T_1 = \dfrac{T_{\text{s,Keg}} + T_{\text{Beer}}}{2}$
2. Beer
$C_p = 4.05\ \text{kJ/kg K}$
$\rho = 1020\ \text{kg/m}^3$
Keg

Throwing neural networks at things is cheating.

We study physics/maths/engineering to understand the universe.

Maybe we can make neural nets smarter this way?

All models are wrong, but some models are useful.

George Box

Q&A

Jack Naylor

jack.naylor@sydney.edu.au

nackjaylor.github.io