THE UNIVERSITY OF SYDNEY

ROBOTIC IMAGING LAB

Robotic Imaging:

From Photons to Actions

ACFR
AUSTRALIAN CENTRE FOR FIELD ROBOTICS

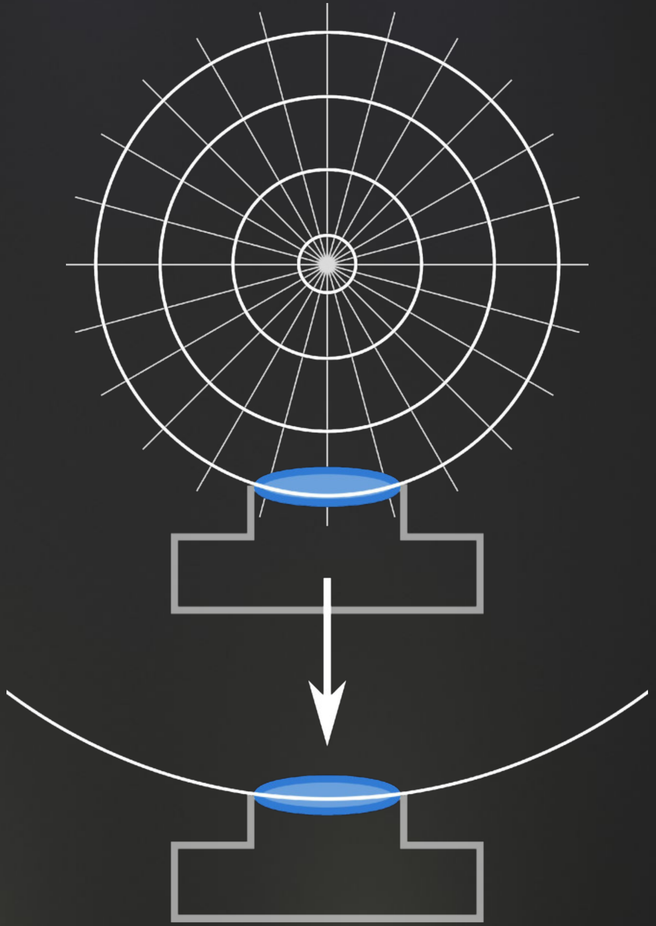
Donald G Dansereau
RVSS
2024·Feb·09

Why don't objects appear dimmer as we move away from them?

Should we expect them to?
Yes, further = less energy into lens
Inverse square law
True for isotropic point sources

$$I \propto \frac{S}{\pi r^2}$$

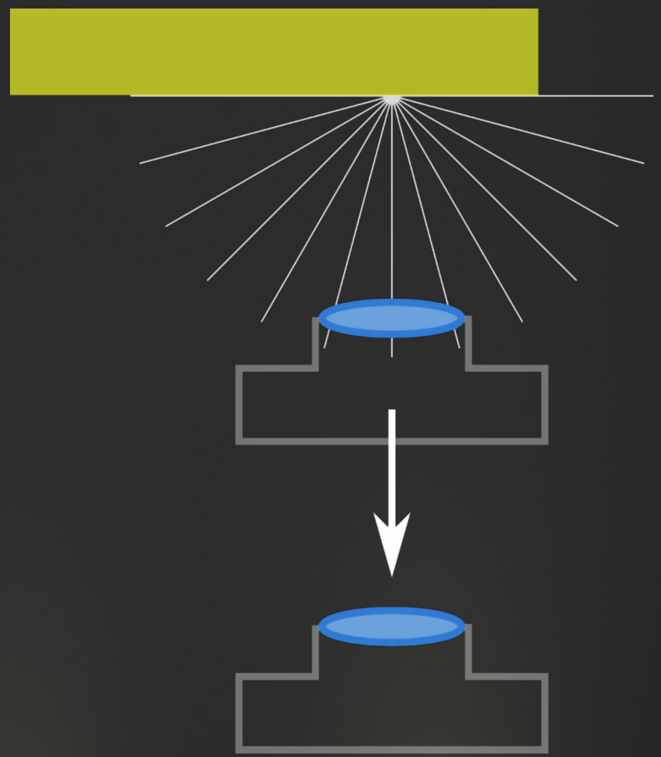
I = measured intensity
S = lens area
r = distance to source



Surfaces are collections of points

Model surface as superposition of points
This doesn't help, still follows inverse square falloff
Hint: What have we missed about the scene?

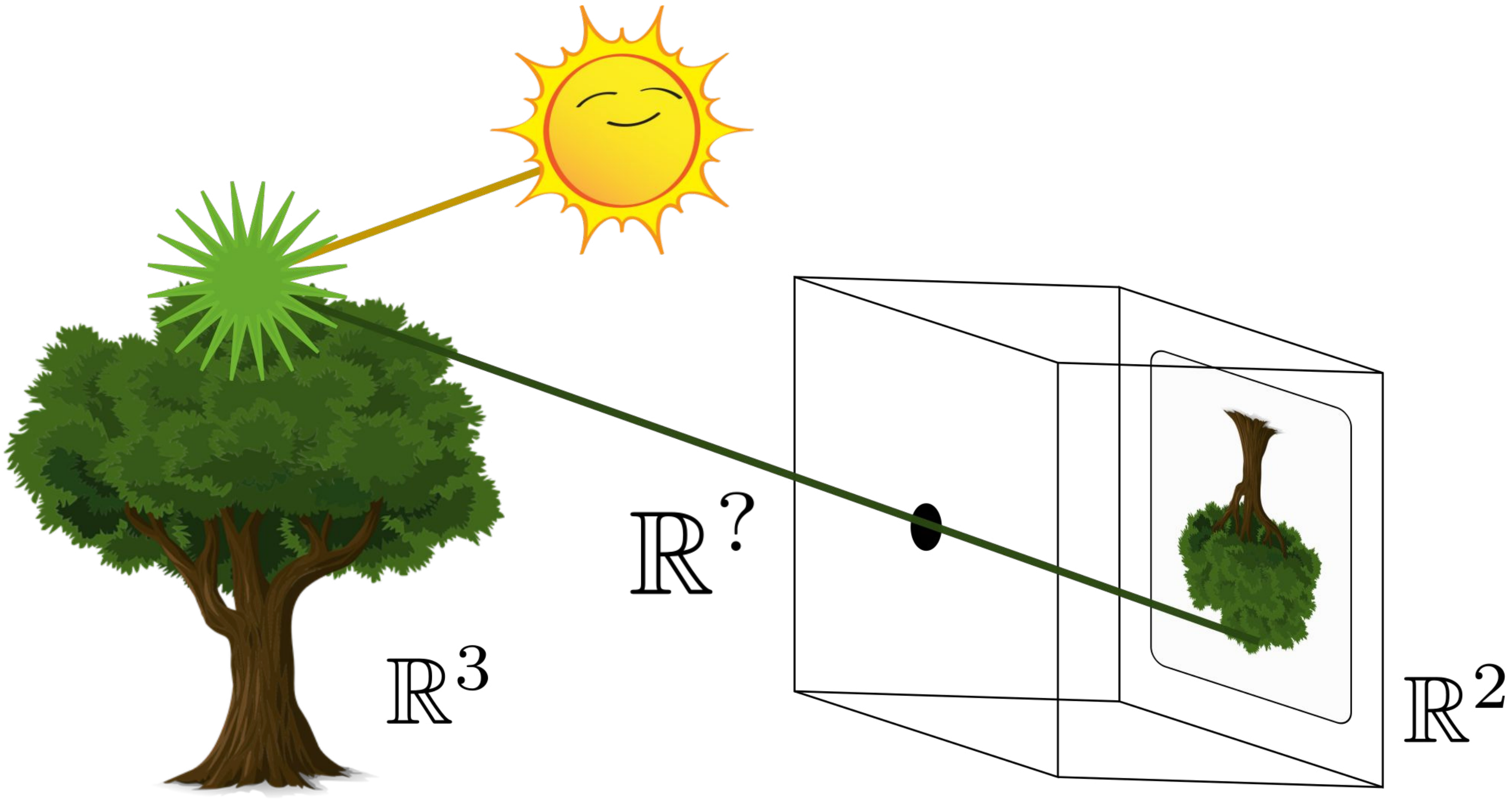
$$I \propto \frac{S}{\pi r^2}$$



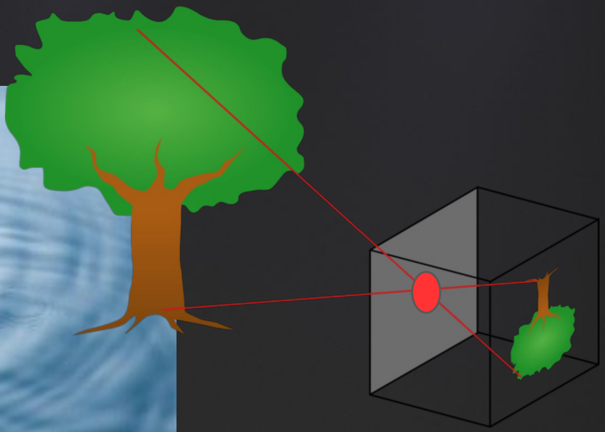
Today's Talk

1. Designing a camera for robots
2. Automating integration of new cameras
3. Making the most of existing cameras
4. Emerging challenges

What Do Cameras Measure?



Light as a Field



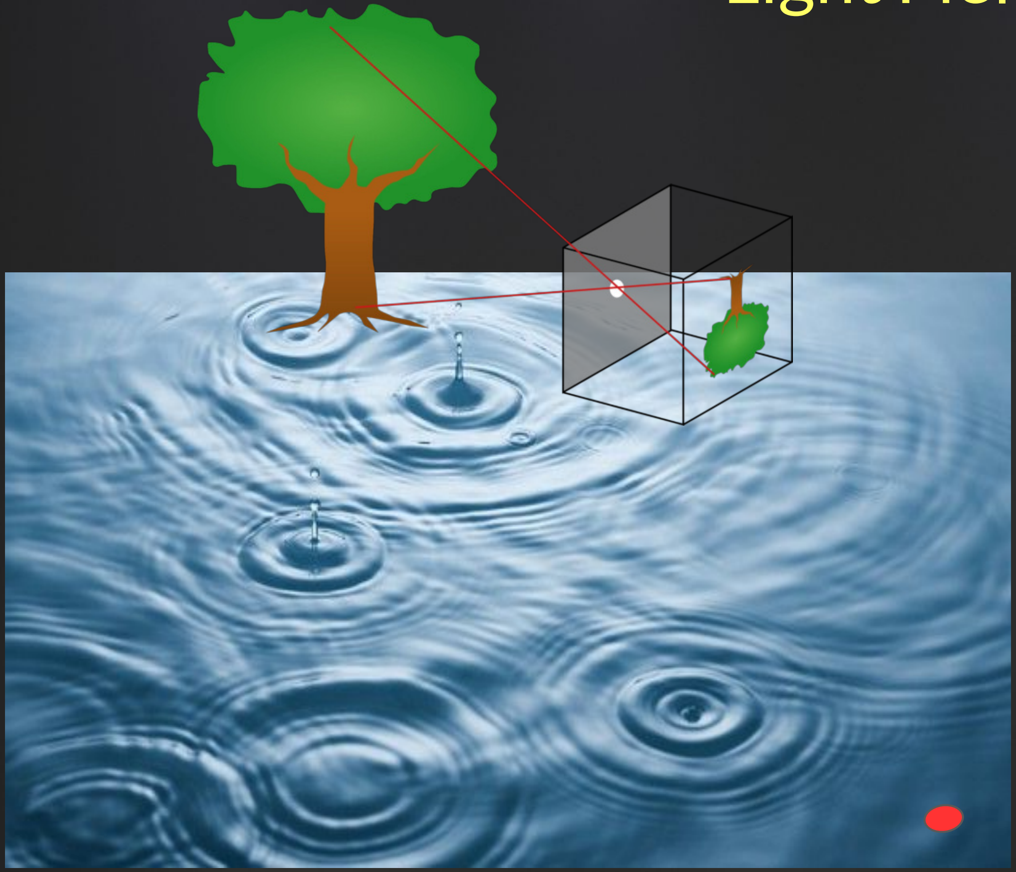
Cameras sample from *a field of light*

- Caused by matter / photon interactions *in the scene*
- Cameras measure light *at the lens*

An analogy: measuring ripple height at some point in a pond

- We measure the ripples, not the drops

Light Field Cameras



Cameras measure waves arriving *at a point*

Light field cameras measure waves arriving *at a surface*

Light Field as Subset of Plenoptic Function

<u>Plenoptic Function $\mathscr{L}(.)$</u>

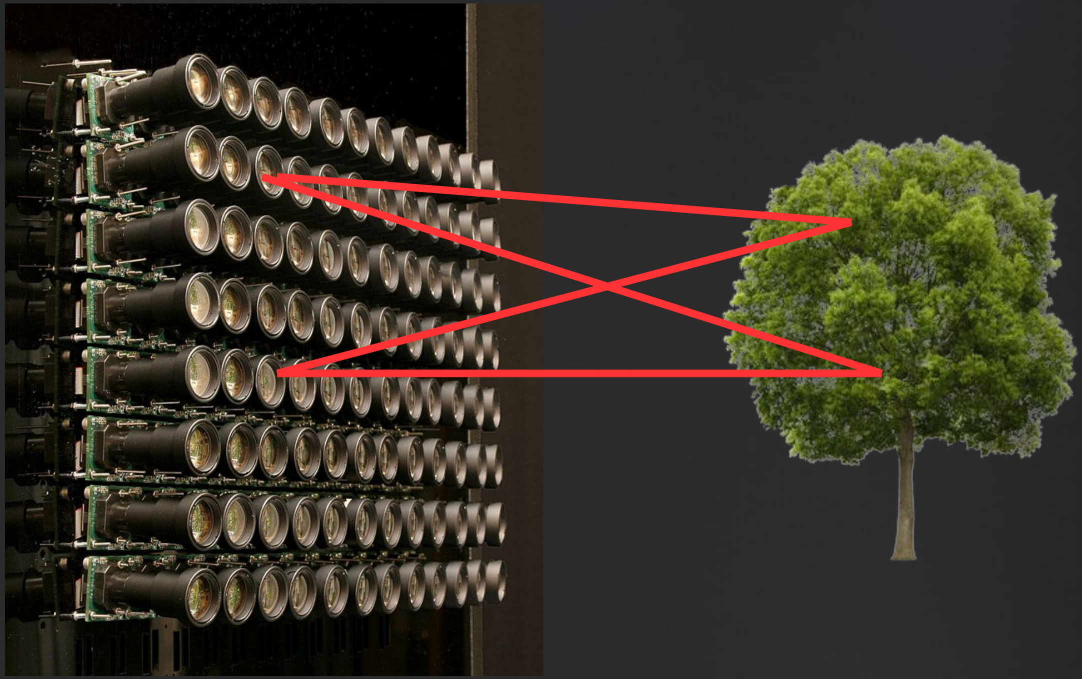
Position x,y,z

Direction u,v

Time

Wavelength

Polarization

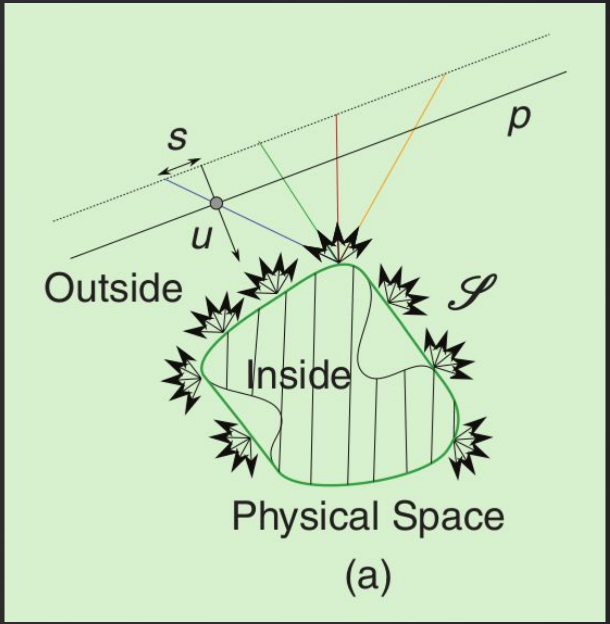


Camera: $\mathscr{L}(u,v)$

NeRF: $\mathscr{L}(x,y,z,\ u,v)$

Light Field: $\mathscr{L}(x,y,\ u,v)$

4D is a Sweet Spot for Light Rays

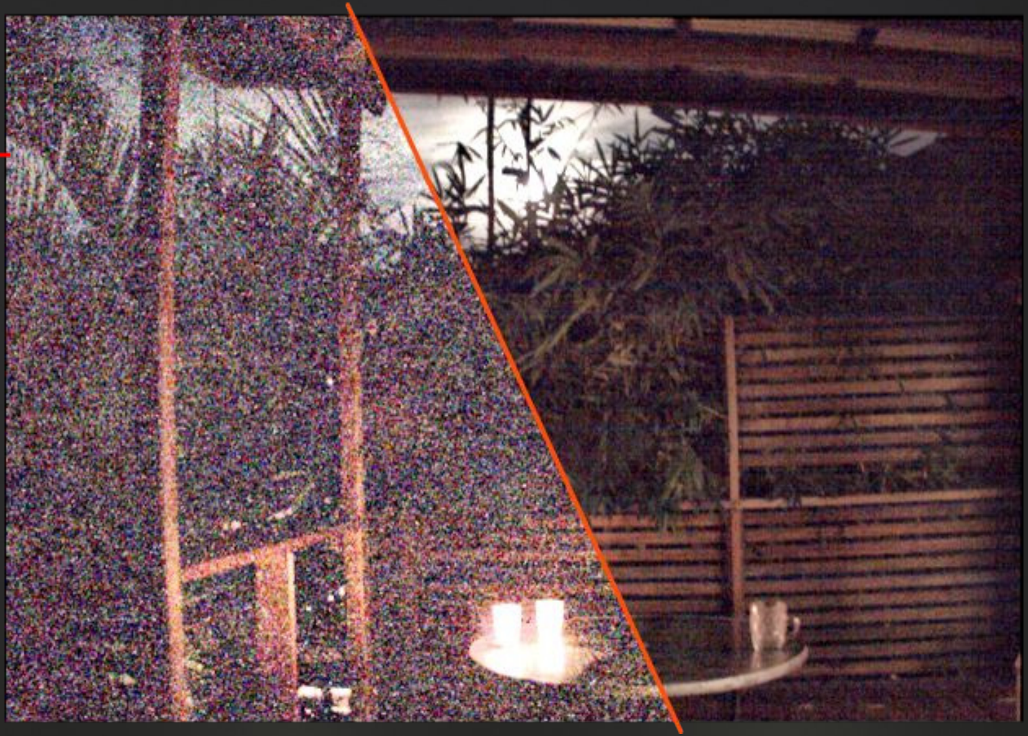


Complete description of light in a volume
1D less than the obvious 5D

Assumes : non-participating medium,
Viewers remain outside the volume

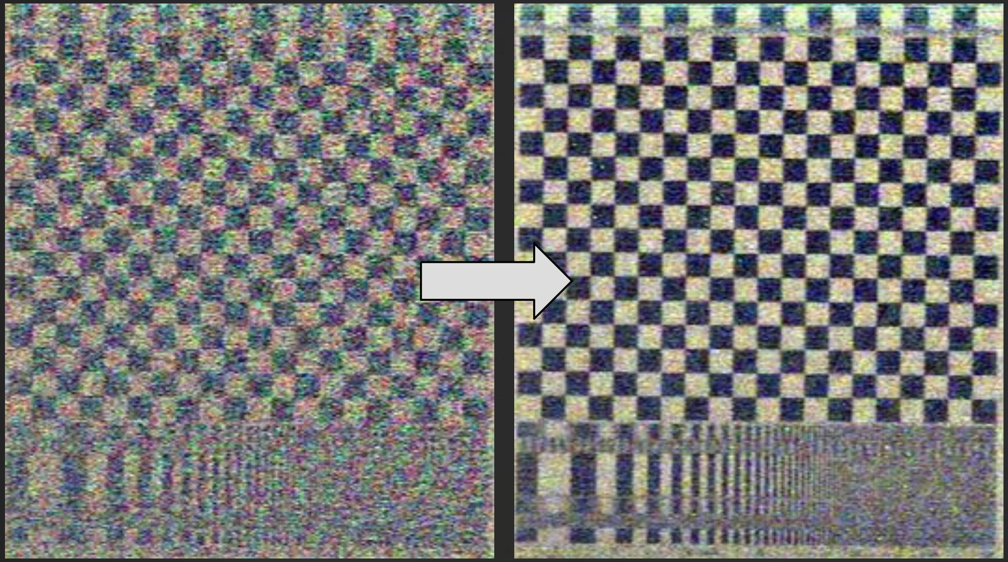
https://augmentedperception.github.io/deepviewvideo/

Light Field Capabilities: Seeing in the Dark

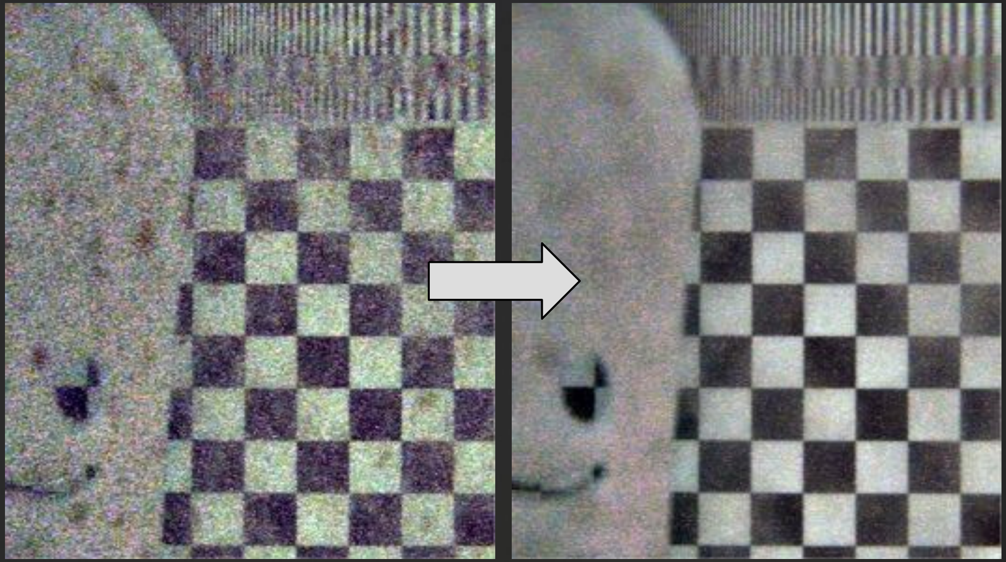


[dansereau2015]

Seeing through Haze and Around Particulate



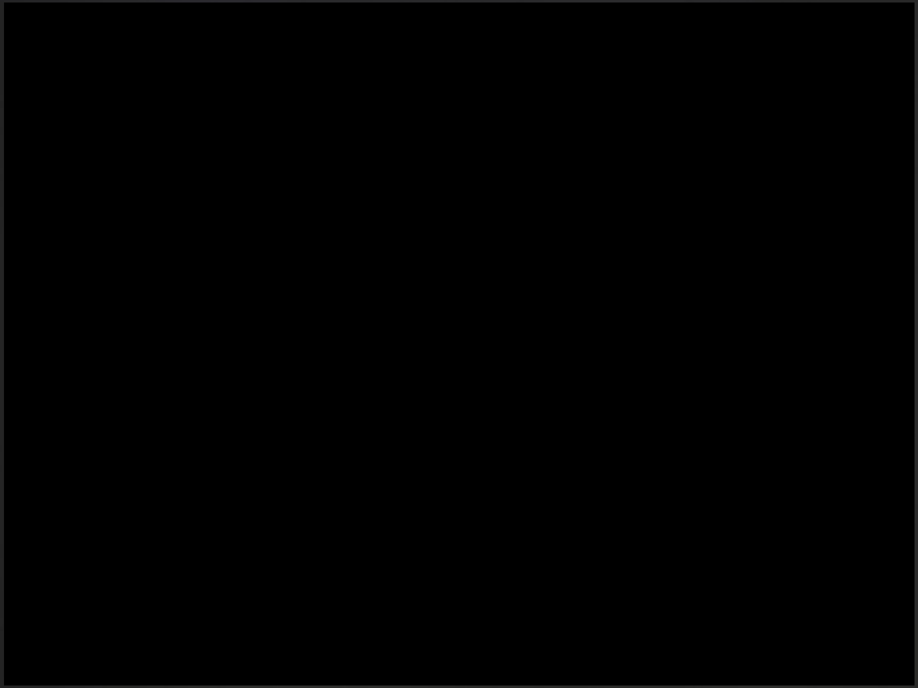
Backscatter (noise-limited)

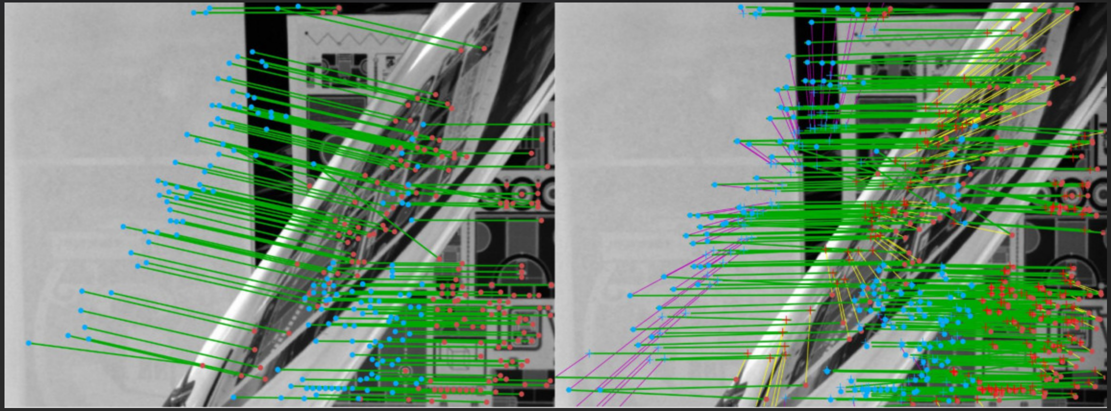


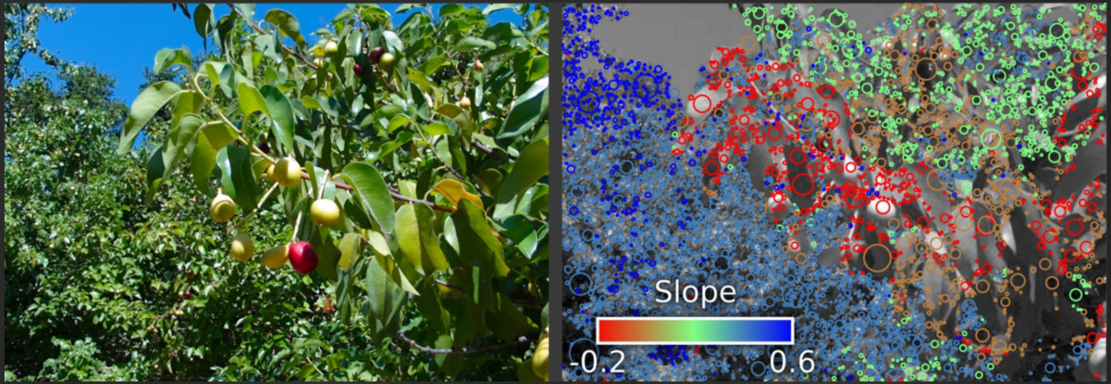
Particulate / occluders

[dansereau2015]

Native Transparency and Dense Structures

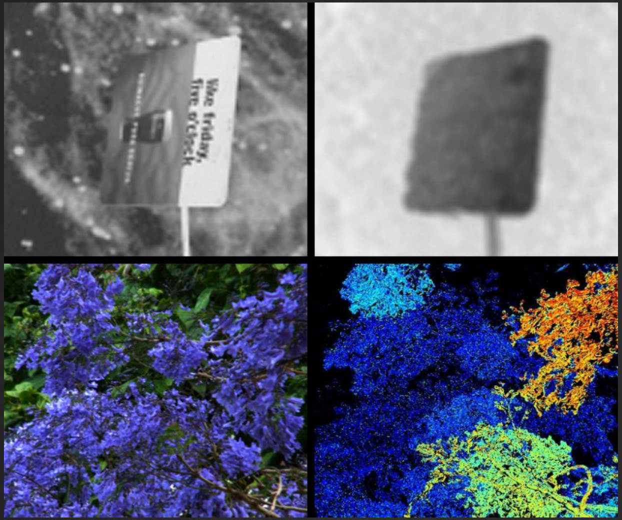






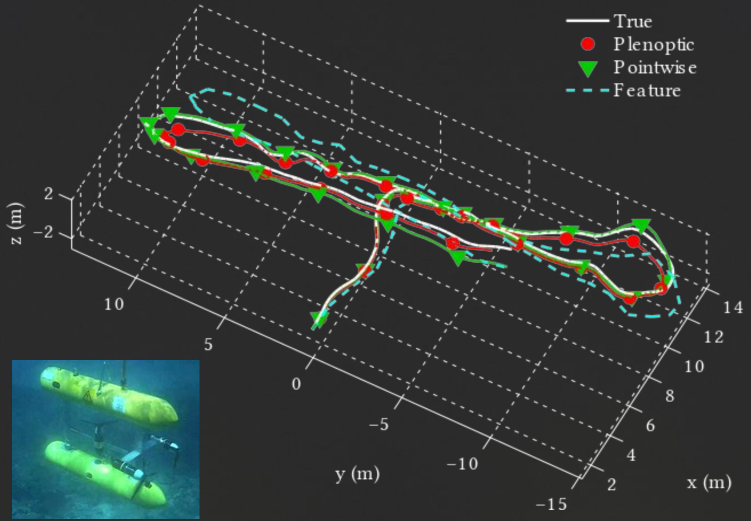
[tsai2021]

3D is Easier in 4D



Fast Depth Estimation

[dansereau2004]



Fast 6D Motion Estimation

[dansereau2011]

Fast 3D Change Detection

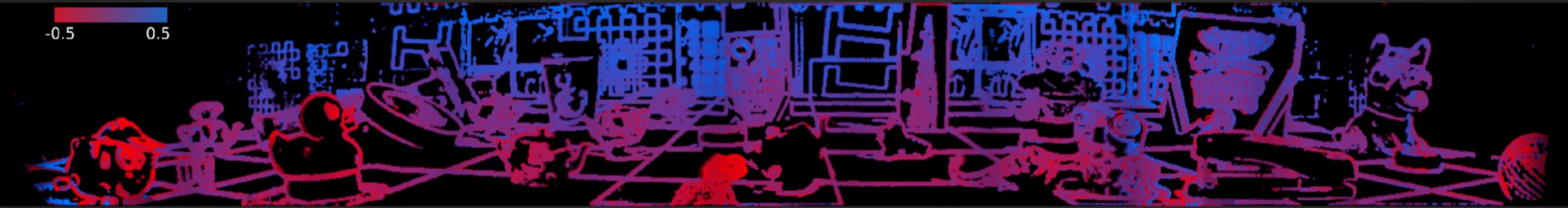
[dansereau2015]

Designing a Camera for Robots: Wide-FOV LF Camera

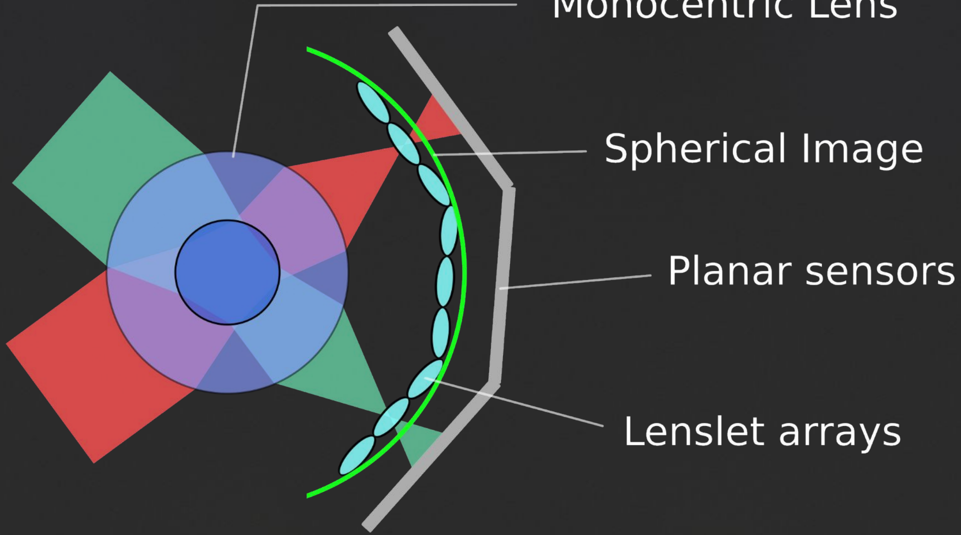
[Gordon Wetzstein
Joseph Ford
Glenn Schuster]

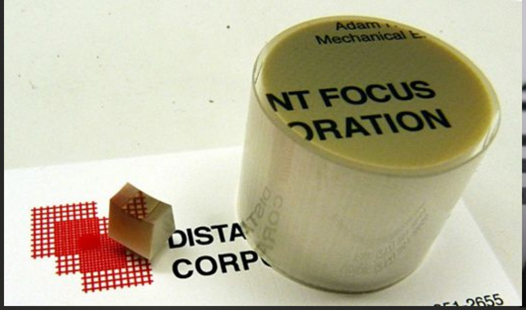
OpEx 2019
CVPR 2017

Low-latency, low-power, hardware-friendly vision



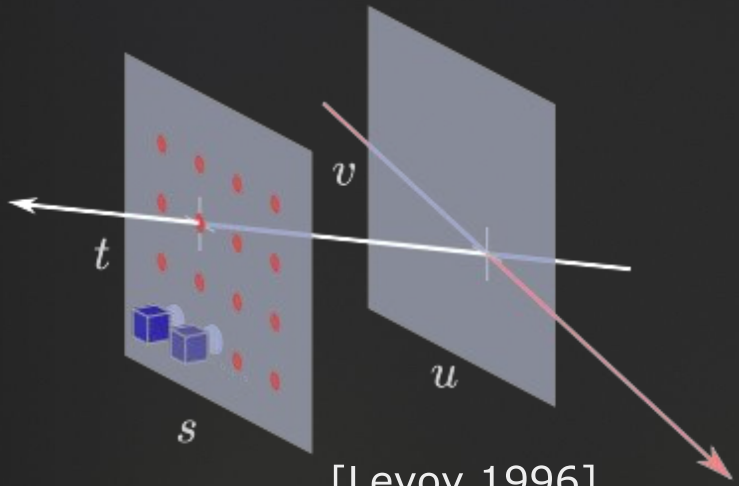
Monocentric-Lens LF Camera



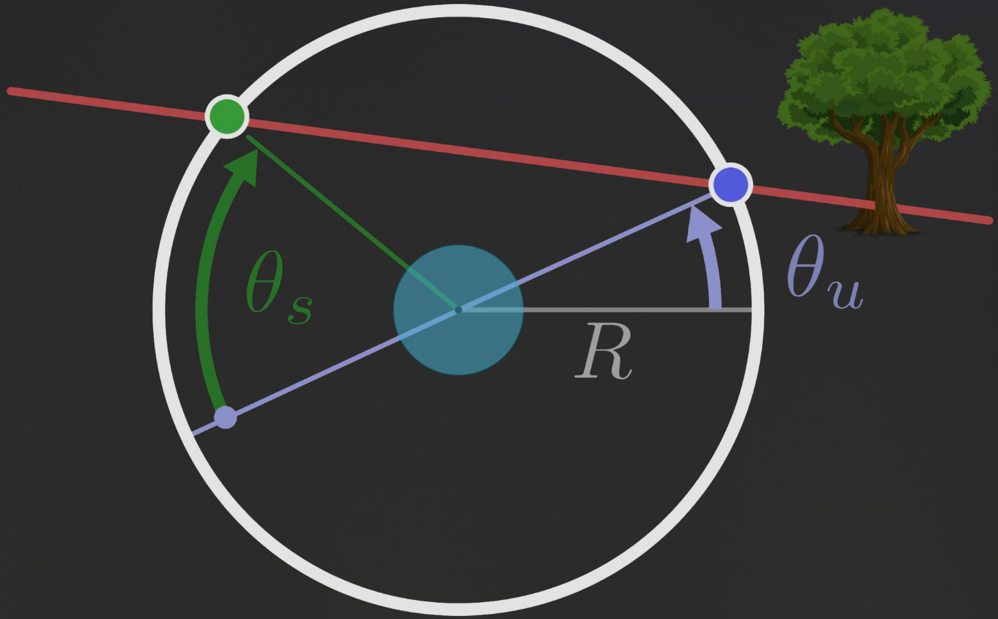


Fiber Bundles
→ Replace with LF processing

Spherical Parameterization



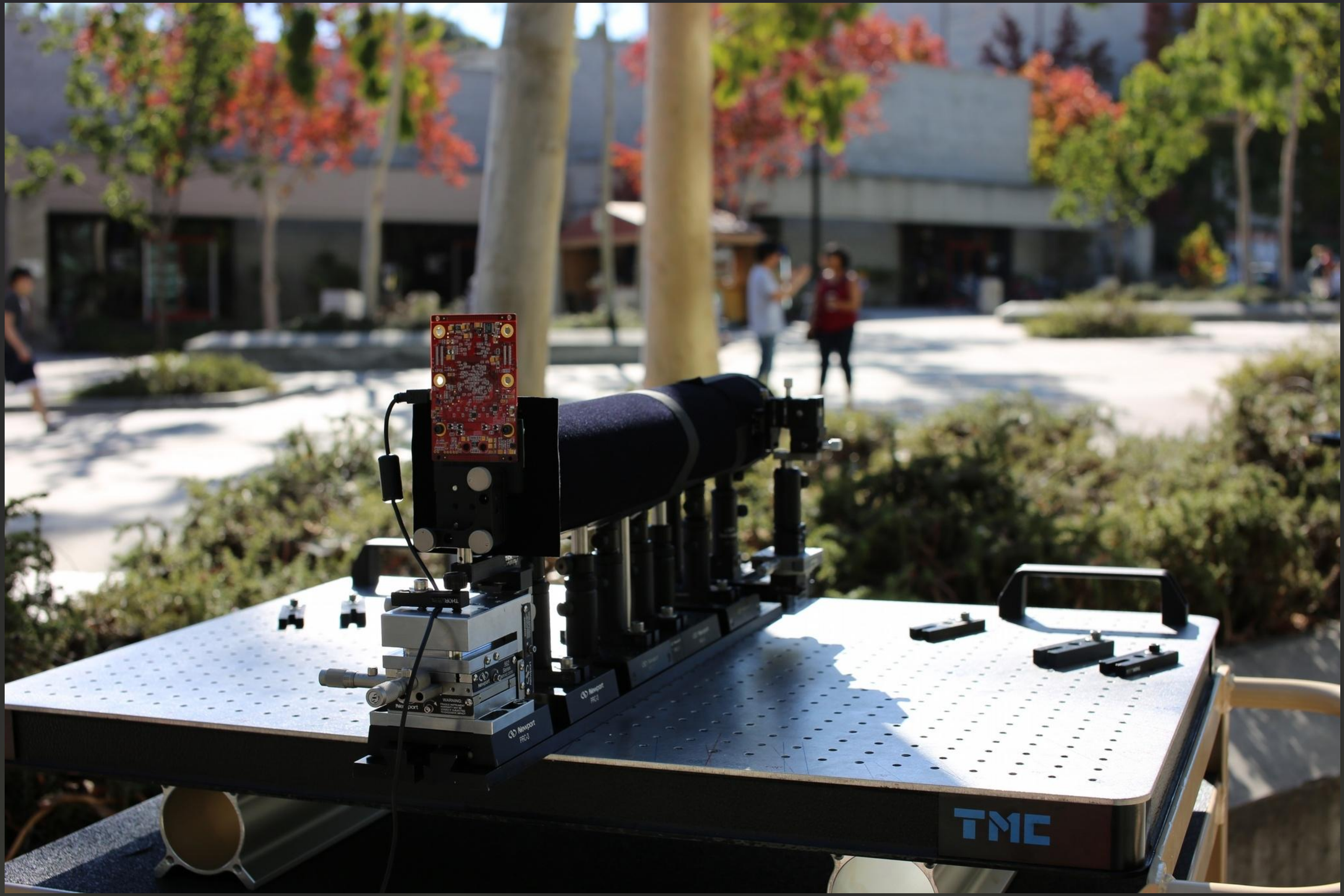
[Levoy 1996]



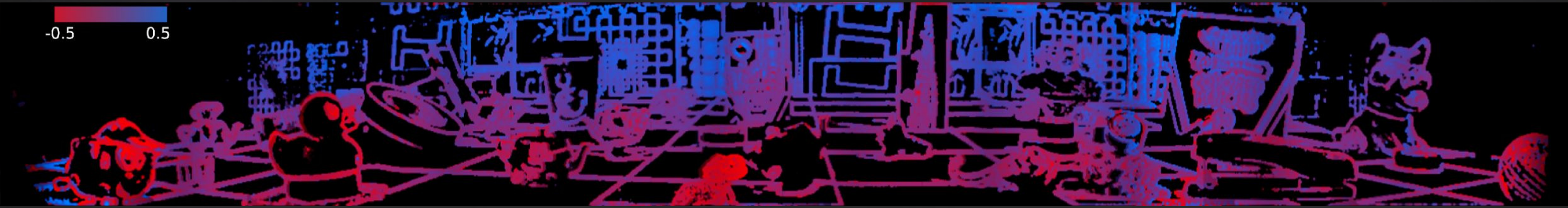
R = foc dist

Locally 2pp-compatible
Existing tools work!

1st Prototype



Panoramas



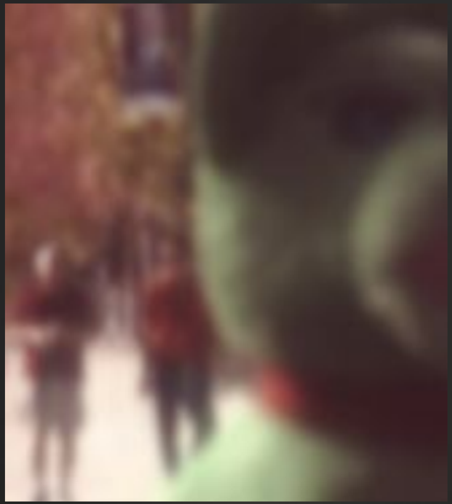
15 × 15 × 1600 × 200 (72 MPix)   138° FOV

Refocus:

Panoramas

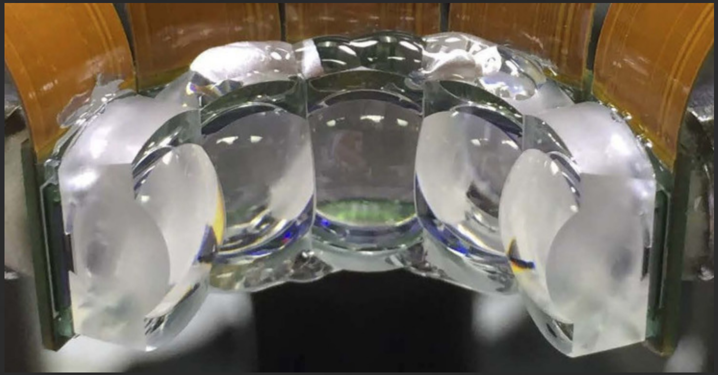
15 × 15 × 1600 × 200 (72 MPix)   138° FOV

Refocus:

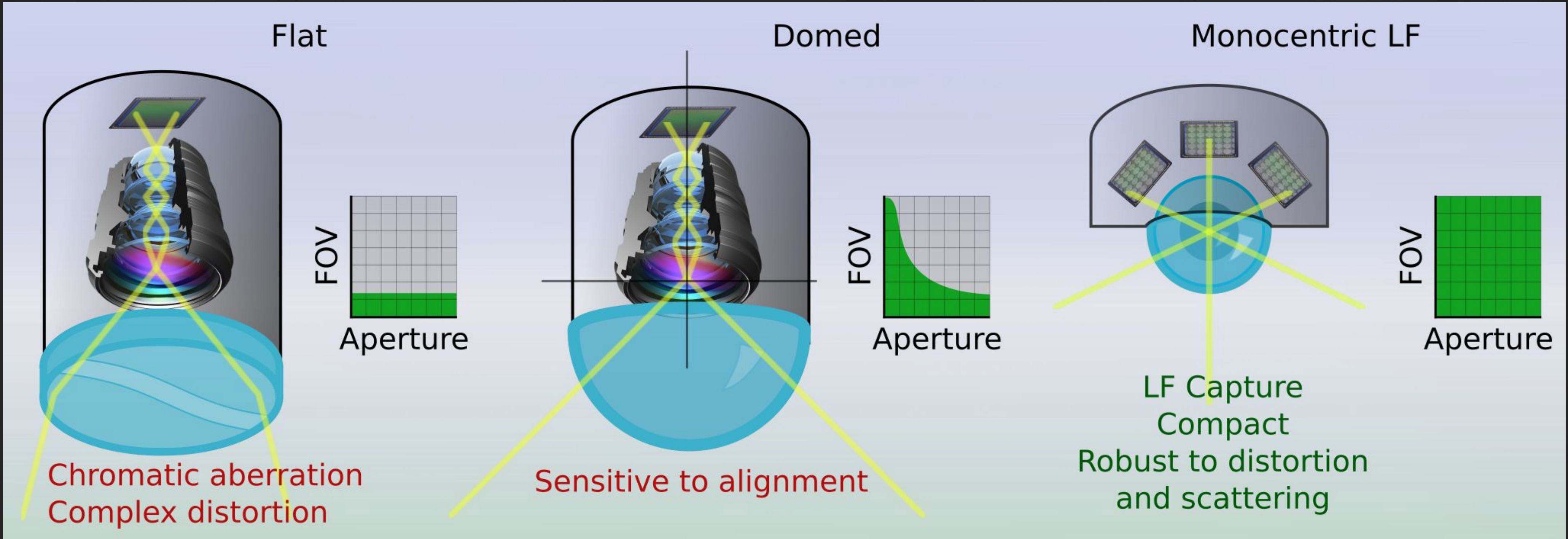


Super-resolution:

Compact Prototype



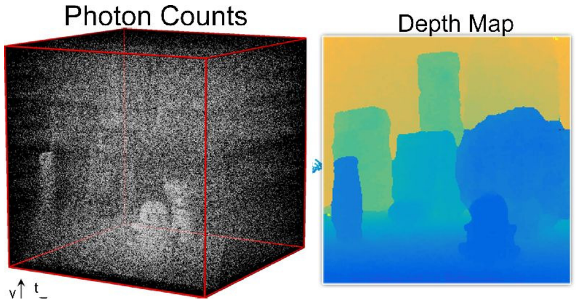
Next Steps: Underwater



Today's Talk

1. Designing a camera for robots
2. **Automating integration of new cameras**
3. Making the most of existing cameras
4. Emerging challenges

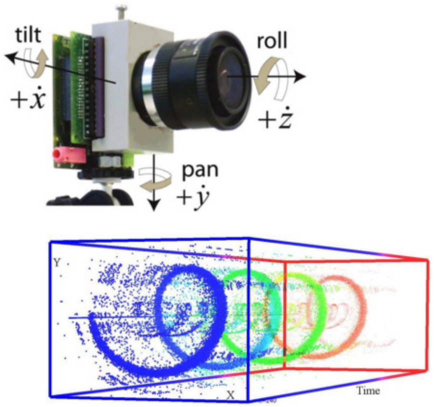
New Cameras = New Ways of Seeing, But…



Single-Photon
Depth (SPADs)

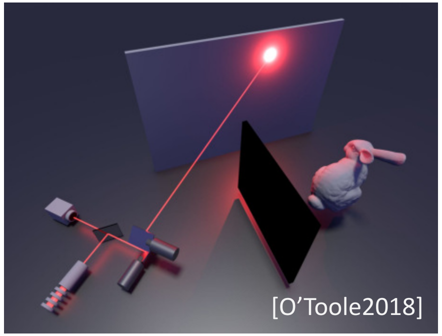
[EPIImaging LLC]

Sparse LF
cameras

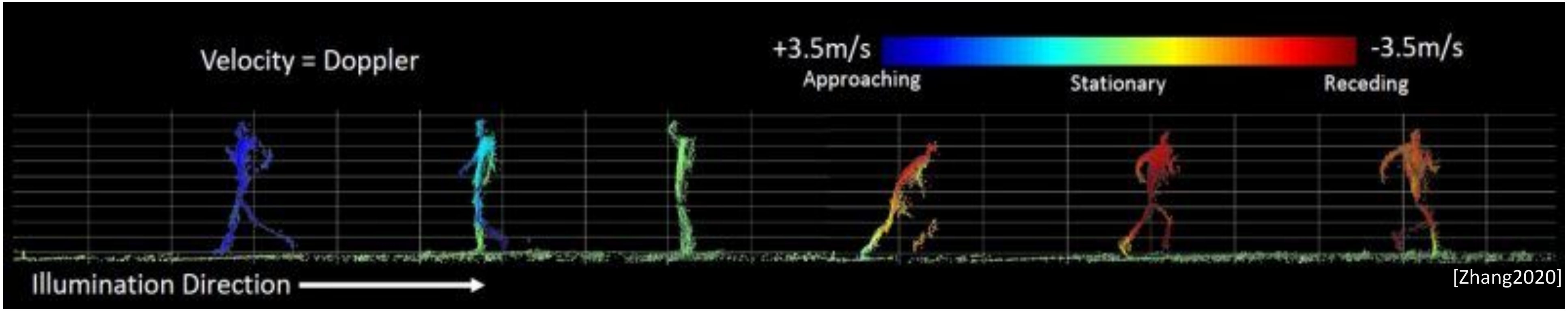


[Gallego2017]

Event cameras



Non-line-of-sight imaging



Doppler LiDAR

Automating Integration of New Cameras

[Teja Digumarti
Joe Daniel
Avie Ravendran
Ryan Griffiths
Jack Naylor
ICRA 2023
IROS 2021]

IROS 2021 : Odometry and 3D from LF cameras


[Raytrix]

Lenslet-based cameras


[K|Lens]

Lens Adaptors

[Bajpayee2018]

Linear Array

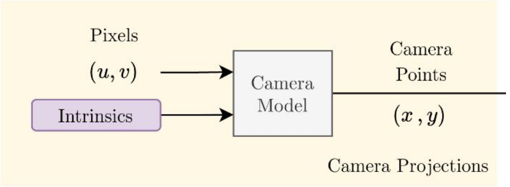
[EPIImaging LLC]

Sparse Array

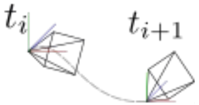

[Pelican Imaging]

Full Array

ICRA 2023: Automating calibration “NOCaL”



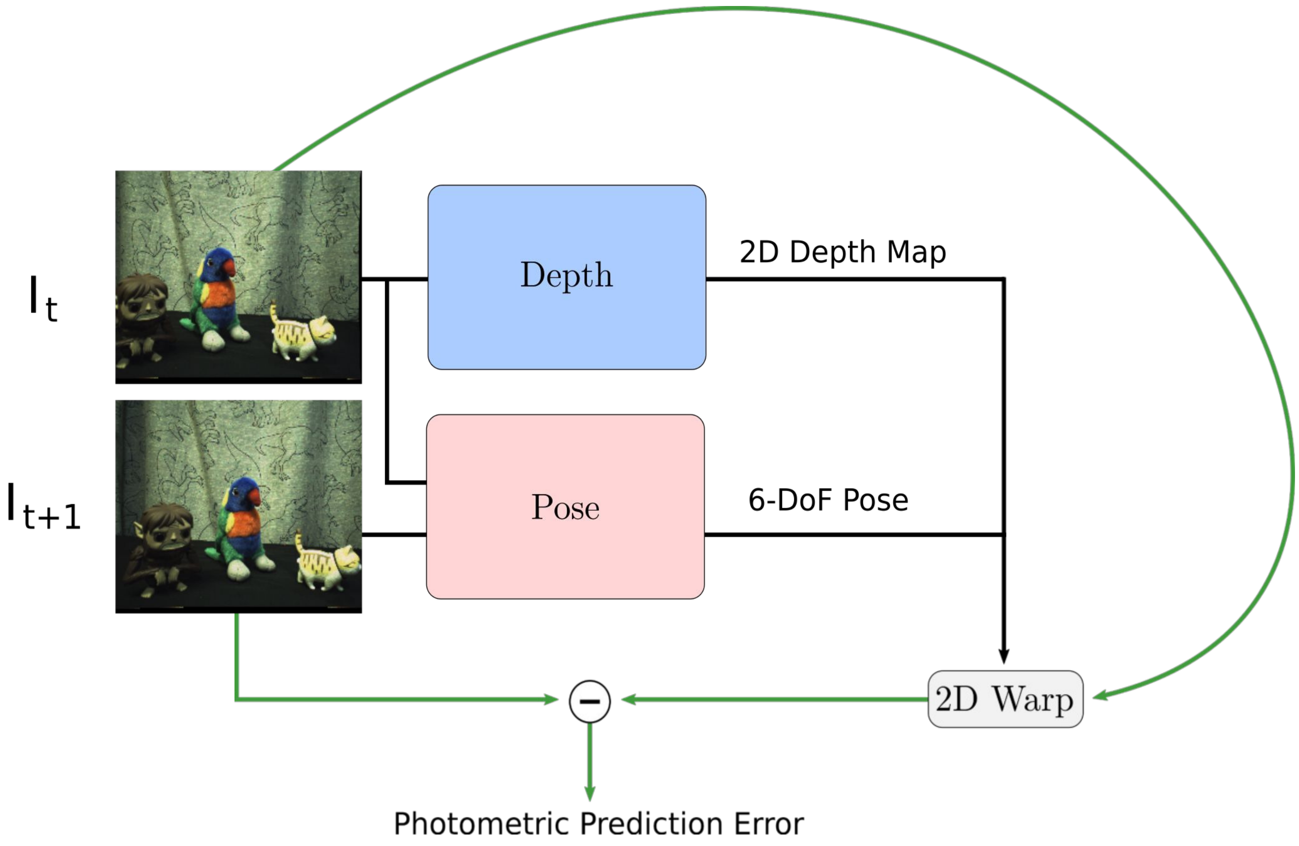
Data-driven
freeform distortion model



Odometry

Novel View Synthesis

Prediction as Supervisory Signal



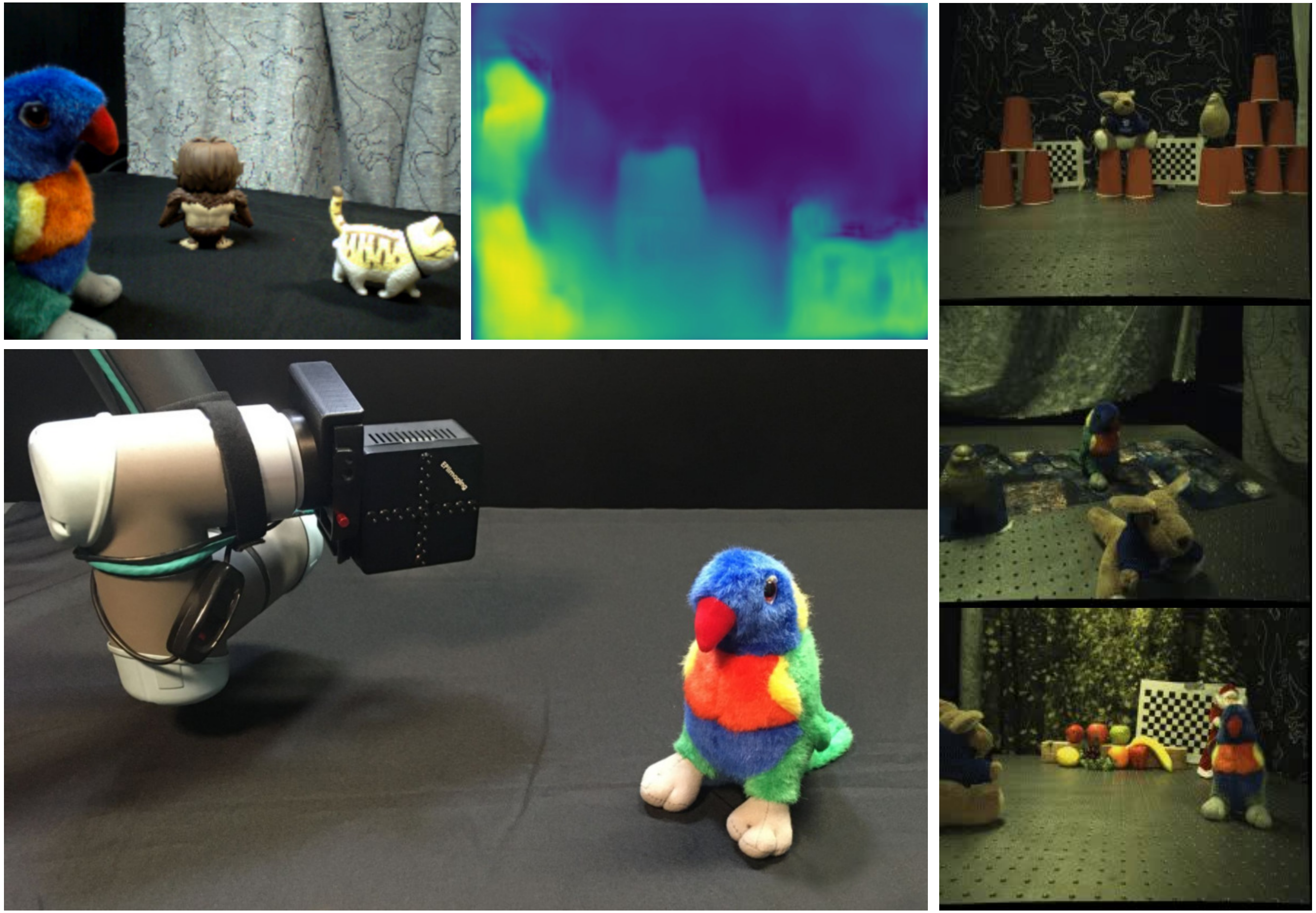
[Zhou2017] : Prediction as supervision

[Zhan2018, Dharmasiri2018] : Learning from Stereo

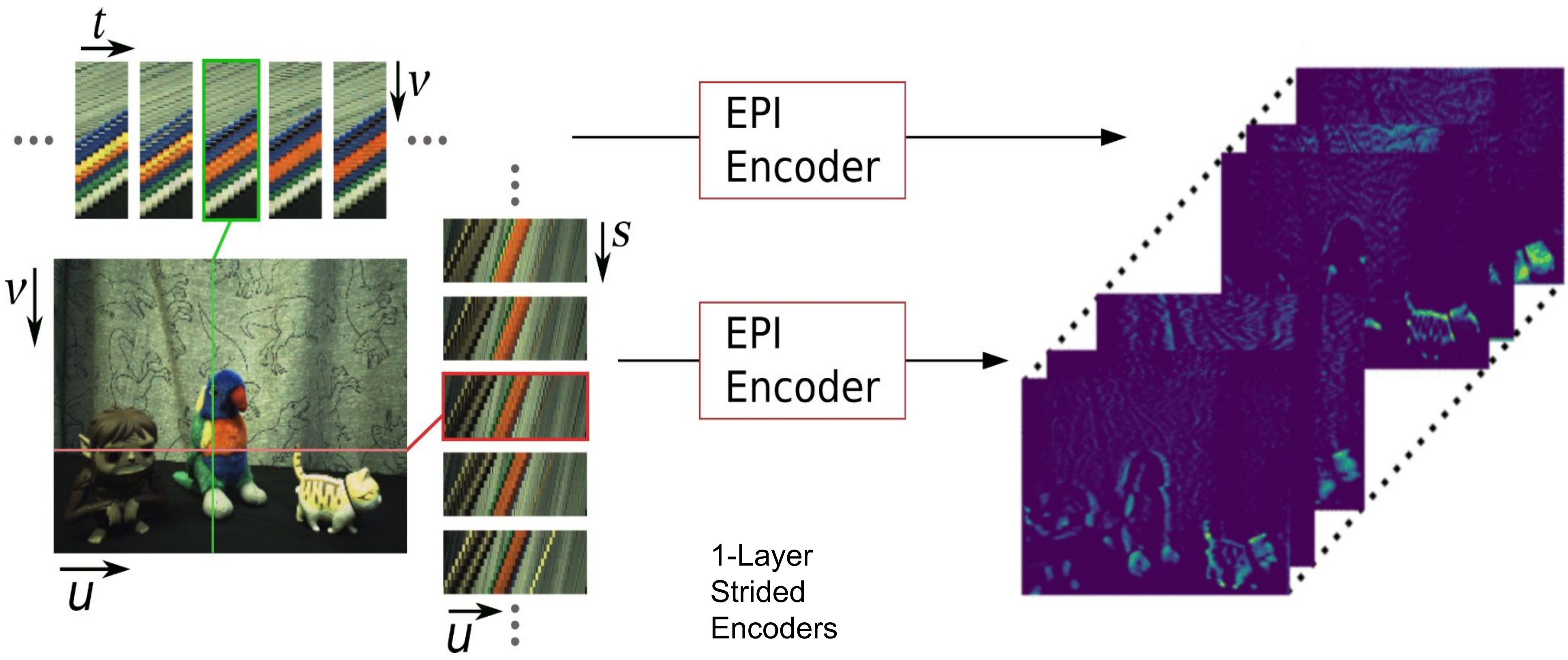
Adapting to Sparse Light Field Cameras

RoboticImaging.org/Projects/LearnLFOdo/

Sparse LF Video
8298 x 17-views
Arm-mounted



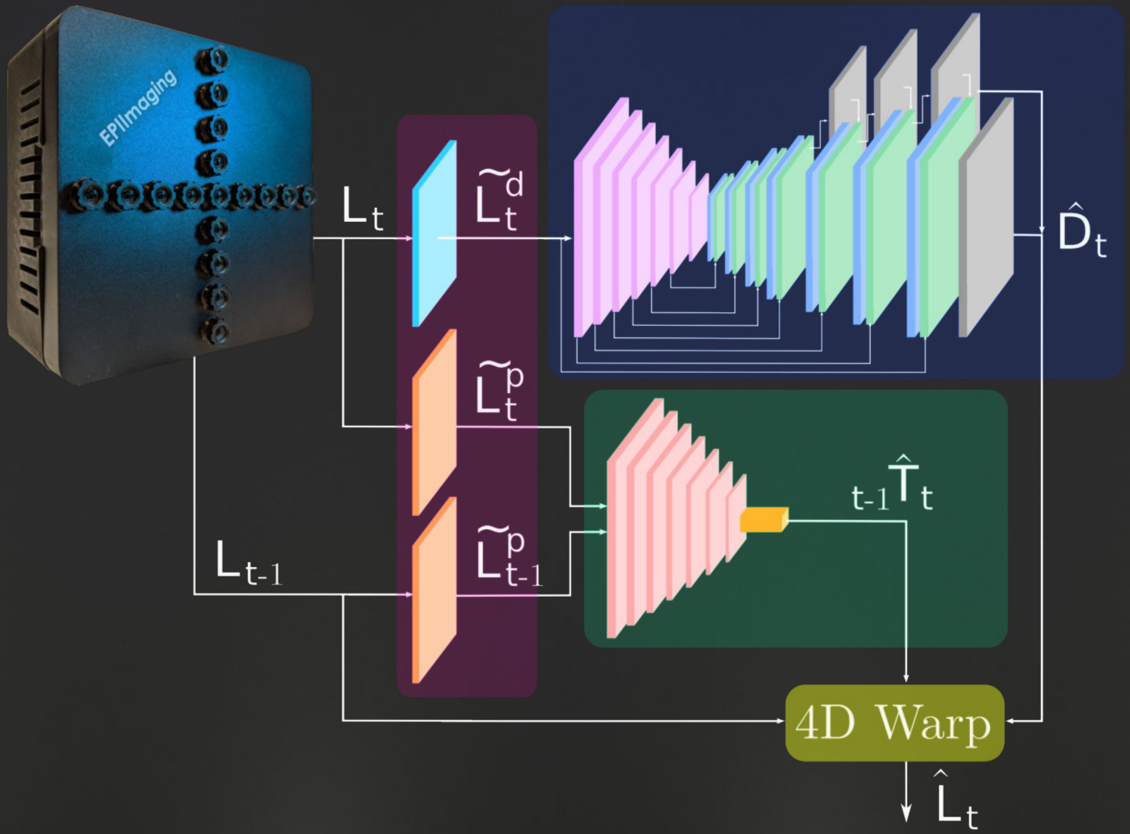
A General LF Encoder



Tiled Epipolar Plane Images

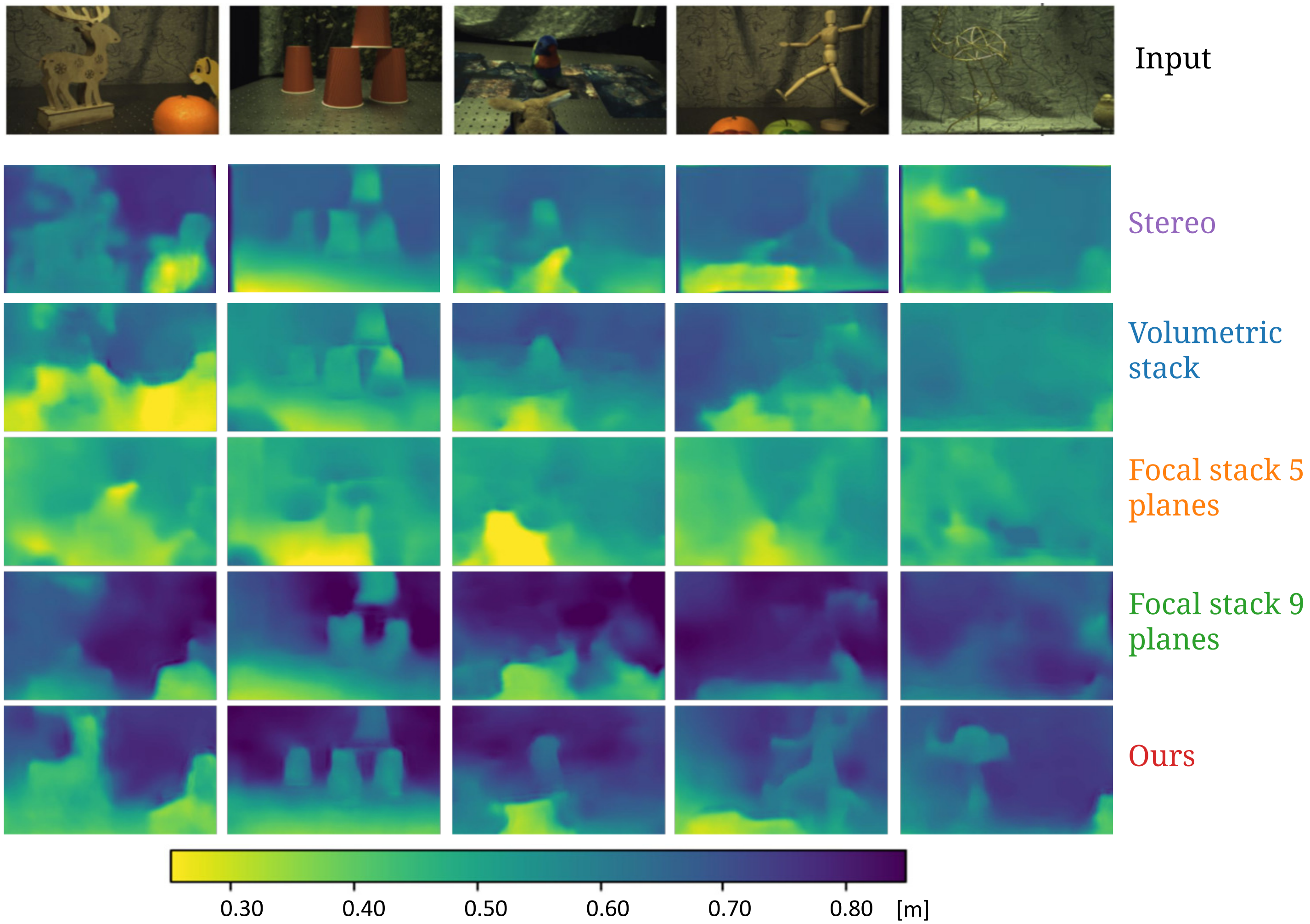
Encoded Epipolar Stacks

Self-Supervised Odometry and Depth



Results

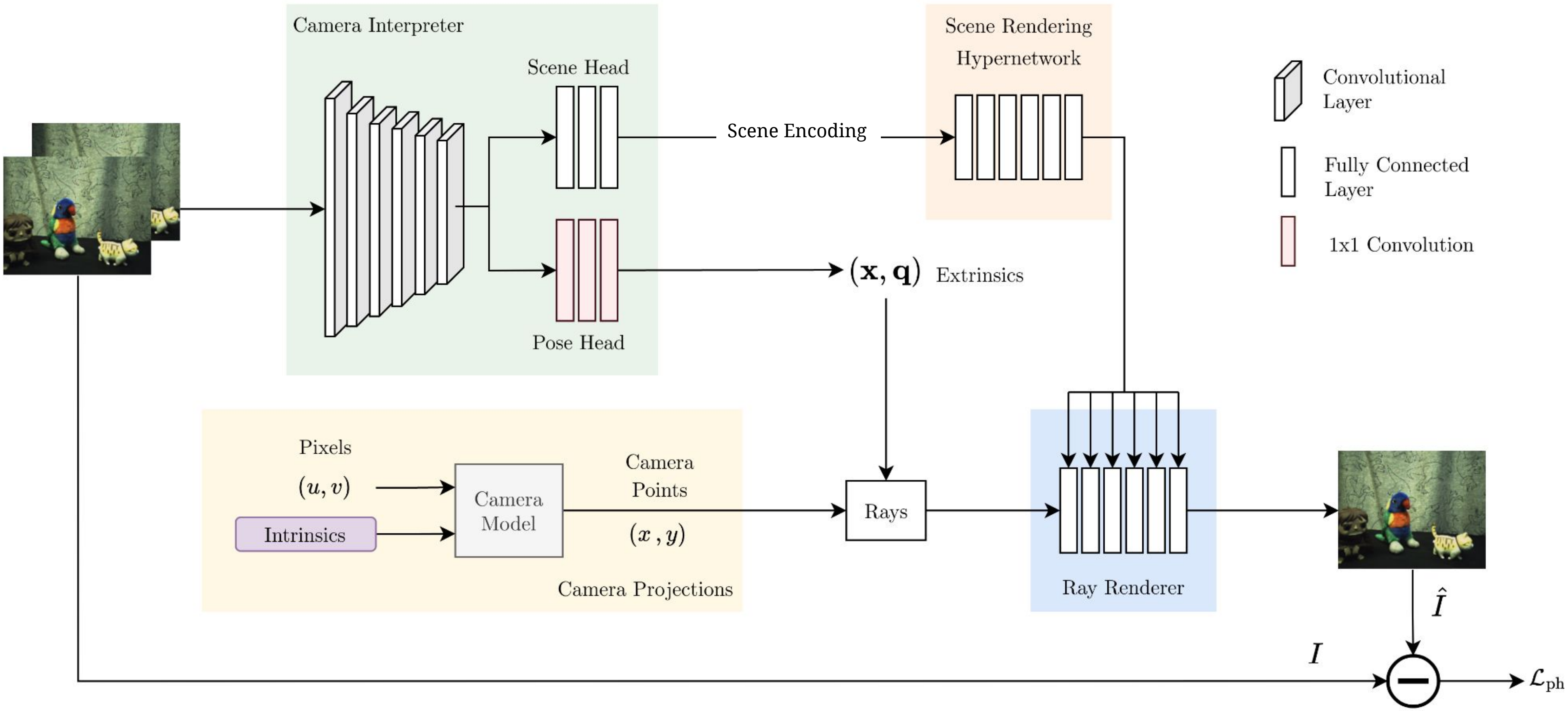
Results: Depth



Results: Odometry

		Relative pose error in translation [m]		Relative pose error in rotation [deg]	
		Mean	Std.dev	Mean	Std.dev
Single-warp	Monocular	0.029	0.016	1.522	0.969
	Volumetric stack	0.028	0.015	1.453	0.880
	Focal stack – 5	0.030	0.015	1.439	0.883
	Focal stack – 9	0.030	0.016	1.452	0.912
	Ours	0.026	0.016	1.302	0.885
Multi-warp	Volumetric stack	0.024	0.016	1.366	0.802
	Focal stack – 5	0.031	0.017	1.457	0.779
	Focal stack – 9	0.035	0.017	1.585	0.868
	Stereo	0.028	0.015	2.413	1.655
	Ours	**0.023**	0.017	**1.282**	1.311

Generalising the Camera with LF Meta-Rendering



LF Rendering Novel Views & Novel Cameras

Input Pair

Novel Views

Novel Camera Geometries

Odometry & Calibration Results

Odometry

Method	Labelled Images	Unlabelled Images	Translation Error [m]			Rotation Error [degrees]		
			Mean	STD	RMSE	Mean	STD	RMSE
Odometry accuracy on captured indoor imagery								
Fully supervised	800	0	0.025	0.009	0.027	1.553	1.847	2.414
Unlabelled calibrated	0	8000	0.029	0.016	0.033	1.522	0.969	1.808
NOCaL (ours)	800	7200	**0.020**	**0.008**	**0.022**	**0.412**	**0.295**	**0.505**
Ablation study using rendered indoor imagery with camera distortion								
Ours no intrinsics or distortion	100	900	0.157	0.060	0.168	8.026	9.180	12.194
Ours no distortion	100	900	0.147	**0.053**	0.156	4.790	2.209	5.275
Ours full	100	900	**0.145**	0.054	**0.154**	**4.024**	**1.971**	**4.481**

Calibration

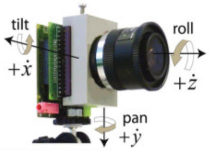
Camera	Initial		NOCaL (ours)		COLMAP [1]	
	f [px]	Error $\overline{\Delta r}$	f [px]	Error $\overline{\Delta r}$	f [px]	Error $\overline{\Delta r}$
Focal: 600px, no distortion	640.0	0.0225	601.1	0.0008	599.4	0.0003
Focal: 600px, large distortion	640.0	0.0440	595.3	0.0186	602.1	0.0010

[Schonberger, 2016]

Next Steps

More Sensors, Modes

IMUs



Mr Potato Head:
Plug-and-play
more imaging technologies

Today's Talk

1. Designing a camera for robots
2. Automating integration of new cameras
3. **Making the most of existing cameras**
4. Emerging challenges

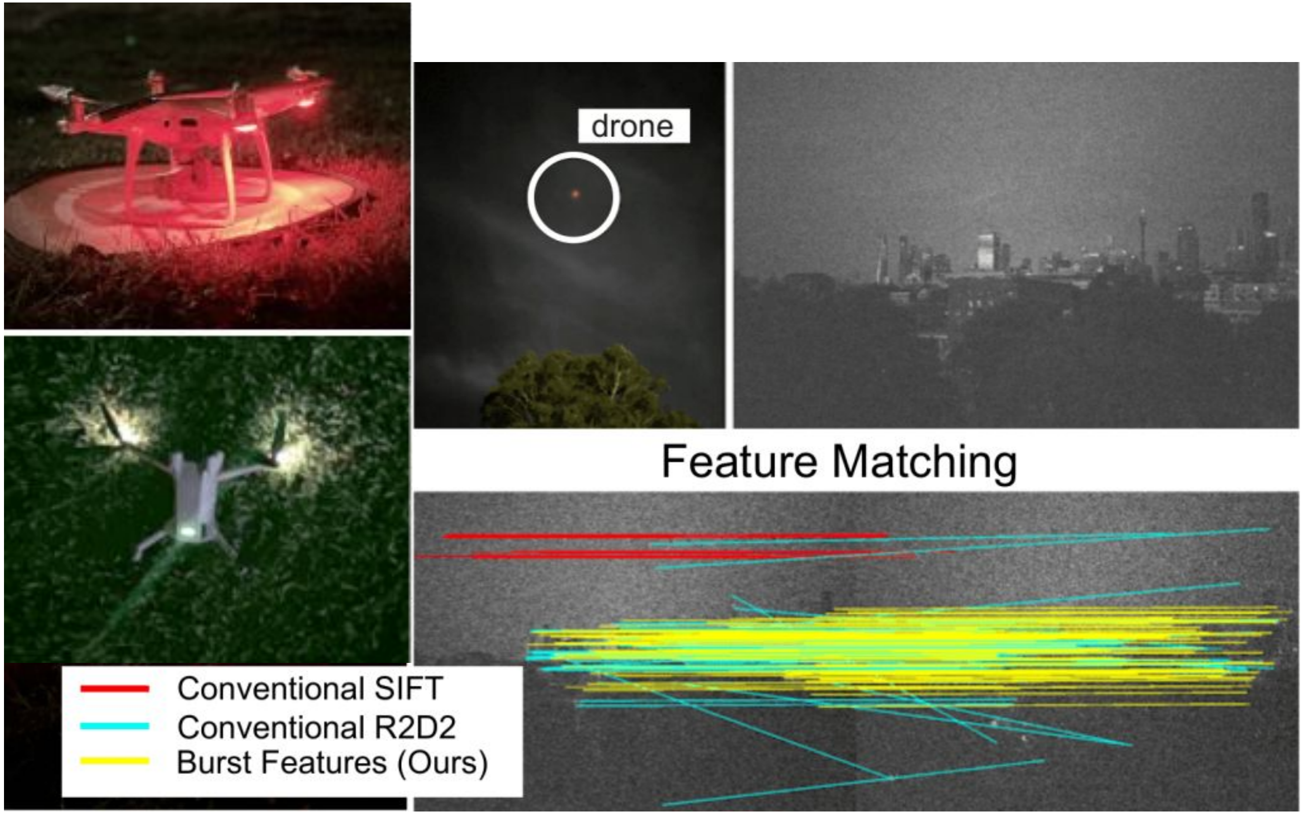
Making the Most of Cameras: Burst

Avie Ravendran
Mitch Bryson

RA-L/ICRA 2024
ICRA 2022

Seeing in mLux conditions

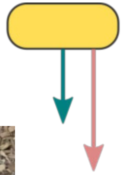
Exploiting statistics of platform dynamics and image formation

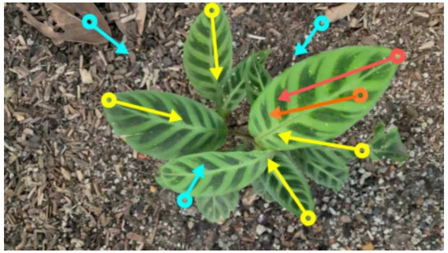


Burst Imaging in Low Light

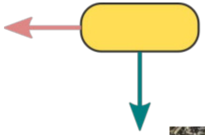
[hasinoff2016, liba2019, ...]

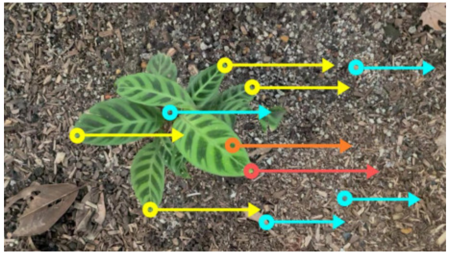
Platform Motion in a Burst





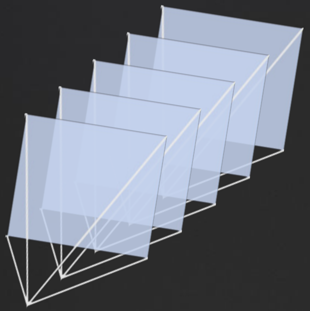
Platform inertia
→ local linear apparent motion
Not all the time
… but much of the time



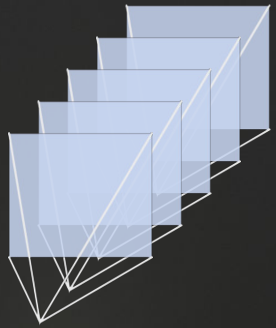


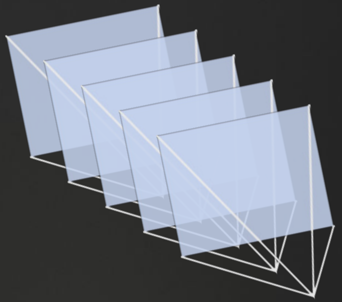
Approach: Hybrid Dense + Sparse

Within burst: 2D+T feature detection, no merging
Exploits visual coherence, platform dynamics



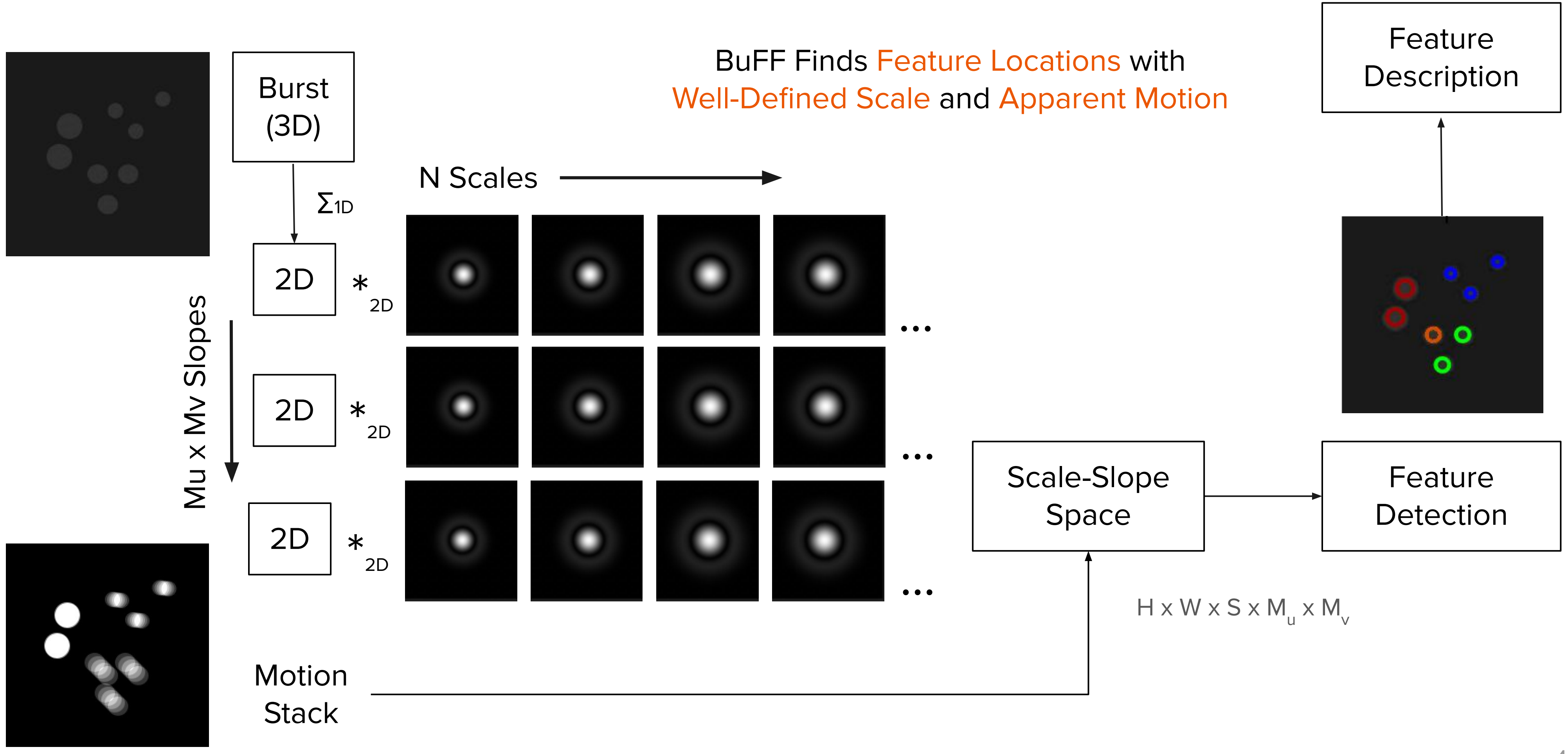
Between bursts: Sparse feature matching
Tolerates appearance changes, fast



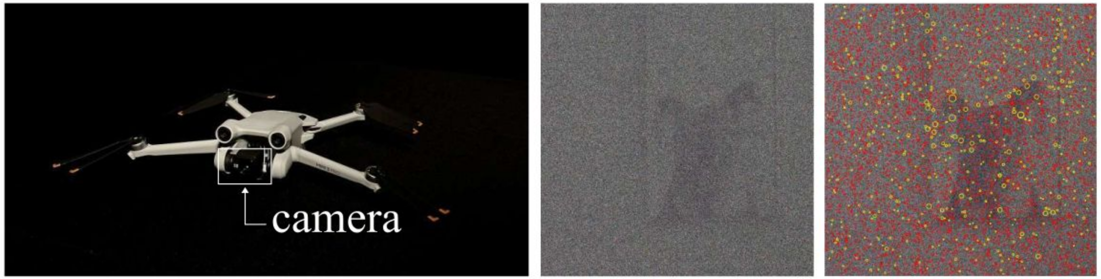


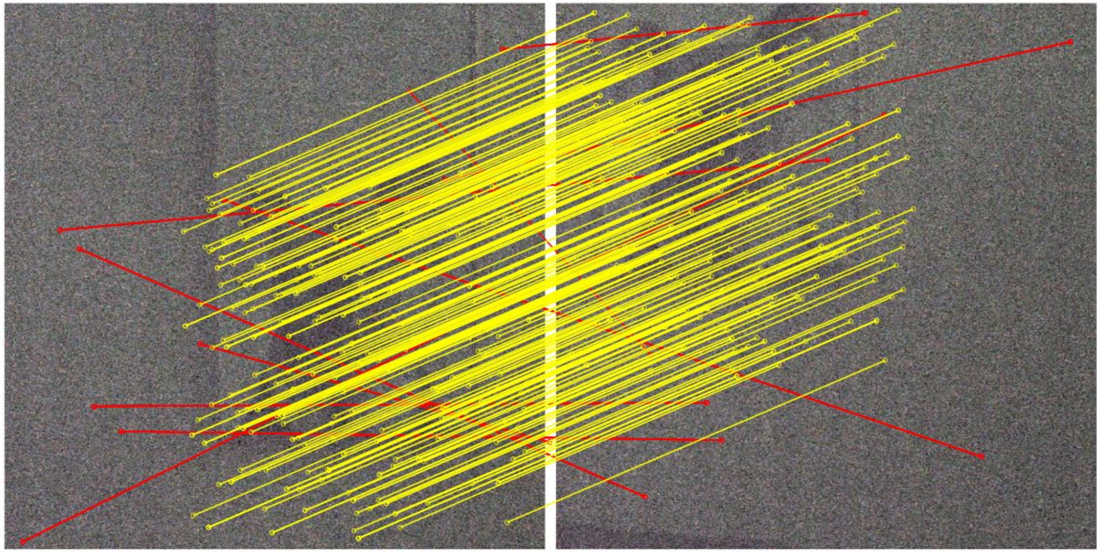
Moderate compute / bandwidth
Large signal gains

Hand-Crafted Burst Feature

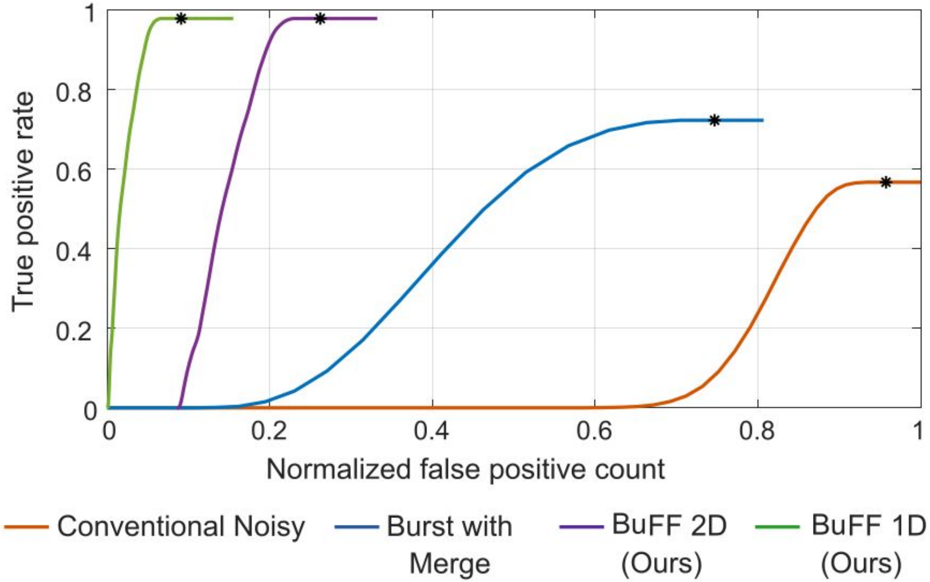


Hand-Crafted Burst Feature



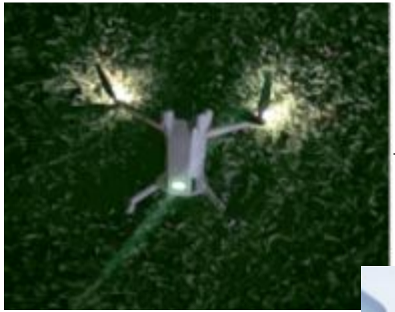


red: SIFT
yellow: ours



Drone Flight by Moonlight: mLux conditions

Images at 0.05 - 0.12 lux

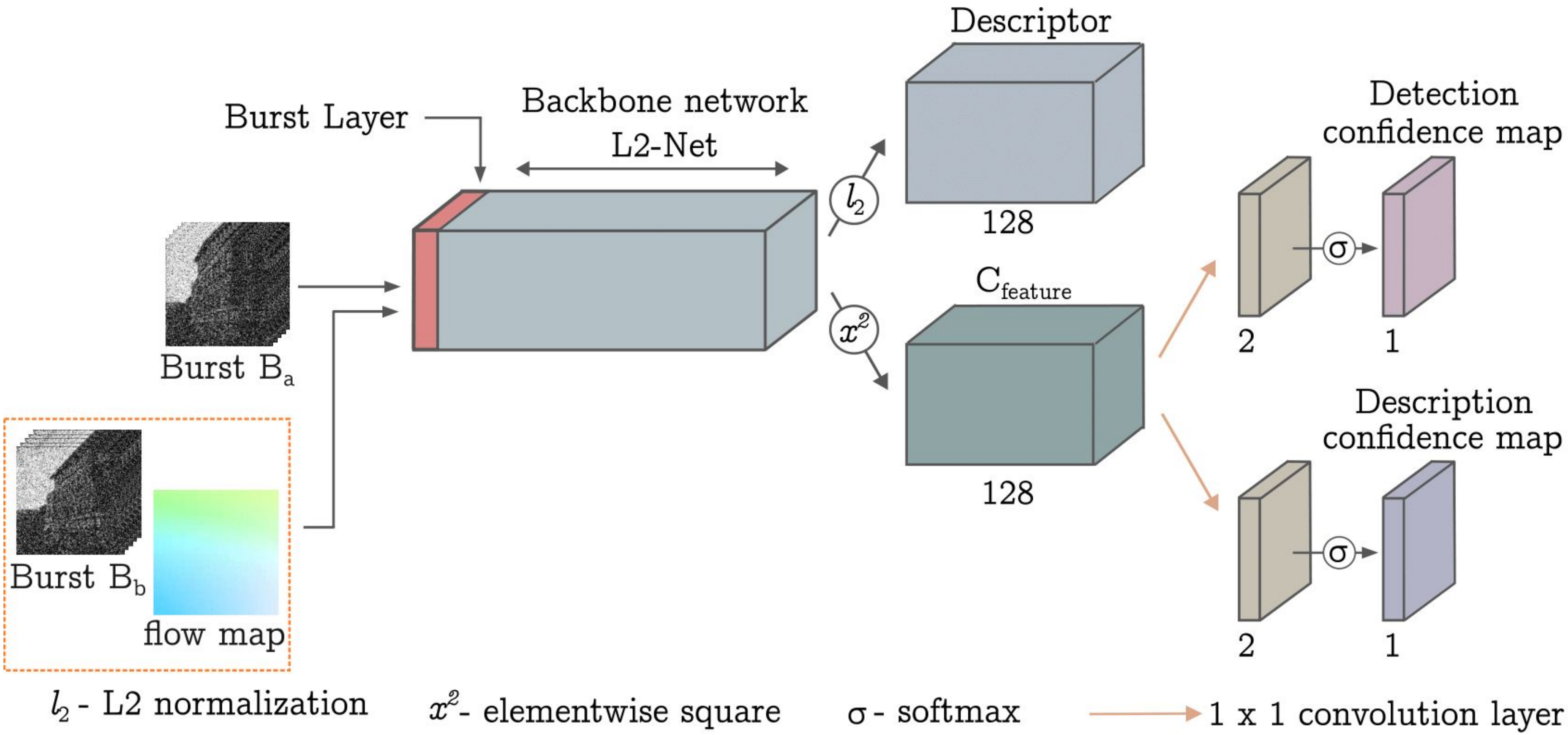


DJI Mini 3 Pro

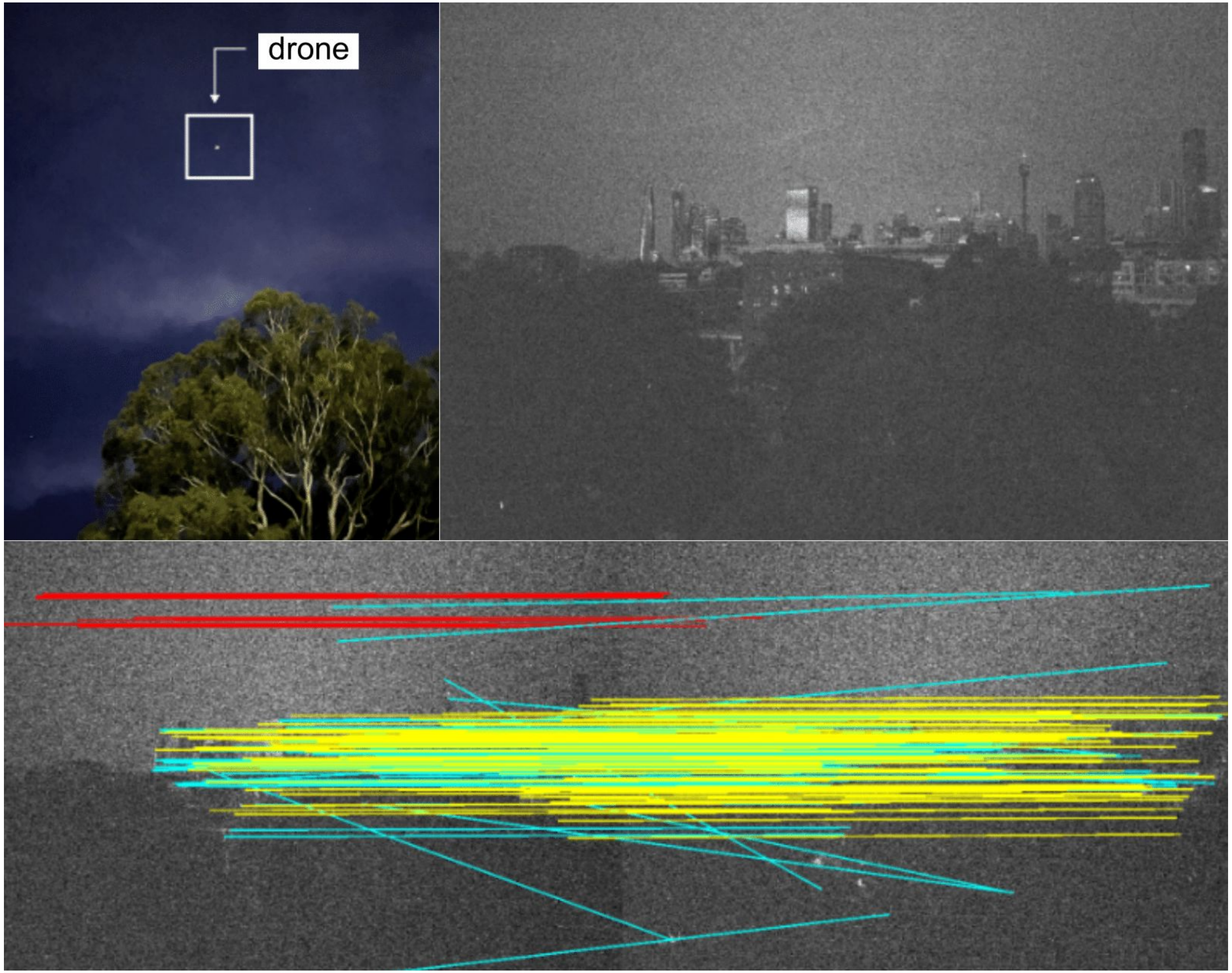
Images at 0.02 - 0.08 lux

DJI Phantom 4

Learning-Based Burst Feature

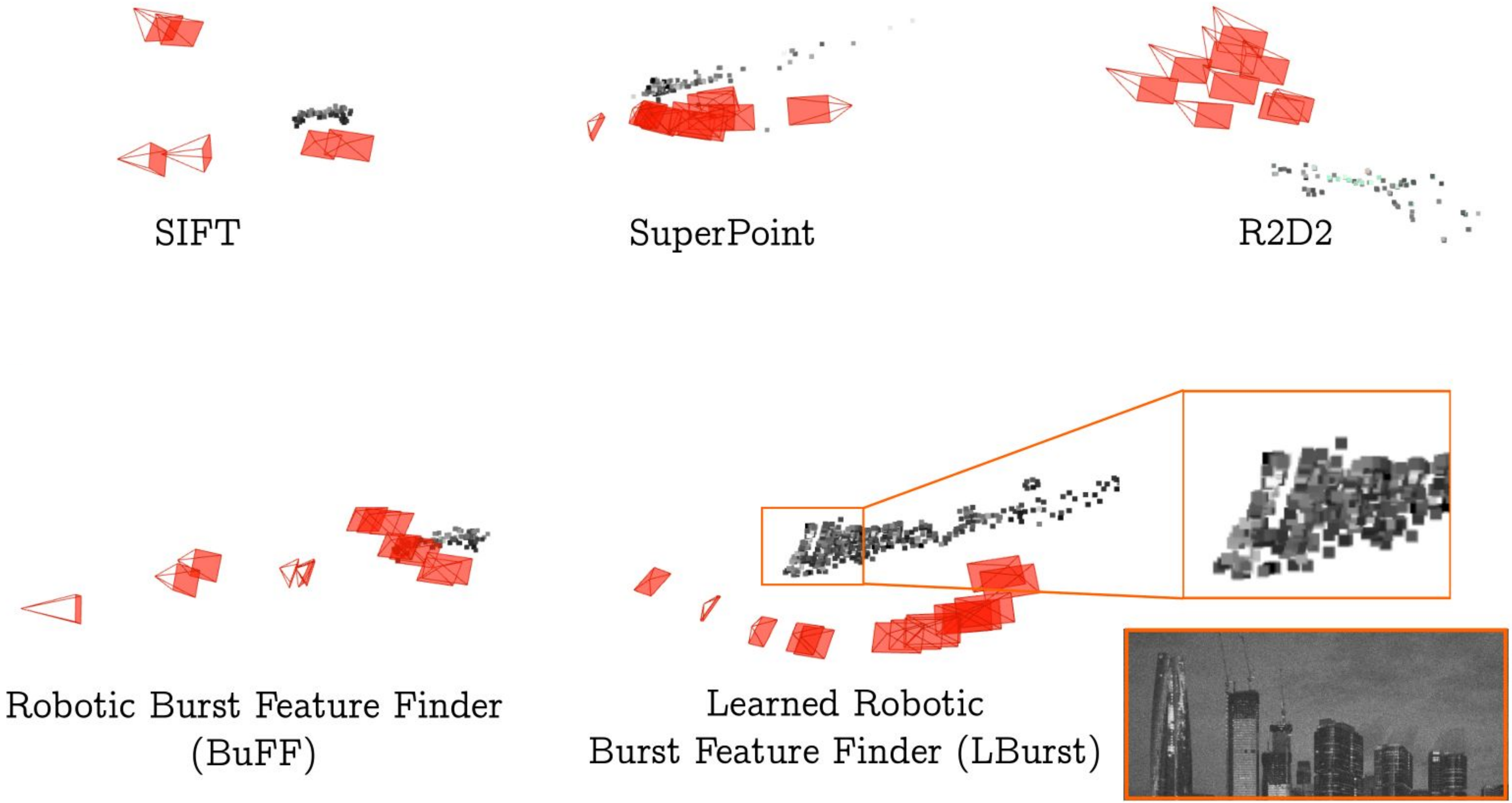


Learning-Based Burst Feature



blue: SIFT
red: R2D2
yellow: ours

Reconstruction



Reconstruction Performance

Average Performance Over 5 Scenes Each with 20 Bursts using COLMAP (Schönberger, et al., 2016)

Apparent Motion	Method		Convergence Rate	% of Images Pass	Match Ratio	Match Score	Precision	3D Points/ Image
2D	Gold Standard		1.0	100.0	3.89	3.87	0.99	305.5
	Conventional Noisy	SIFT	0.0	0.0	0.00	0.00	0.86	0.0
		SuperPoint	0.0	0.0	0.01	0.00	0.00	0.0
		R2D2	0.0	0.0	**0.65**	0.10	0.15	0.0
	Burst with Merge	SIFT	0.8	77.0	0.18	0.17	0.95	32.2
		SuperPoint	0.4	19.0	0.24	0.07	0.24	10.8
		R2D2	0.4	41.0	0.62	0.10	0.16	4.0
	BuFF		**1.0**	98.0	0.26	0.26	**0.98**	47.2
	LBurst		**1.0**	**100.0**	0.44	**0.43**	0.97	**73.1**

Next Steps

Adaptive Sampling

Sensor Fusion / IMU

Direct RAW input

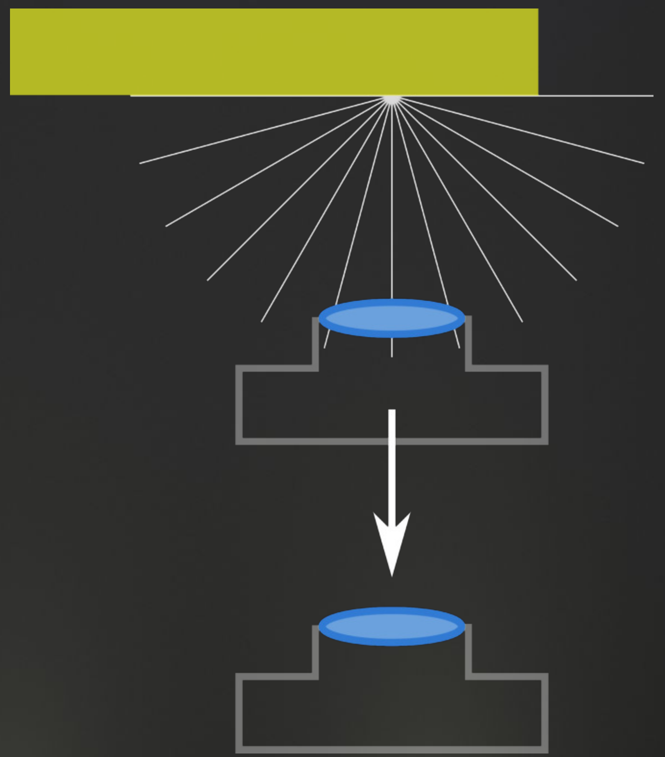
Imaging Physics

Beyond Burst

Why don't objects appear dimmer as we move away from them?

Hint: What have we missed about the scene?

$$I \propto \frac{S}{\pi r^2}$$

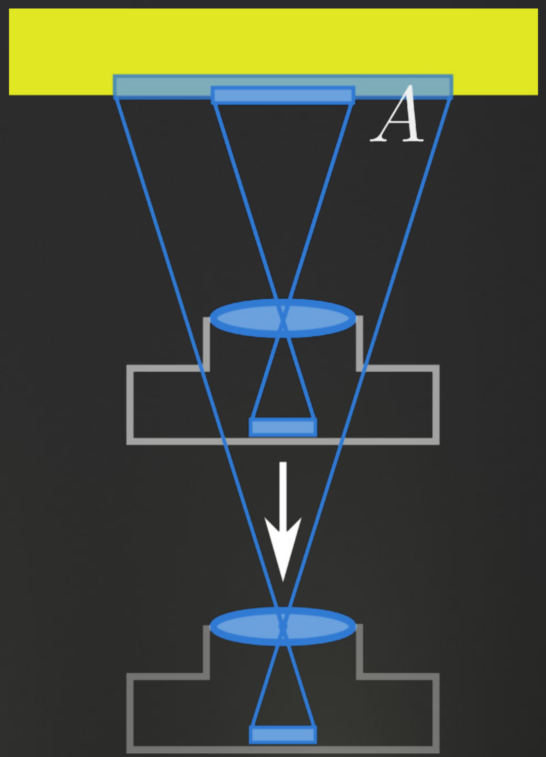


Why don't objects appear dimmer as we move away from them?

Hint: What have we missed about the scene?

The camera!
Projected pixel area grows with r^2
The r^2 cancel
→ Constant apparent intensity

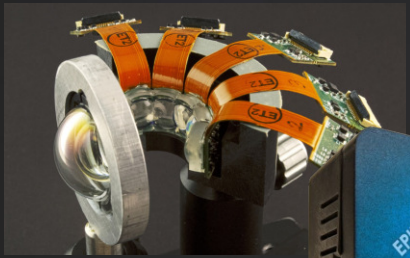
$$I \propto \frac{S}{\pi r^2}$$

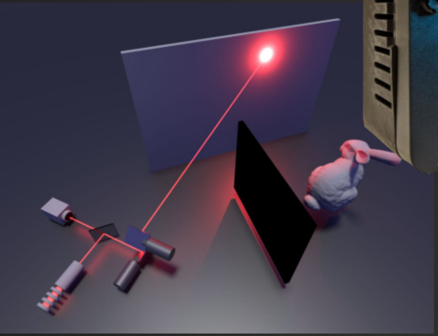


Today's Talk

1. Designing a camera for robots
2. Automating integration of new cameras
3. Making the most of existing cameras
4. **Emerging challenges**

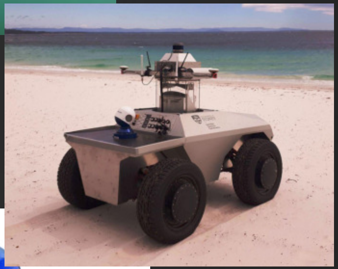
Representations for Eyes, Brains, and Bodies



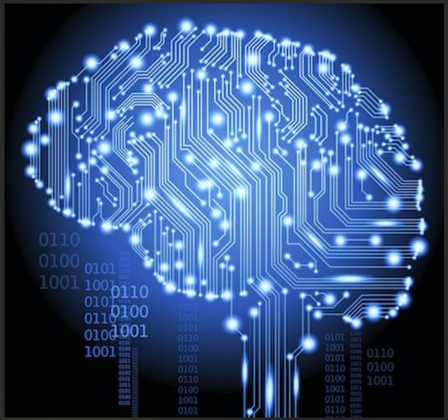


Light:
NeRFs,
Gaussian Splats,
LFs

Geometry,
Semantics,
Uncertainty

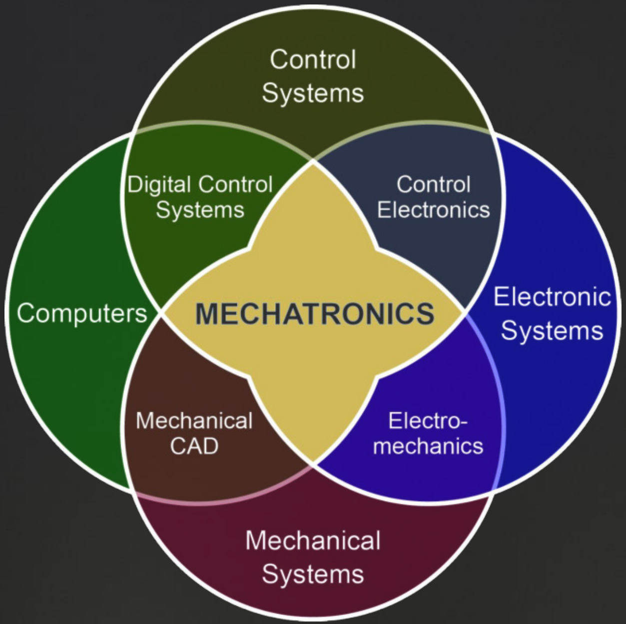


Language, Vision,
Action & Interaction



Automating Perception System Design...

Co-designing optics, sensing, embodiments, and algorithms
to let robots work in new ways



+

Robotic Imaging Lab

Dr Don Dansereau

Bina Rajan

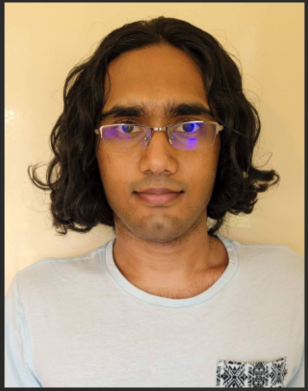
Jack Naylor

Ryan Griffiths

Nikolai Goncharov

Oliver Yan

Alex Cardaillac



Bhargava Gowda

Dr Ahalya Ravendran

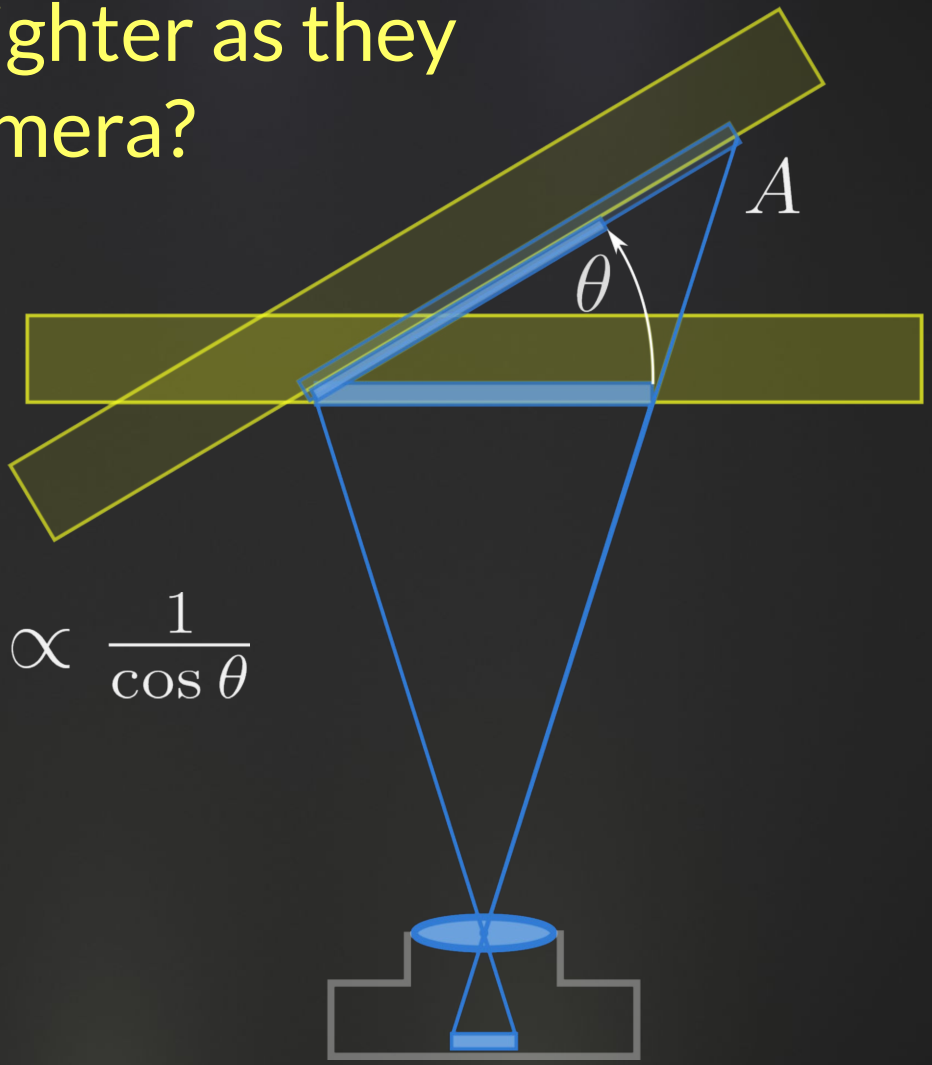
Dr Teja Digumarti

Why don't surfaces appear brighter as they turn away from the camera?

Should we expect them to?

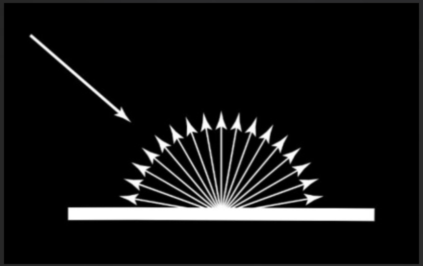
Yes, more angle = more projected pixel area
= more energy into the lens

$$I \propto A \propto \frac{1}{\cos\theta}$$

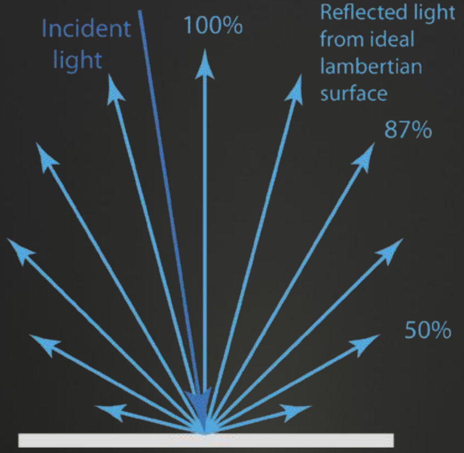


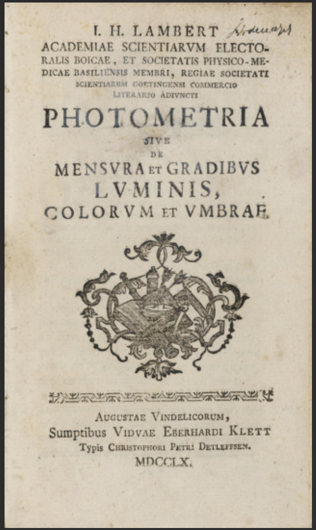
Why don't surfaces appear brighter as they turn away from the camera?

Hint: this is useful but wrong (or at least misleading) :



Lambert's cosine law:





I. H. LAMBERT
ACADEMIAE SCIENTIARVM ELECTORALIS BOICAE, ET SOCIETATIS PHYSICO-MEDICAE BASILIENSIS MEMBRI, REGIAE SOCIETATI SCIENTIARUM GOETINGENSI COMMERCIO LITERARIO ADIVNCTI

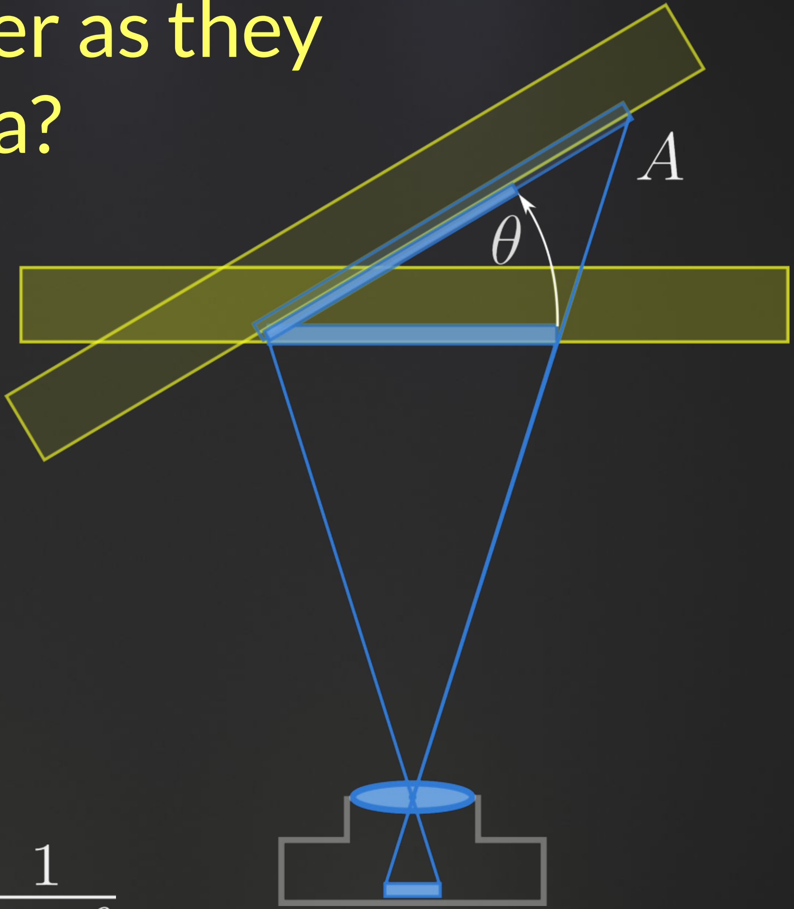
PHOTOMETRIA
SIVE
DE
MENSVRA ET GRADIBVS
LVMINIS,
COLORVM ET VMBRAE.

Augustae Vindelicorum,
Sumptibus Viduae Eberhardi Klett
Typis Christophori Petri Detleffsen.
MDCCLX.

[lambert1760]

$$I \propto A \propto \frac{1}{\cos\theta}$$

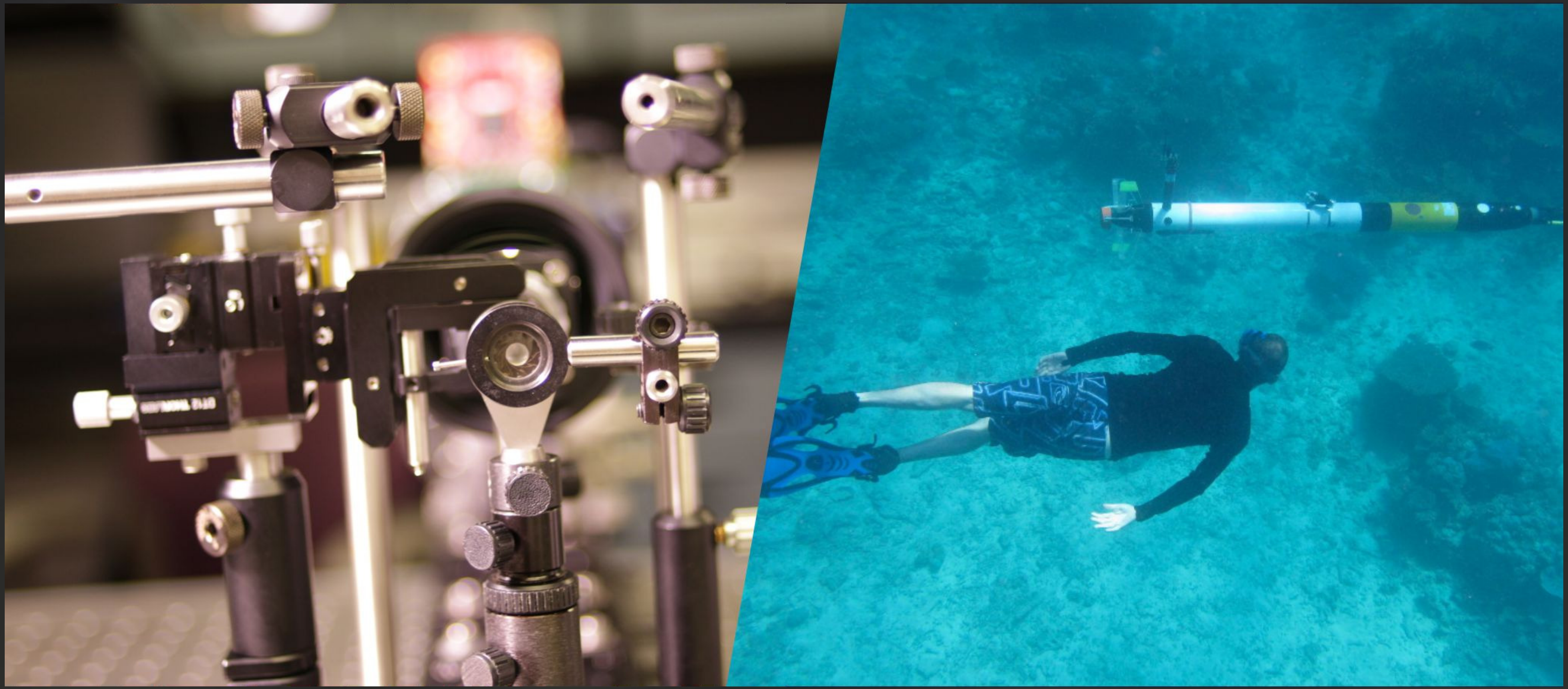
$$I \propto \cos\theta$$



To Ponder:
Is our moon Lambertian? How would you prove it?

Robotic Imaging Lab

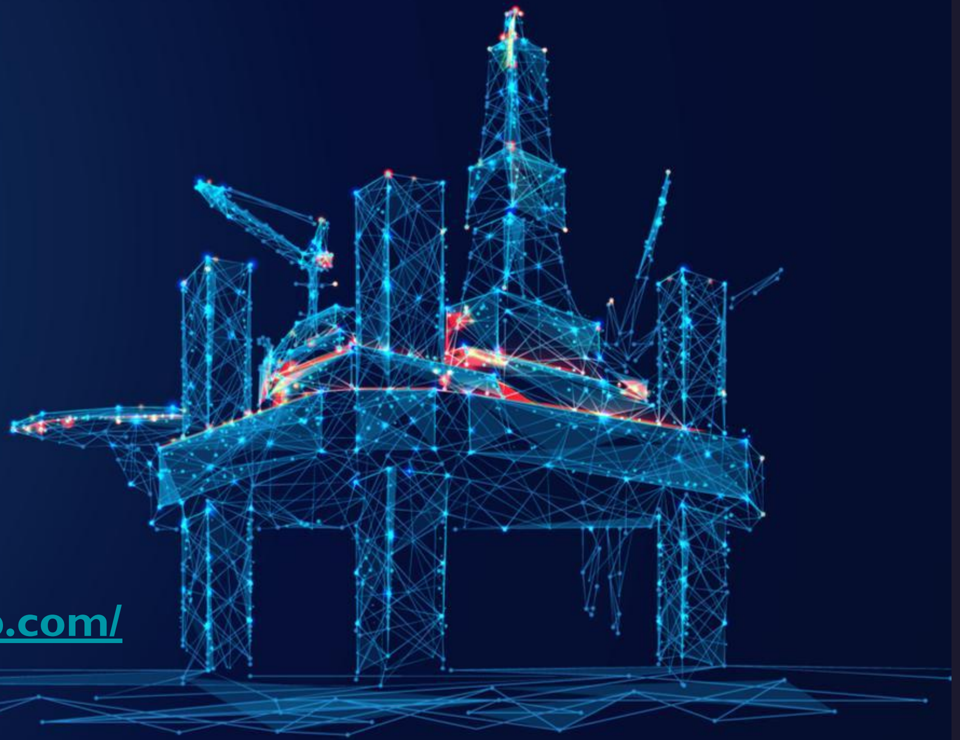
RoboticImaging.org



Australian Robotic Inspection & Asset Management Hub (ARIAM)

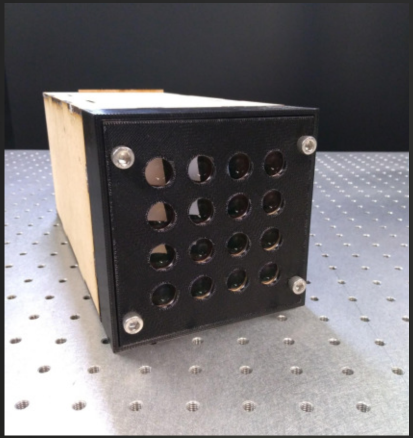
USyd, QUT, ANU   +10 industry partners
5-year ARC Research Hub
Fundamental Problems in Sensing, Perception & Control

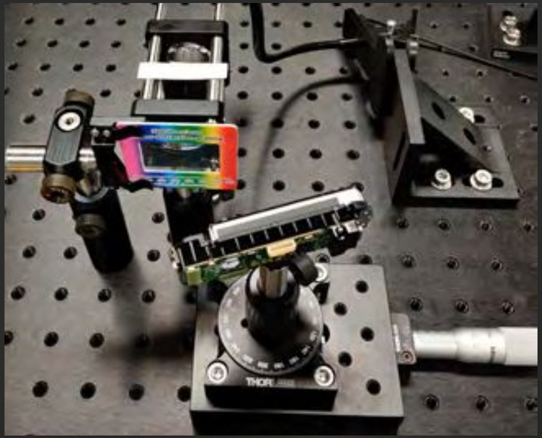
Autonomy to Manage Infrastructure & Assets
Underwater, Ground, Air

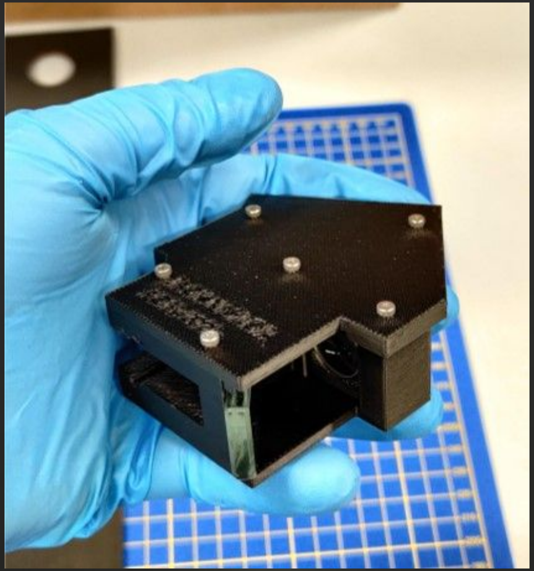


Multiple PhD & Postdoc Roles Open   https://ariamhub.com/

Robotic Imaging Lab @ ACFR



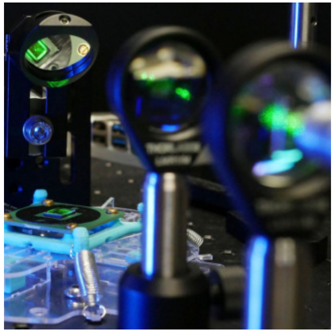




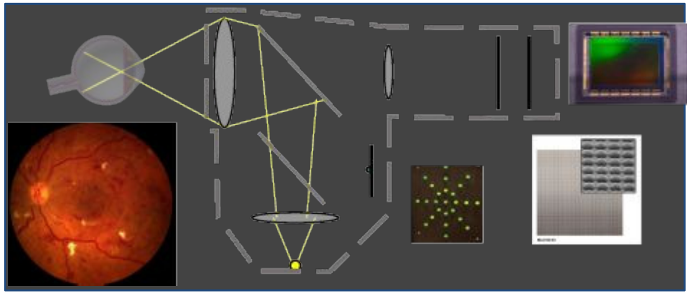
New Cameras, Representations & Algorithms...

Spherical-lens light field (LF)
w/UCSD, Stanford

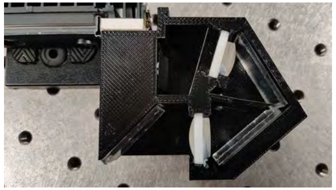
Sparse LF
w/EPIImaging LLC

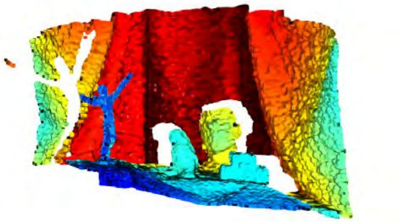


Open-source LF adapters,
Light Field Toolbox

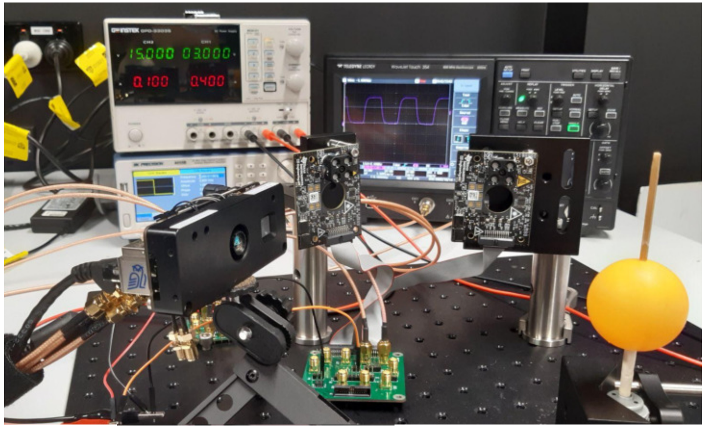


Medical Fundus LF imager
w/QUT, Spun out as Integral Scopes





Hyperspectral 3D Adapter



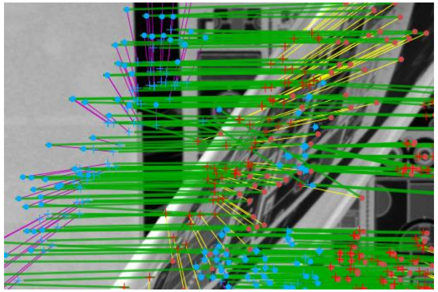
Emerging technologies:
Doppler, Single-photon, Vision-on-chip

… to Make Machines Smarter

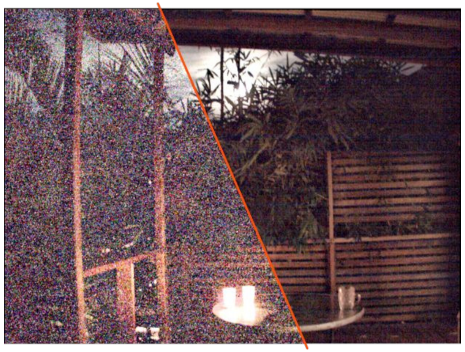
Underwater Robotics

Vision for the Blind

All-Weather Autonomous Driving

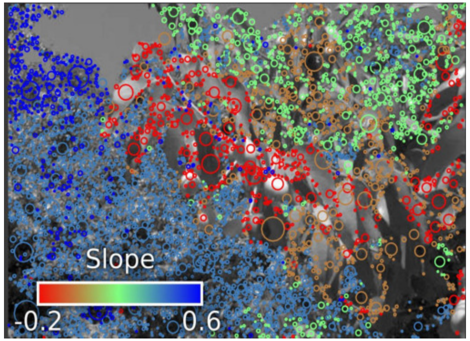


Automation around Transparency



Night-time Operation

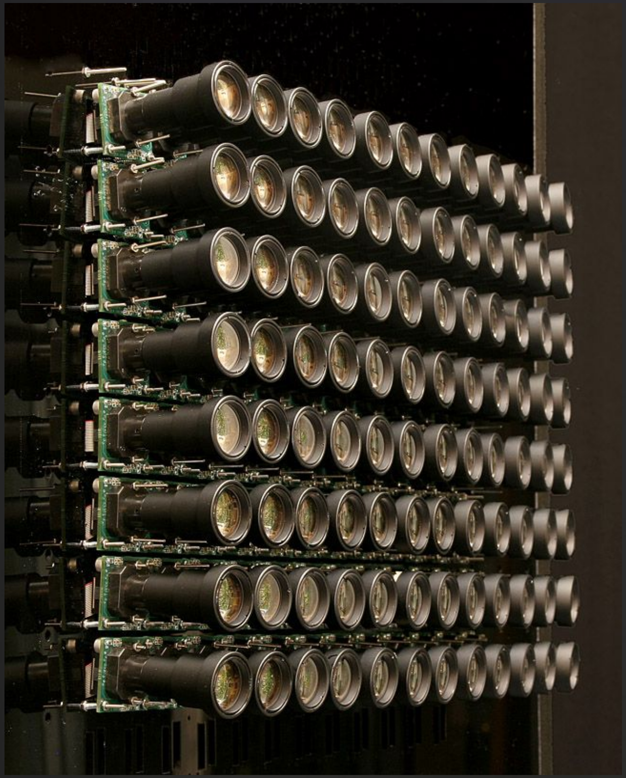
VR, AR, MR



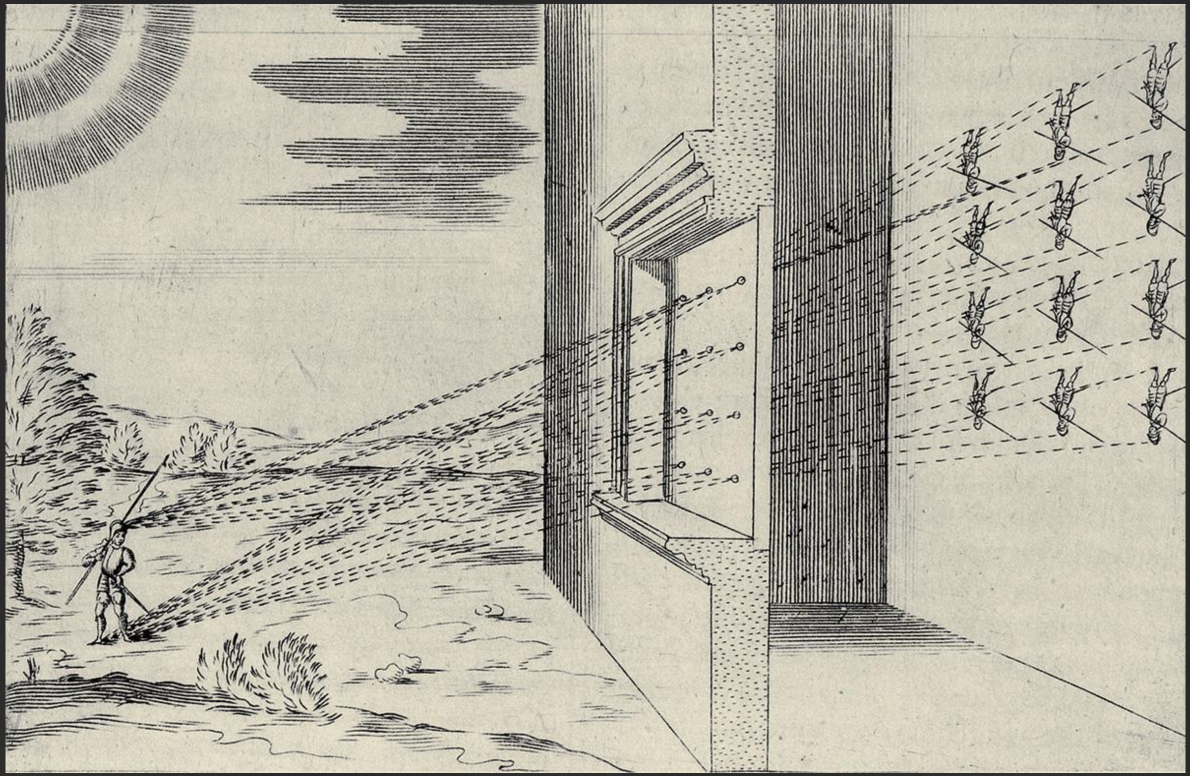
Natural Environments

Privacy-Enhanced Robotic Vision

Light Field Cameras

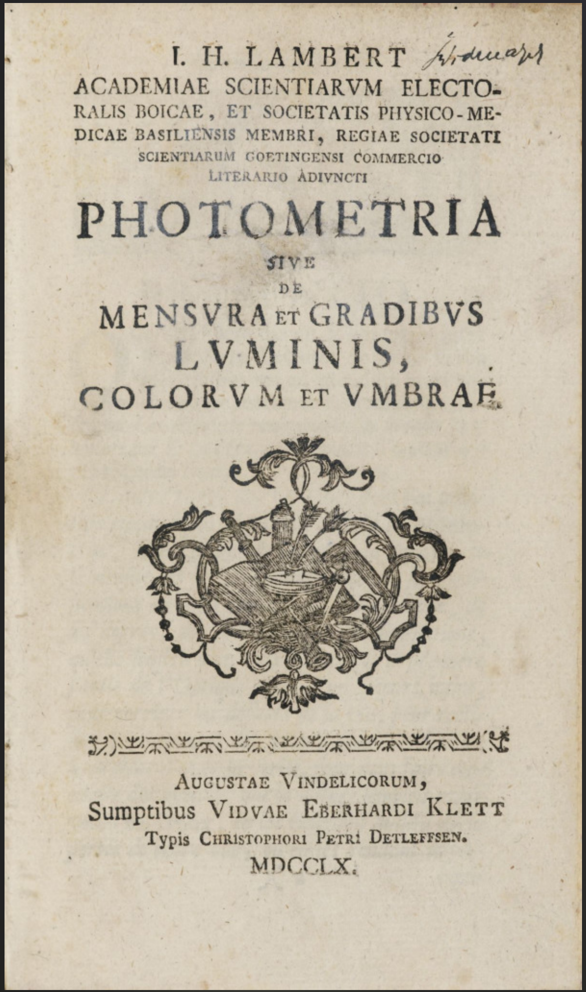


Stanford camera array
2005



Bettini 1642

Thinking About Light



I. H. LAMBERT
ACADEMIAE SCIENTIARVM ELECTO-
RALIS BOICAE, ET SOCIETATIS PHYSICO-ME-
DICAE BASILIENSIS MEMBRI, REGIAE SOCIETATI
SCIENTIARUM GOETINGENSI COMMERCIO
LITERARIO ADIVNCTI

PHOTOMETRIA
SIVE
DE
MENSVRA ET GRADIBVS
LVMINIS,
COLORVM ET VMBRAE.

AUGUSTAE VINDELICORUM,
Sumptibus VIDVAE EBERHARDI KLETT
Typis CHRISTOPHORI PETRI DETLEFFSEN.
MDCCLX.

```
@book{ lambert1760photometria,

     title = {Photometria, sive de Mensura et
     gradibus luminis, colorum et umbrae},

     author = {Lambert, Jean-Henri},

     year = {1760},

     publisher = {Eberhard Klett}

}
```

“Photometry, or the measure and gradation of light, colour, and shadow”