Tackling Extreme Visual Conditions for Autonomous UAVs In the Wild

Hayk Martiros
Head of Autonomy, Skydio
September 2020

Skydio 2

Q4 2019 Launch

Consumer video drone

Built for autonomous flight

Six navigation cameras with 200 deg FOV

User camera w/ 3-axis gimbal

NVIDIA Tegra X2 chip-down design

$999

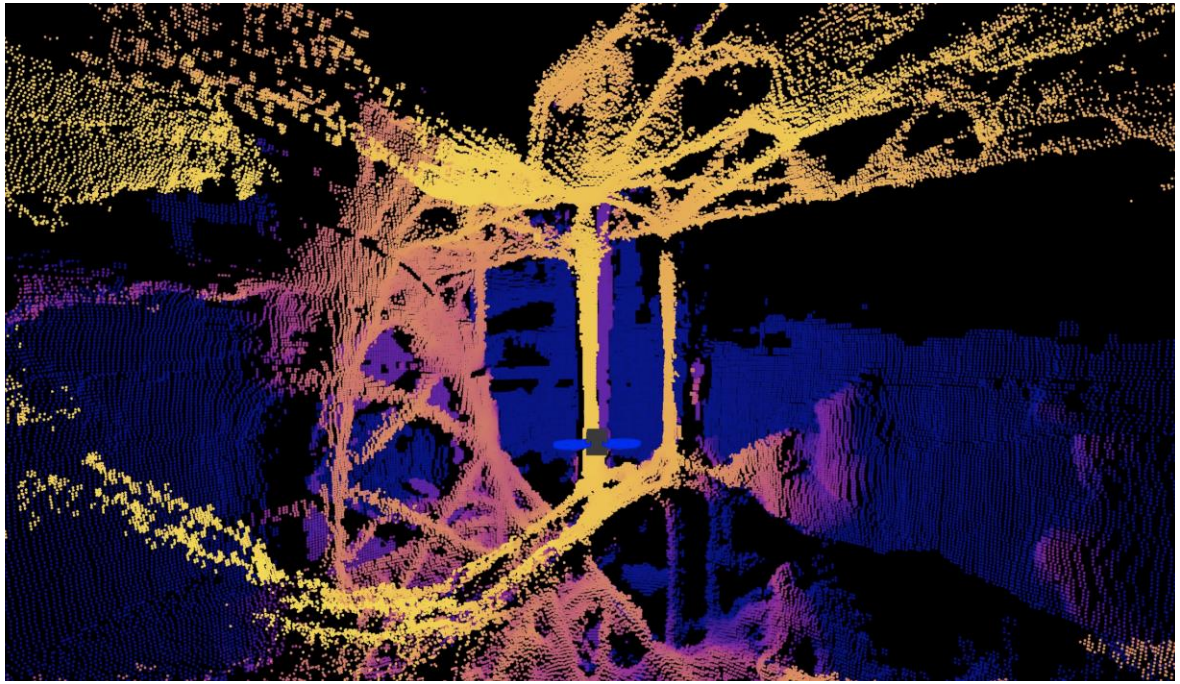
First autonomous robot warrantee of its kind:

If it crashes within our operating guidelines, we'll replace it.

Skydio R1

Q1 2018

Skydio 2



"Skydio 2 is one of the most impressive robots that I have ever seen. . . . This drone has left me feeling like its fundamental autonomous capability is so far beyond just about anything that I've ever experienced that I'm questioning why I would ever fly anything else ever again. . . . When you see this technology in action, it's (almost) indistinguishable from magic." - **IEEESPECTRUM**

"Skydio 2 never crashed. It didn't even come close. No matter how dense the foliage, no matter how many branches stood in its way, it always did what was necessary to avoid a collision. This is a huge confidence booster, and puts your mind at ease during filming sessions." - **Digital Trends**

Skydio

Skydio X2

COMING Q4 2020

INTRODUCING

SKYDIO X2™

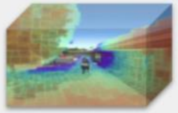
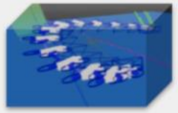
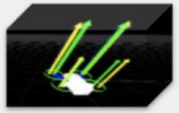
Unmatched autonomy meets enterprise performance.

- **Field-tested AI** building upon Skydio 2's groundbreaking technology foundation
- **Ultimate data capture** via 12MP color camera and 320x256 thermal camera
- **Fly anytime, anywhere** with GPS night flight capability, 35+min endurance, 6 km range, and rucksack portability
- **Enterprise (X2E) and Defense (X2D)** configurations available

Skydio

Skydio Autonomy Engine

Skydio Autonomy Engine

Criteria	Skydio Advantage	Why should you care?
Real-time 3D Mapping	 *>1 million point / second model refresh rate*	Sees everything
Object Recognition	 *9 On-board deep learning networks*	Context-aware flight paths
360° Obstacle Avoidance	 *Full 360 detect and avoid*	Avoids crashes
Motion Prediction	 *500 iteration/sec model to action update rate*	Navigates itself, freeing the pilot
Advanced AI Pilot Assistance	 *360 Superzoom, 180 Vertical View and Precision Mode*	More effective piloted flight
Workflow automation	 *Range of autonomous work flows for specific use cases*	One pilot to many drone operations

Extreme Visual Conditions

The challenge of an autonomous drone: Customer pays for the robot to fly itself in a place we've never seen and quickly gains unconditional trust. Every limit is pushed.

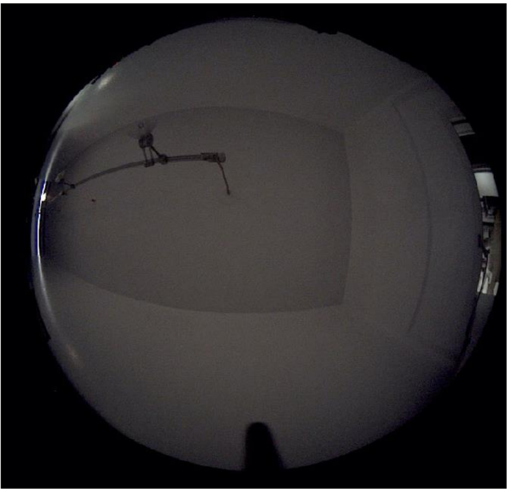
- Few semantic priors - fly inside a tree, in a factory, at Burning Man, on a mountain
- No room for missing obstacles or failing state estimation, robot will crash or fly away
- No room for false positives, drone will be erratic and unpredictable
- Time is ticking! Limited battery life, limited connectivity, high speed motion

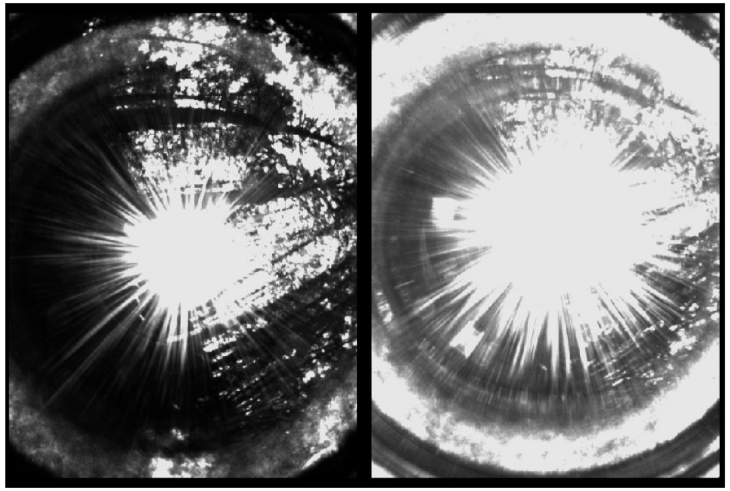
High-speed Follow + Film Scenario

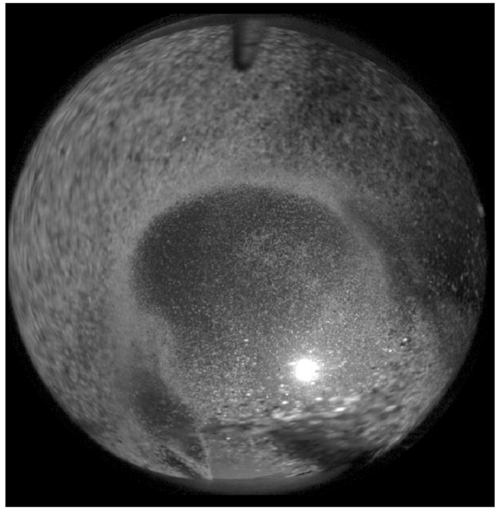
Extreme Visual Conditions

Daunting set of challenges to visual navigation:

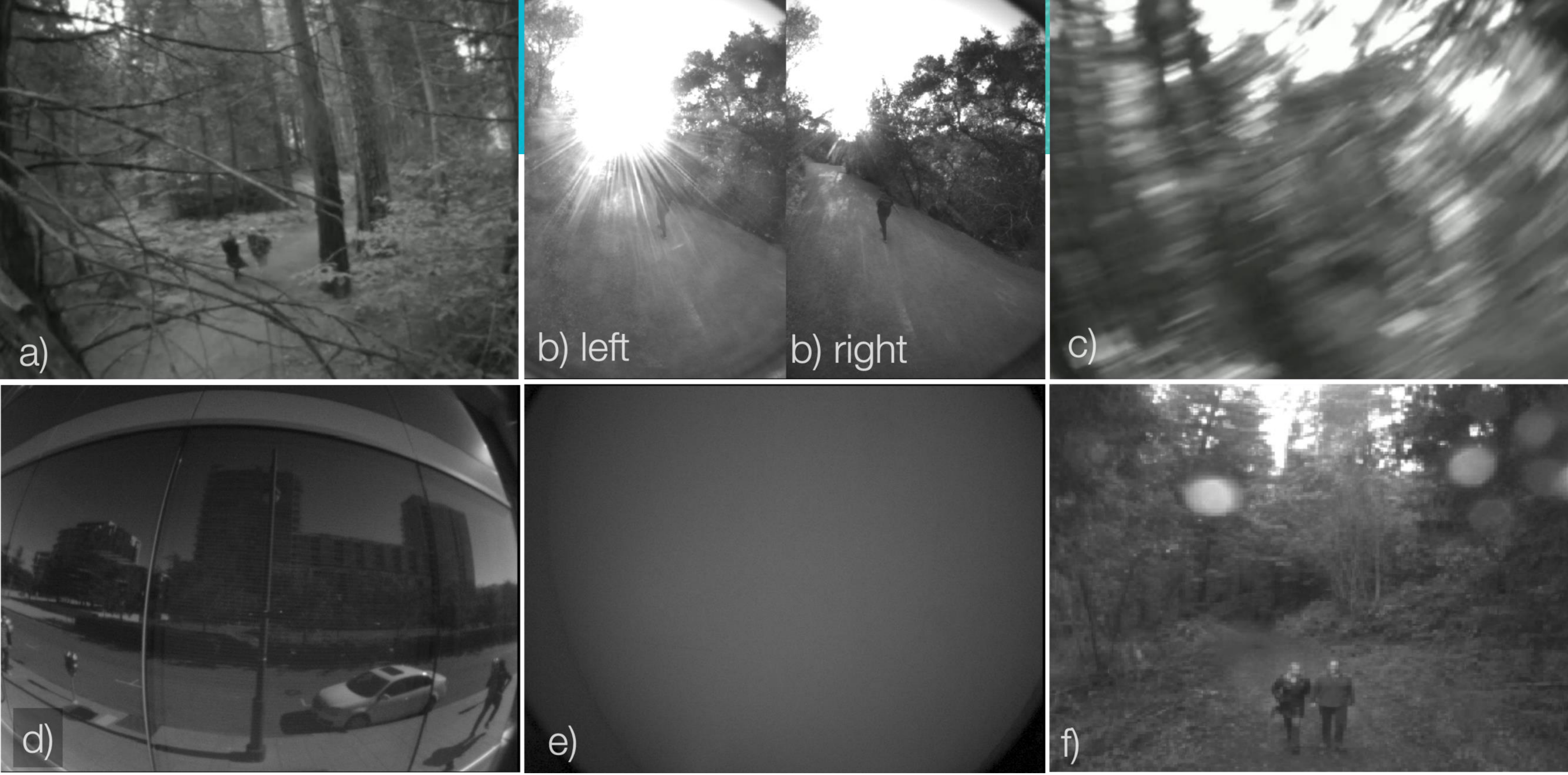
- Thin objects, extreme glare, shadows, camera artifacts, motion blur, dirt, smudges
- Waves, snow, rain, fog, desert dunes, salt flats, reflective surfaces, textureless walls/ceilings
- Calibration errors, rolling shutter errors, time-sync errors, vibrations, electrical interference







Skydio



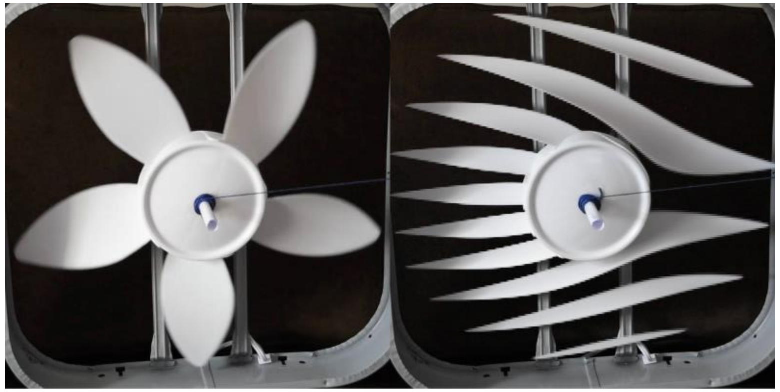
a) thin branches, b) sun glare with dirty lens, c) severe motion blur,
d) reflections, e) textureless sky or ceiling, f) water droplets

Rolling Shutter

Problem: Skydio 2 uses rolling shutter navigation cameras (higher quality, lighter, cheaper). Need to do geometric processing during high speed, high rotation motions close to objects!

Solution

- Careful modeling of rolling shutter effects, accounting for translation and rotation
- Accurate time synchronization between camera triggers, exposure periods, IMUs



Global Shutter Rolling Shutter

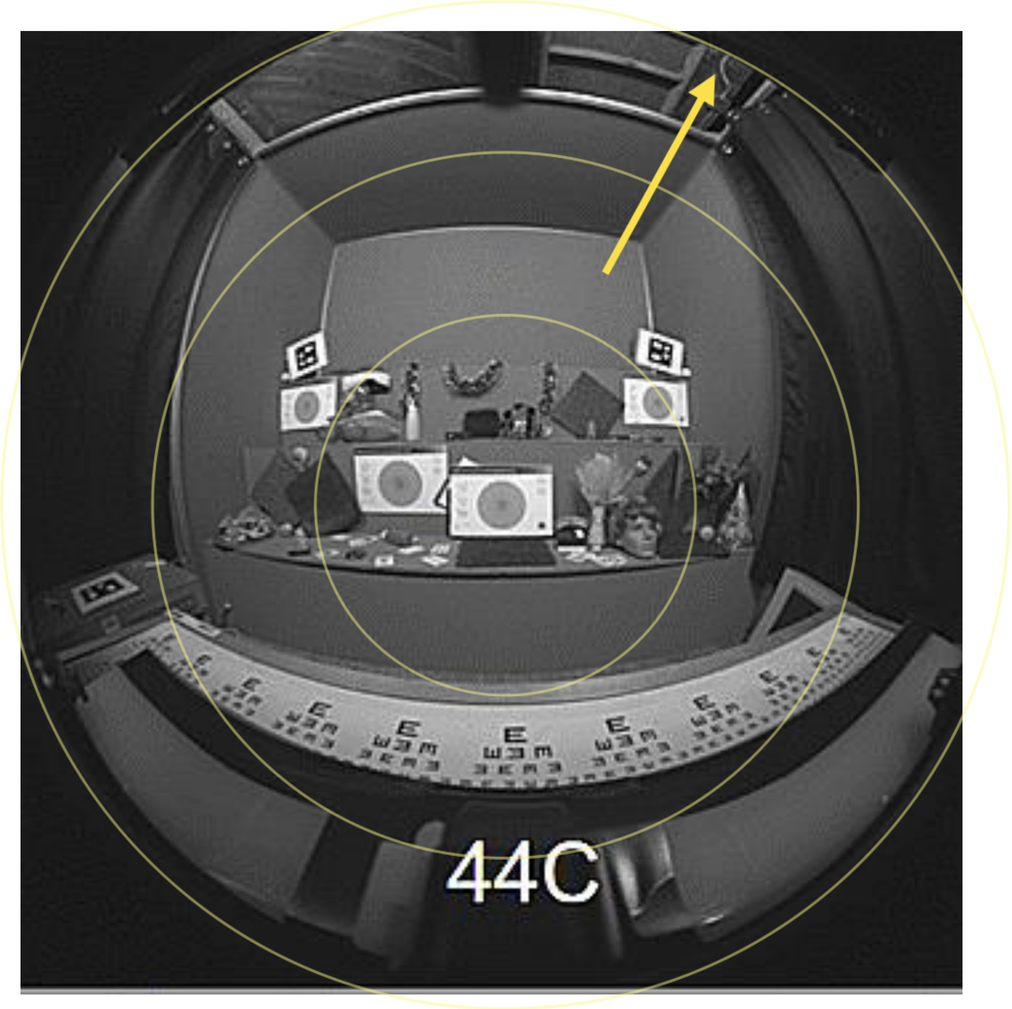
pixel shift through exposure period

Skydio

Lens Intrinsics

Problem: Environmental conditions significantly change intrinsic lens properties. Need to estimate online to maintain high quality mapping.

Solution: Visual-inertial odometry system jointly estimates intrinsics, extrinsics, IMU biases, and world features during flight, accounting for rolling shutter.



stationary lens in temperature chamber

Moving Objects

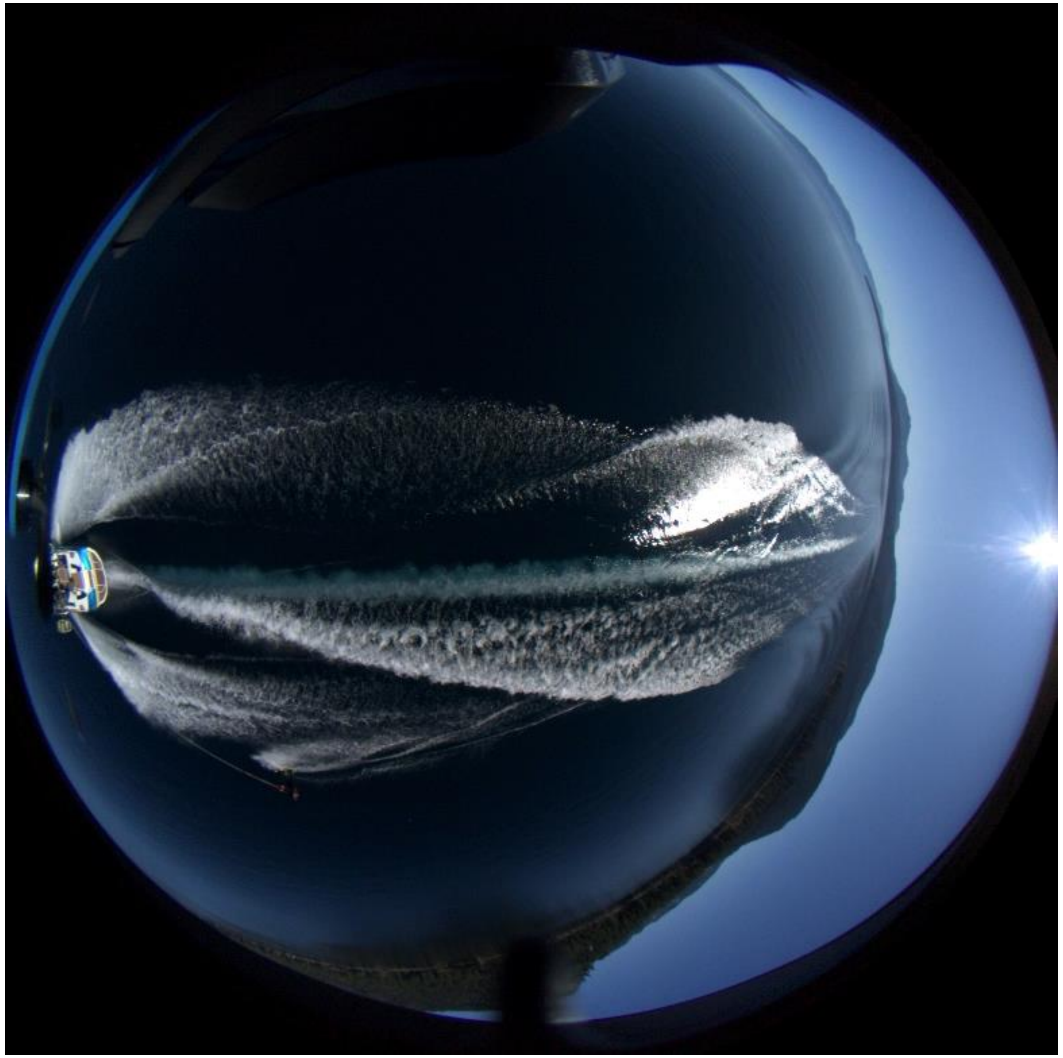
Problem: All of the close visual features in this scene are moving and malicious for state estimation purposes.

Solutions:

+ Careful feature gating relative to IMU

+ Joint consideration of visual and GPS uncertainty

+ Semantic information

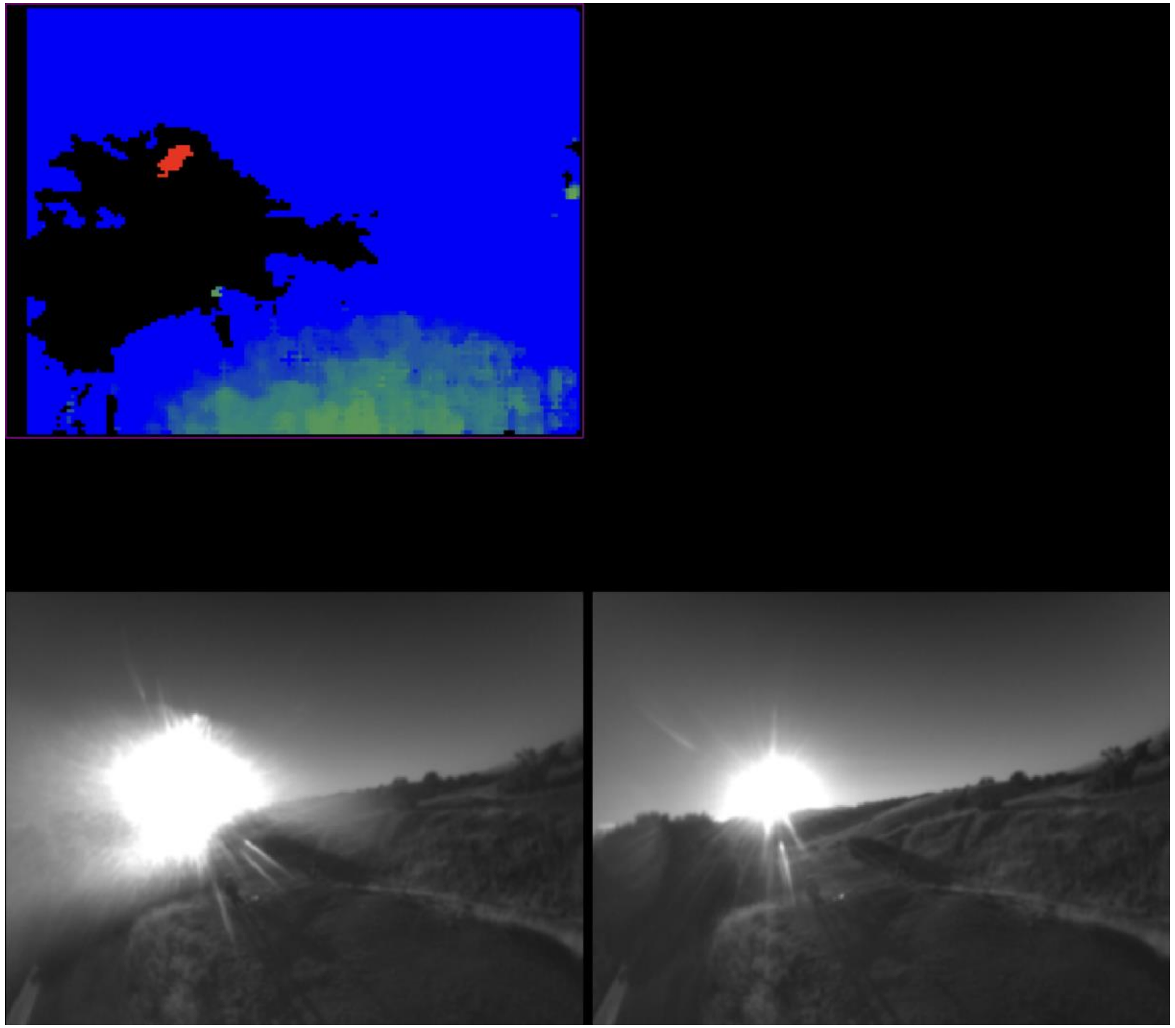


Skydio

Real World Problems: Dirty Camera Detection

Problem: Dirt, dust, fingerprints, water on the lenses can ruin photometric consistency and cause false matches.

Solution: Estimate dirty regions of cameras in flight.



Real World Problems: Dirty Camera Detection

Idea: Regions with poor photometric consistency over a variety of background content are likely from dirty cameras.

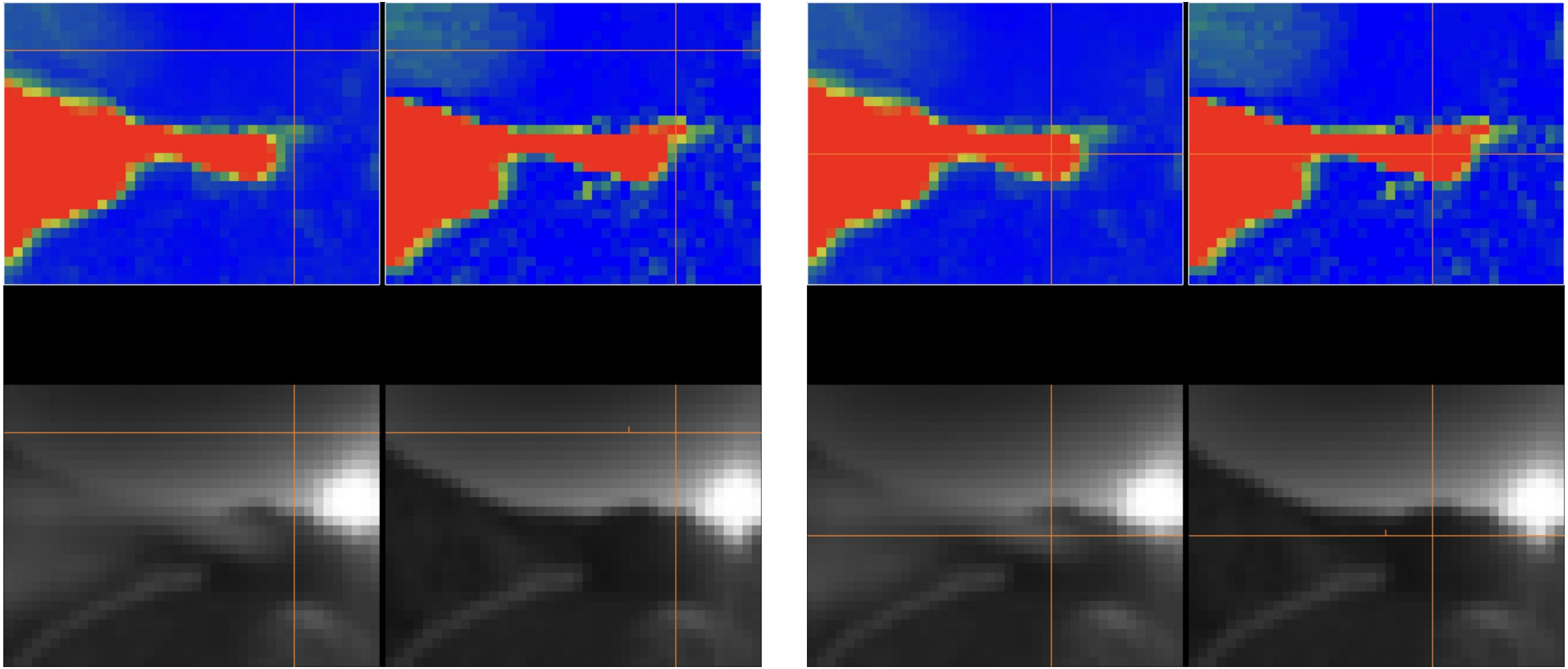
Aggregate from our matching algorithms across time and vehicle rotations, use as an invalid mask.

Continue flying, applying the mask, but warn the user to land and clean lenses!

If the mask is a large portion of a camera, go into a high-level error state and tell the pilot (if possible) to land quickly.

Skydio

Real World Problems: Dirty Camera Detection



Challenges for Learning Depth

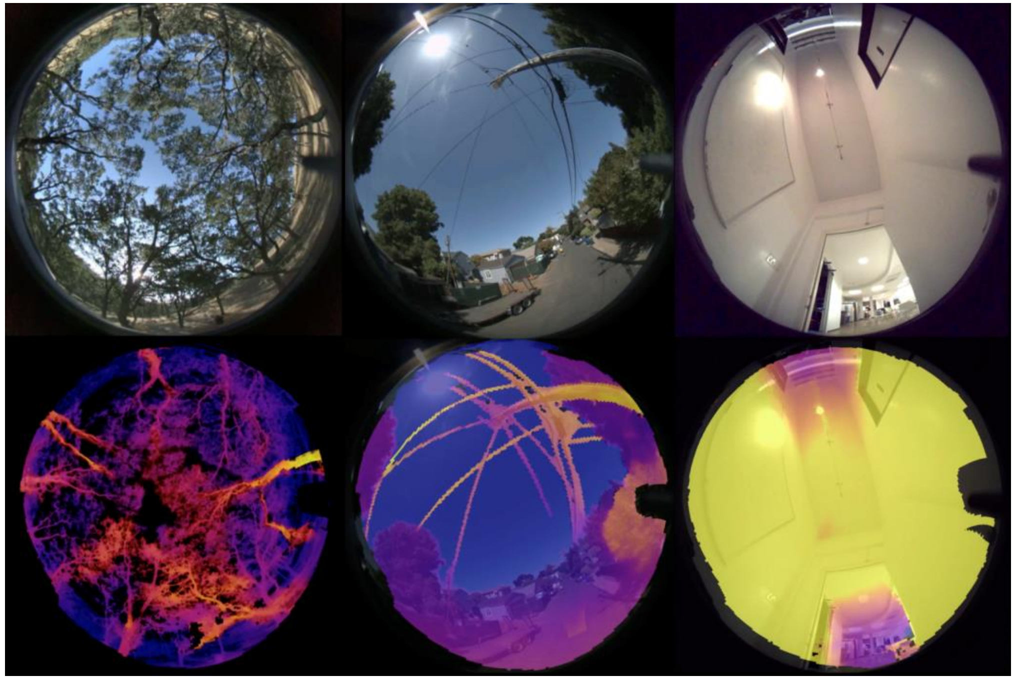
No ground truth - infeasible to get human labels

Billions of images captured by S2 to learn from —> unsupervised learning!

- SOTA has not made this work yet. Photometric consistency fails.

Need to find ways to use leverage large amounts of in-domain unlabeled data for extreme visual conditions.

More acausal reasoning? Better features? Cycle consistency?



Synthetic Generation

Synthetic augmentations of real data allows forms of supervision without knowing ground-truth.

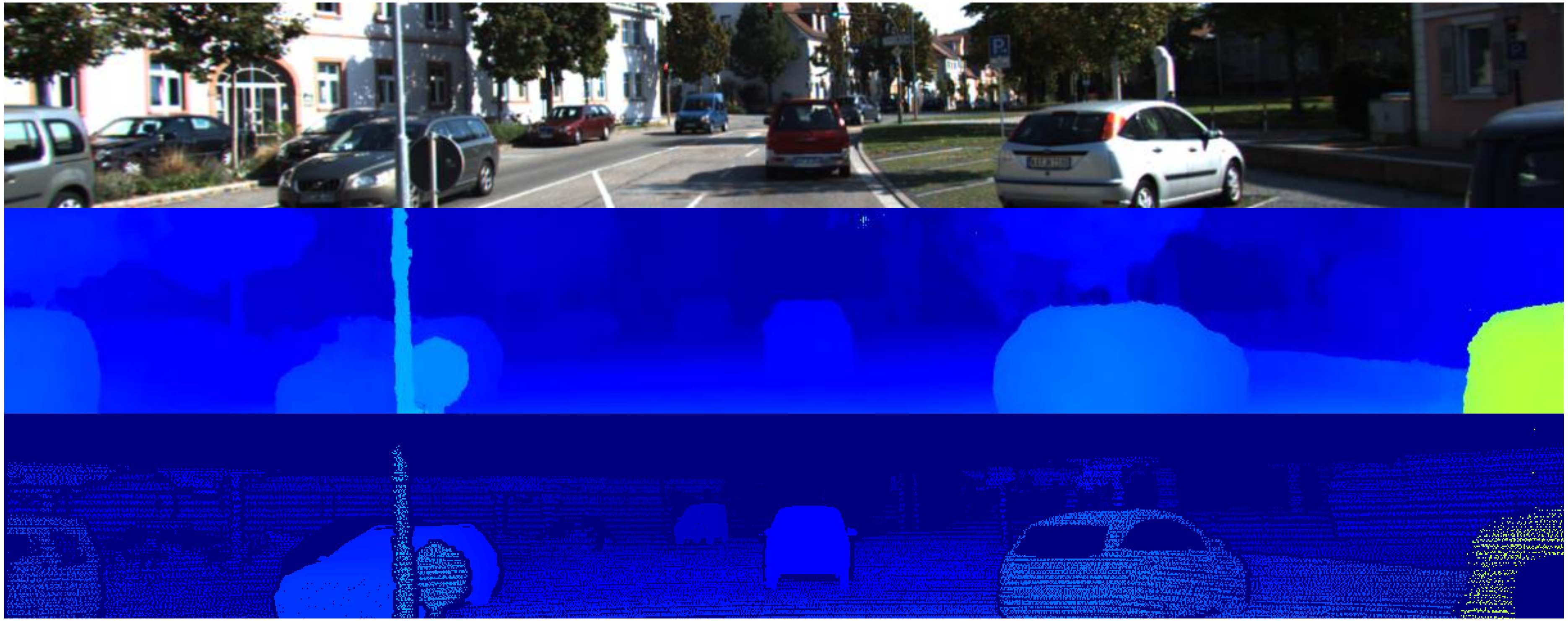
Example: Generate suns on real data, penalize prediction errors between original and augmented images (or expect uncertain prediction).

a) Examples of synthetic sun glare generated by simulating diffraction of light through lens imperfections. b) Real sun (left) and augmented sun (right) on a real image.

Deep-Learned Depth Estimation

End-to-End Learning of Geometry and Context for Deep Stereo Regression
Kendall et al. ICCV 2017
KITTI 2012: #1 KITTI 2015: #2

~1 fps on a Titan X GPU



Deep-Learned Depth Estimation

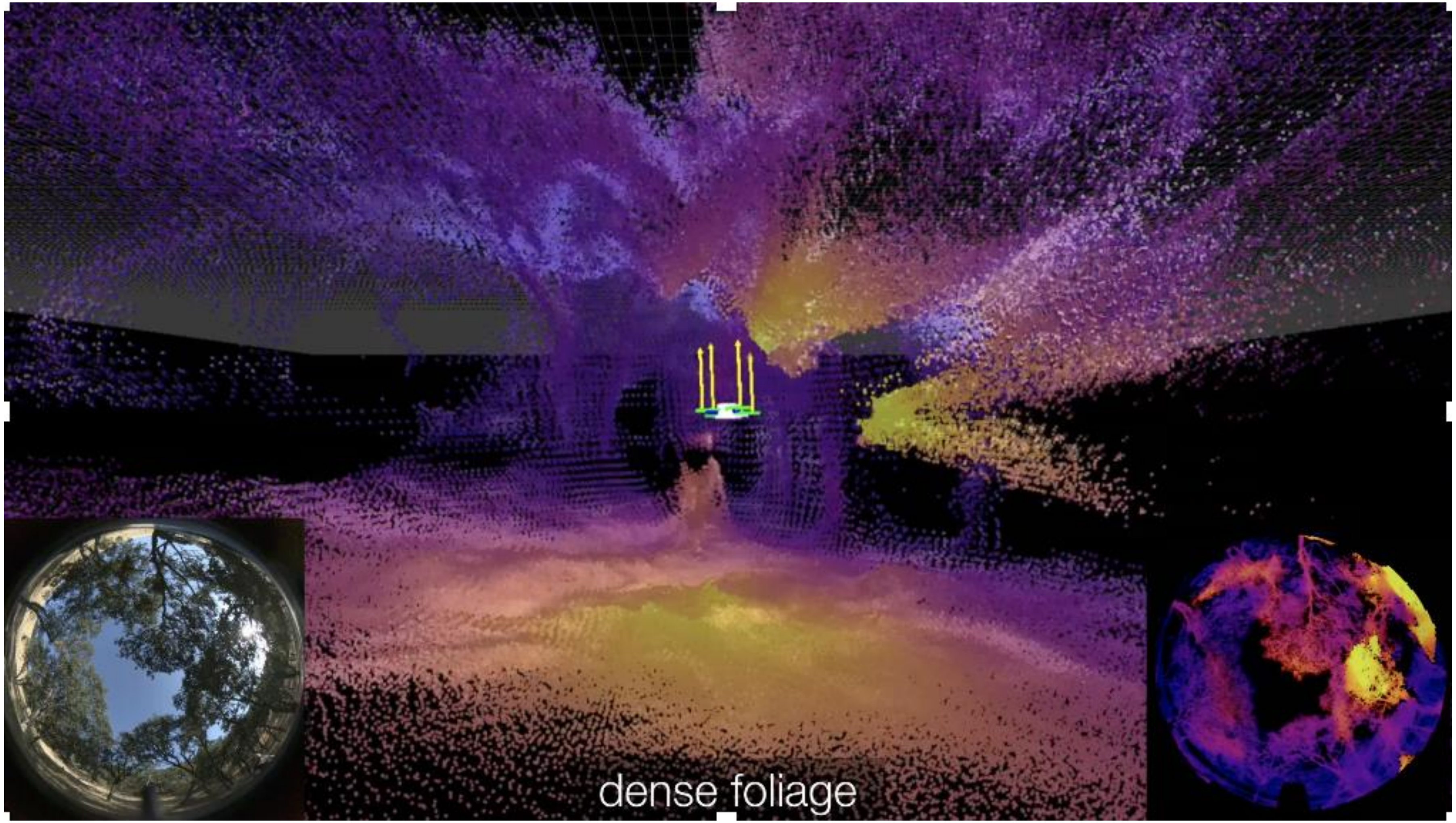
First demonstration of complex autonomous flight using deep stereo vision

July 2018

~1000 fps on a Titan X GPU

a) replaying previous collision b) autonomous navigation with deep stereo vision

Deep-Learned Depth Estimation Results



Skydio

Intelligent Error Handing + Graceful Degradation

Comprehensive error handling is **extremely** difficult and requires complex integration of the whole stack. This what separates lab demos from real products – there is no substitute.

Dirty Cameras	Too Dark	Dangerous Environment	High Wind
Collision	Gap Too Small	CPU Too High	Gimbal Obstructed
Takeoff Obstructed	Battery Low	Lost Comms	Control Diverged

Conclusions

Autonomous robots, especially UAVs, encounter extreme visual conditions - handling them is the key to mainstream adoption. We are not close to the ceiling for this.

Requirements:

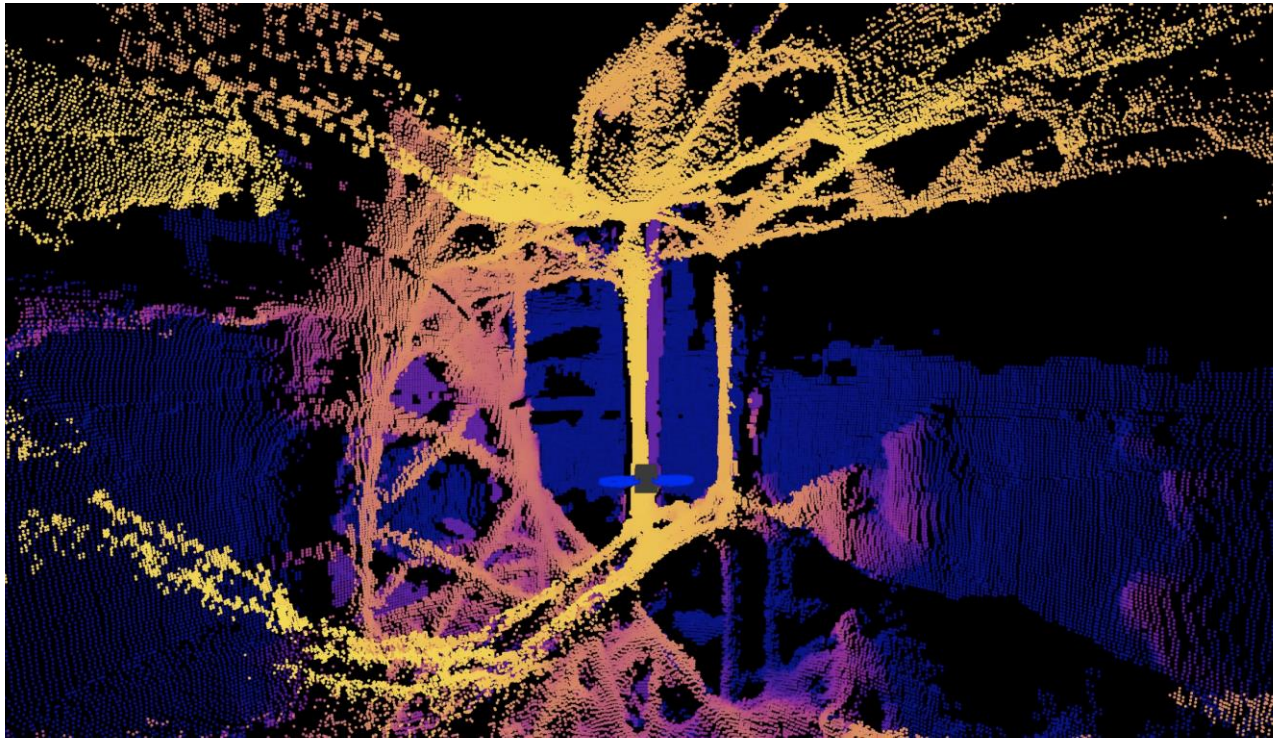
Product / user mindset in research.

Accurate modeling AND learning from data.

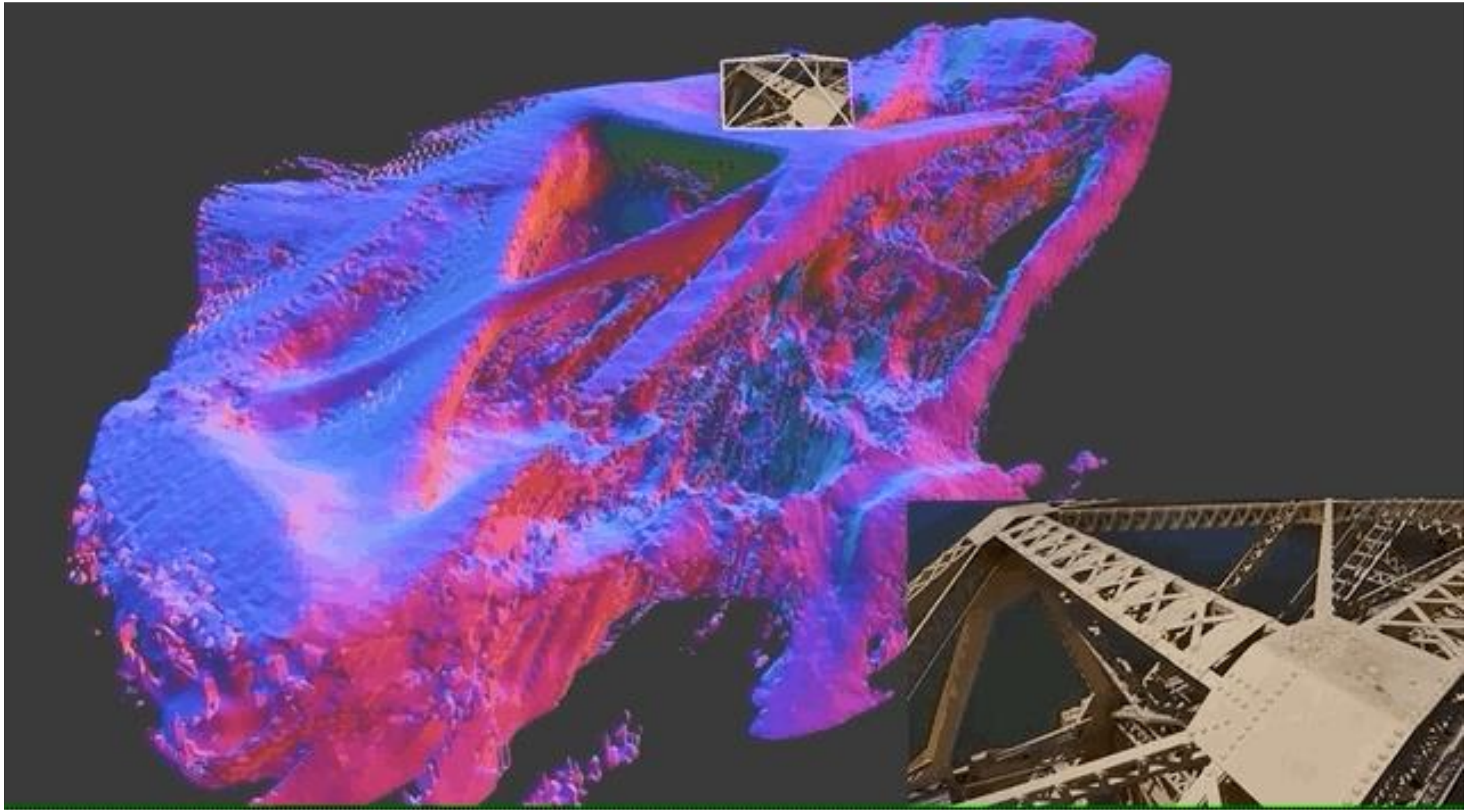
Dedicated time and iteration.

Hardware and rapid fearless testing ability.

Great software engineering.



3D Reconstruction



Resources

Skydio product info

https://www.skydio.com/

Skydio Blog

https://medium.com/skydio

Skydio 3D Scan

https://youtu.be/VxLXTeycyeE

Skydio House Scan

https://youtu.be/MSy_06aOBzg

Recorded MIT lecture on autonomy

https://youtu.be/R5K-IV3J8XM

"End-to-End Learning of Geometry and Context for Deep Stereo Regression"

https://arxiv.org/abs/1703.04309

Hiring contact:

https://skydio.com/jobs