University of Zurich[UZH]

ROBOTICS & PERCEPTION GROUP

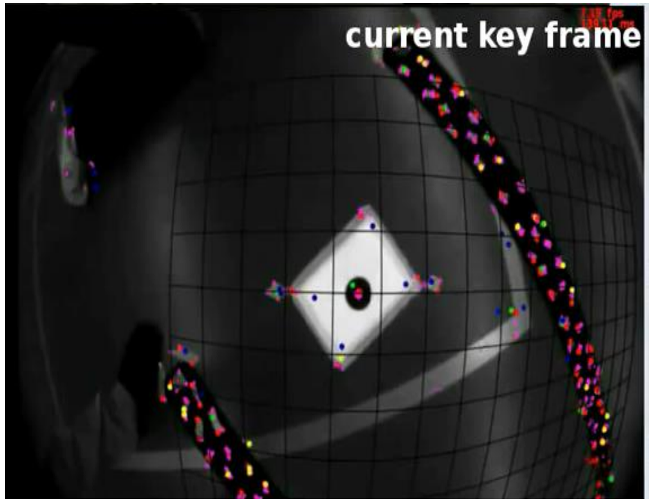
Learning Agile Vision-based Flight

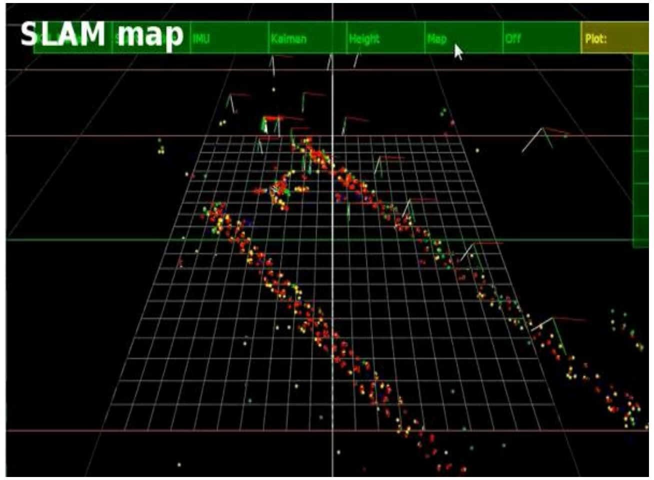
Davide Scaramuzza

Elia Kaufmann | Antonio Loquercio | Leonard Bauersfeld | Yunlong Song

http://rpg.ifi.uzh.ch/

13-years ago: First Vision-based Autonomous Flight





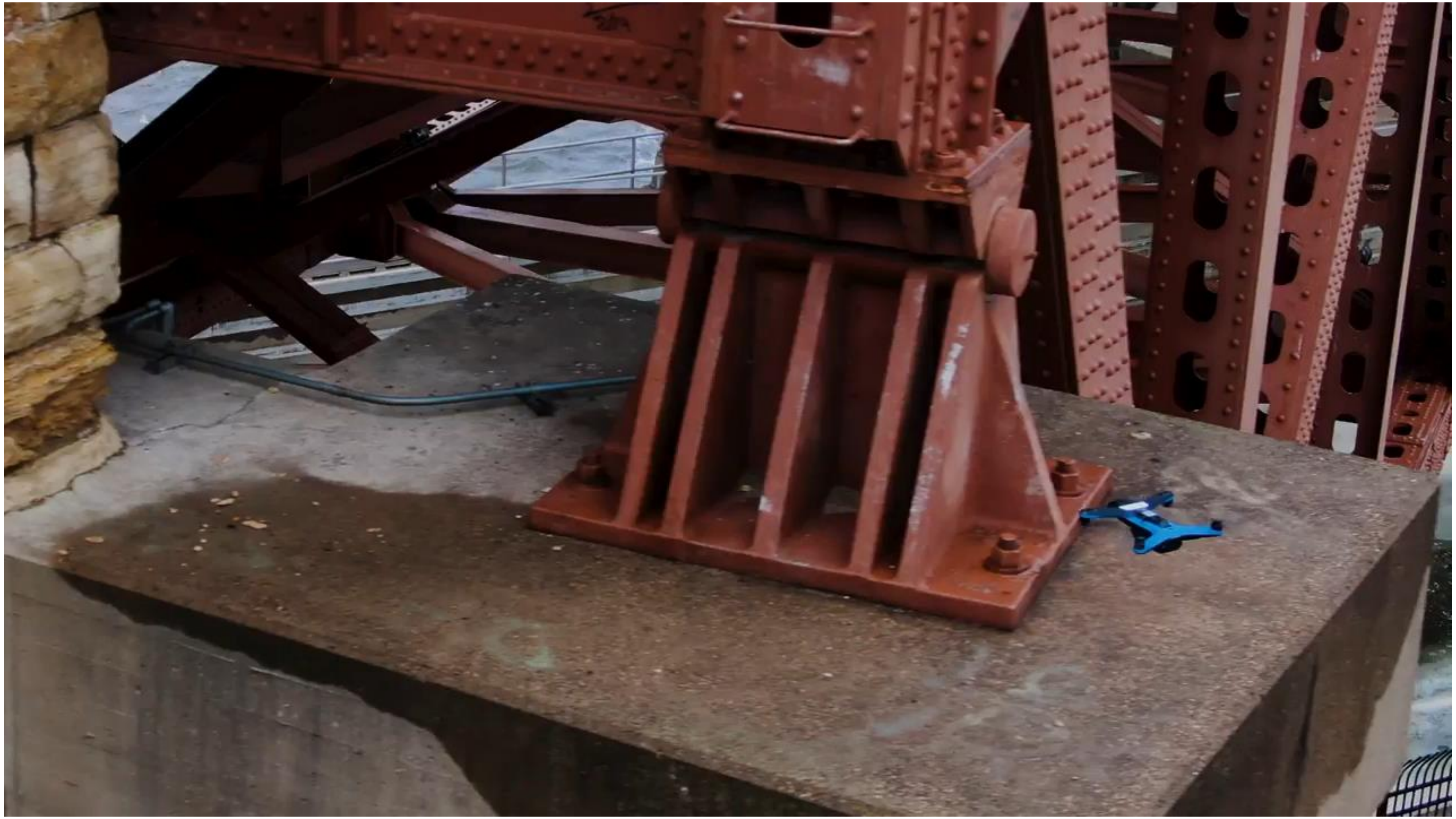
European Micro Aerial Vehicle competition, Sep. 9, 2009

Bloesch, Weiss, Scaramuzza, Siegwart, *Vision Based MAV Navigation in Unknown and Unstructured Environment*, **ICRA'10** [PDF]

Today

NASA Ingenuity helicopter performing autonomous vision-based flight on Mars

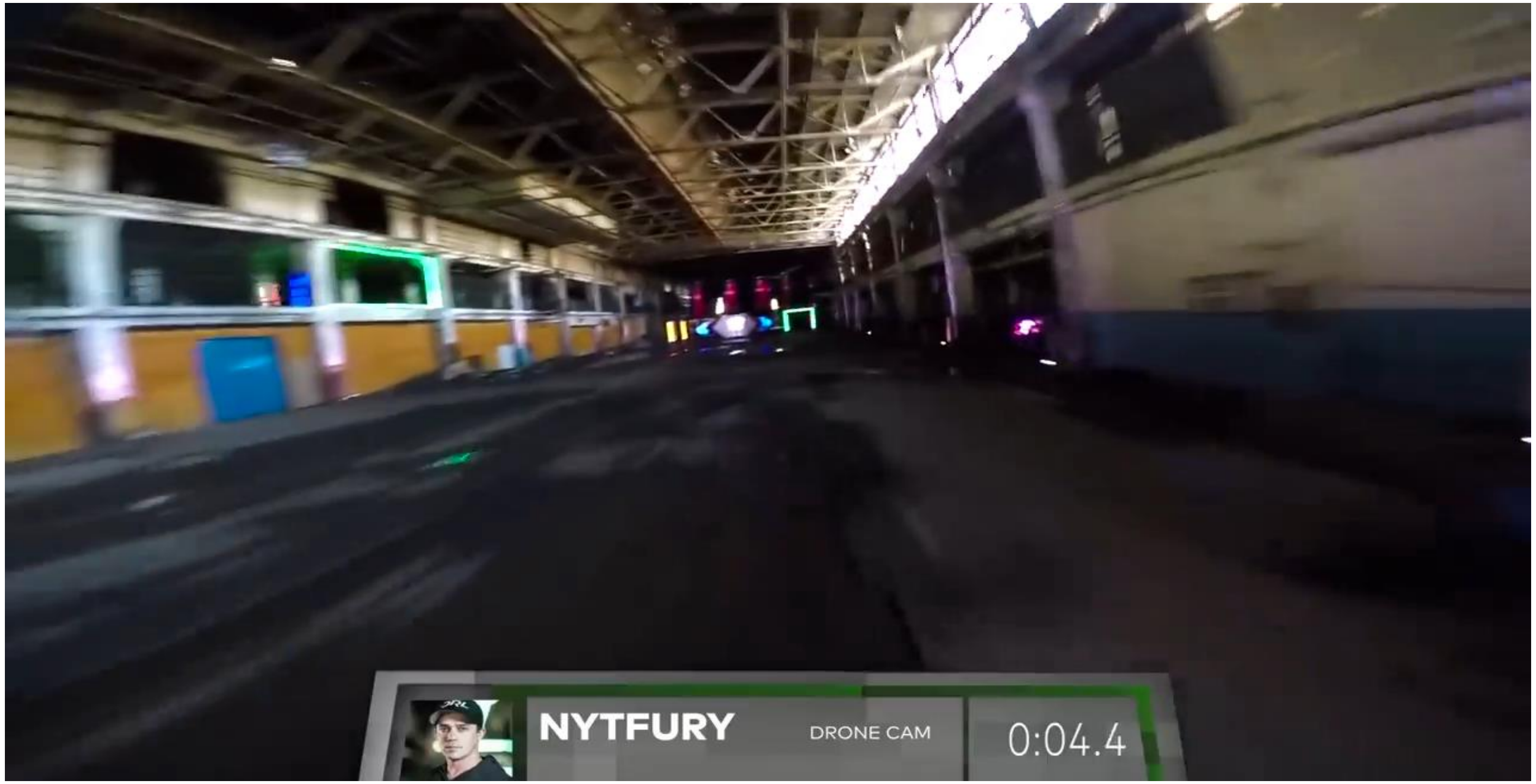
Today



Skydio drones use vision-based navigation for autonomous person following and 3D mapping

What’s Next?

What does it take to fly as **good as or better** than human pilots?

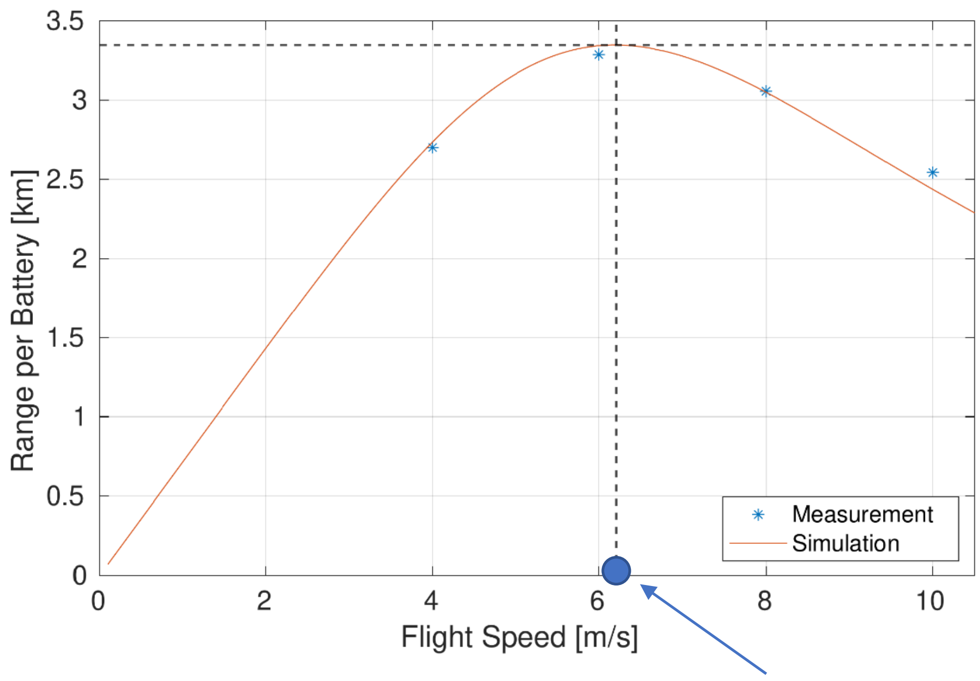


WARNING! This drone is NOT autonomous; it is operated by a human pilot.

Human pilots take years to become **agile**!

Why do bother about Agile Flight?

- **Making drones faster increases their range** (limited by battery life) [1,2]



Optimal speed

[1] Bauersfeld, Scaramuzza, *Range, Endurance, and Optimal Speed Estimates for Multicopters*, IEEE RAL, 2022. PDF.

[2] Karydis, Kumar, *Energetics in robotic flight at small scales, Interface Focus*, 2017. PDF.

Why Agile?

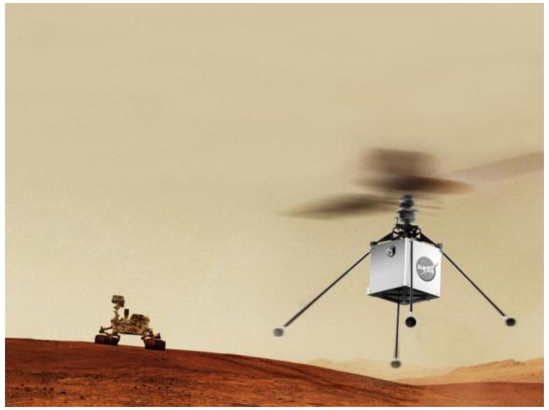
- **Making drones faster increases their range** (limited by battery life) [1,2]
- **Applications**: search & rescue, delivery, exploration, inspection, space
- **Raises fundamental challenges for robotic research**: perception, planning, learning, control
- **Pushes the limits** of vision-based navigation

Search & rescue

Delivery

Exploration & Inspection

Personal air vehicles

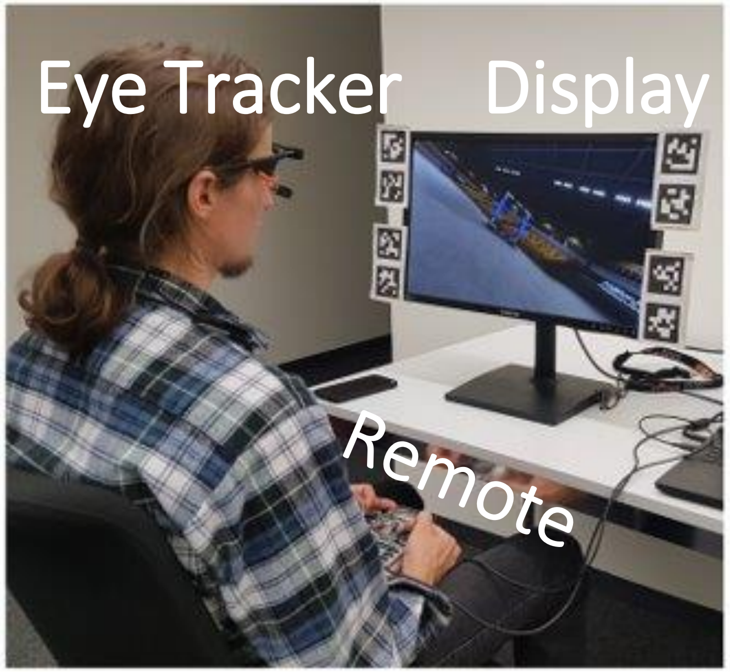


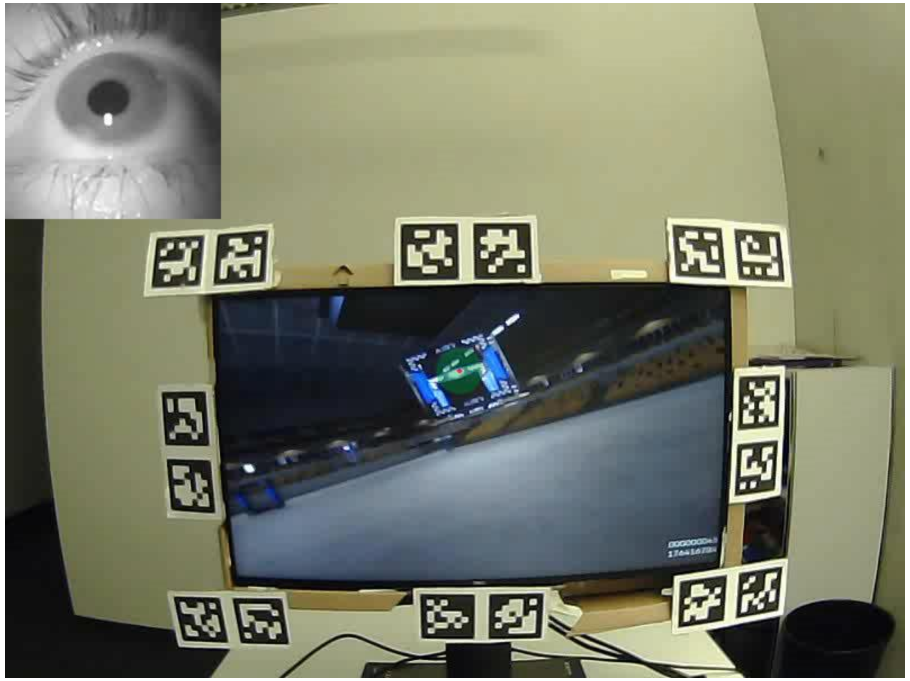
Space

How do Human Pilot Control Drones? **Vision-based!**

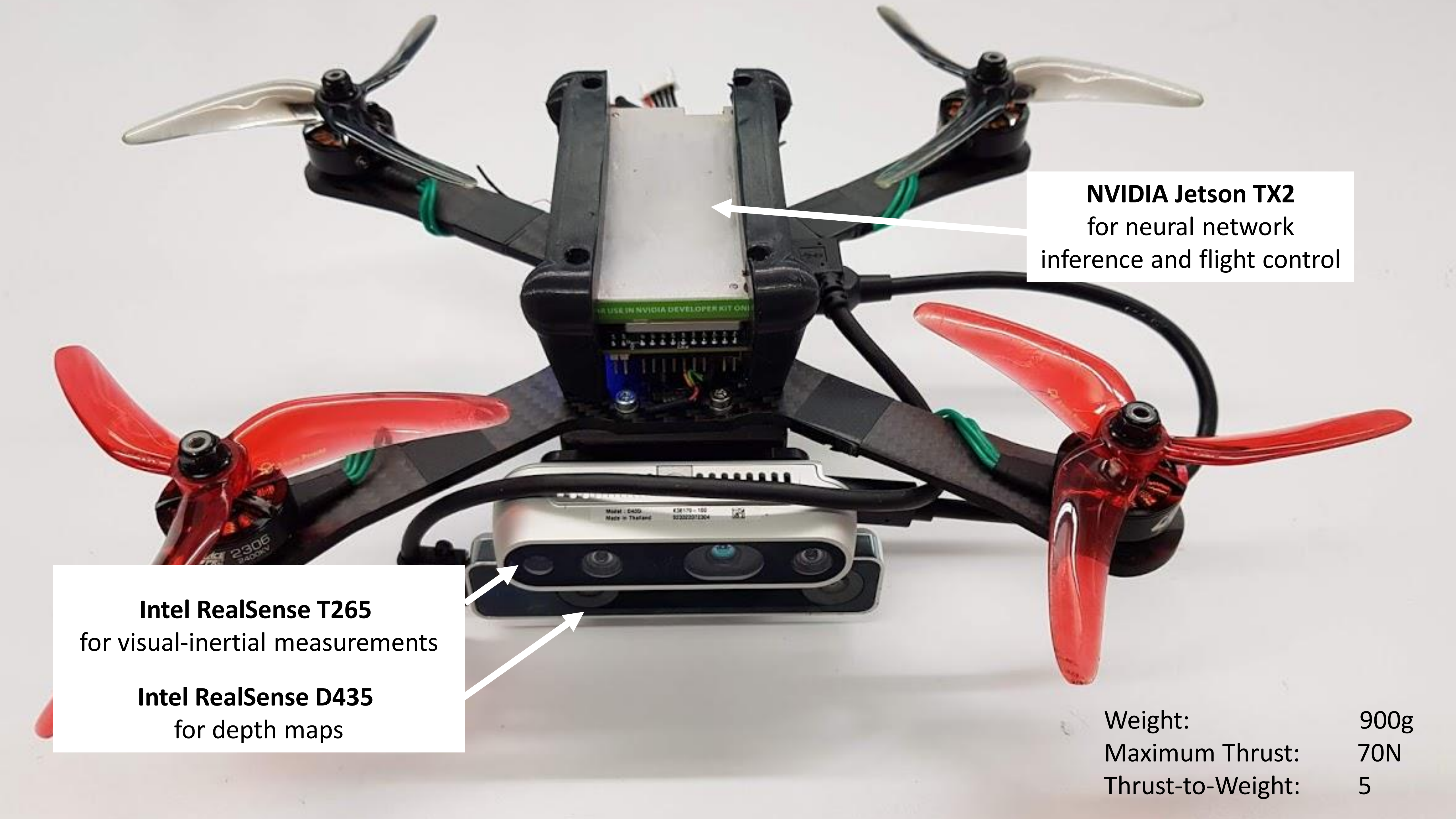
- We recorded **eye-gaze** and **control commands of expert pilots**
- **Our finding: Sensorimotor reaction time** of **220 ms** (i.e., *brain processing* + *behavioral response*)

The human brain is incredibly is slow but better than machines in complex situations!

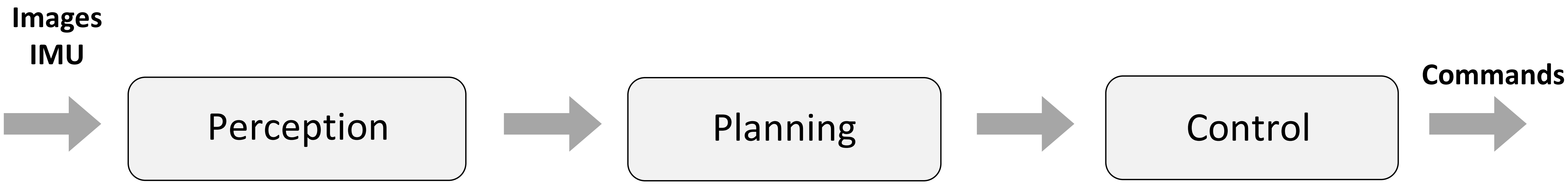




Pfeiffer, Scaramuzza (2021) *Human-piloted drone racing: Perception and control*, **RAL'21**. PDF. Dataset.


NVIDIA Jetson TX2
for neural network
inference and flight control
Intel RealSense T265
for visual-inertial measurements
Intel RealSense D435
for depth maps
Weight: 900g
Maximum Thrust: 70N
Thrust-to-Weight: 5

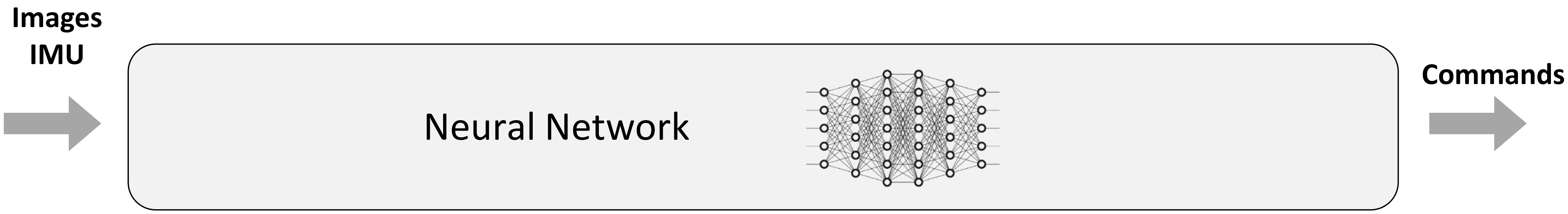
Traditional Drone Control Architecture



Key issues with this architecture:

- **Sensitive to inaccurate models** of sensors, actuators, environment, battery voltage, aerodynamic effects
- Very **sensitive to imperfect perception** due to high-speed motion: motion blur, limited FOV
- Ignores **perception & action coupling**
- **High latency** due to sequential structure

Can we Learn a Navigation Policy?



Key issues with this architecture:

- **Too sample inefficient** to be trained on a physical drone
- Limited **interpretability**

How can we augment the traditional robotic cycle with learning-based methods?

Key Questions

- Should we train it with or without supervision?
- How do we get enough training data?
- Can we learn in simulation?
- How do we address the simulation to reality gap?

Tasks

Autonomous Drone Acrobatics

Navigation in the wild

Autonomous Drone Racing

Deep Drone Acrobatics

GitHub

Code available

This AI-controlled drone is fully autonomous and uses onboard vision and computation

Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** PDF. Video. Code

Trained only in simulation: One-Shot Sim-to-Real Transfer

Source code: https://github.com/uzh-rpg/deep_drone_acrobatics

Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** PDF. Video. Code

How do we address the Simulation-to-Reality Gap?

Simulated

Real

ScienceRobotics

Does computer vision matter for action?

Brady Zhou[1,*], Philipp Krähenbühl[1,2] and Vladlen Koltun[1]

+ See all authors and affiliations

Science Robotics 22 May 2019:
Vol. 4, Issue 30, eaaw6661
DOI: 10.1126/scirobotics.aaw6661

Abstract

Controlled experiments indicate that explicit intermediate representations help action.

Neural networks fed with intermediate representations **train faster**, **achieve higher task performance**, and **generalize better** to previously unseen environments!

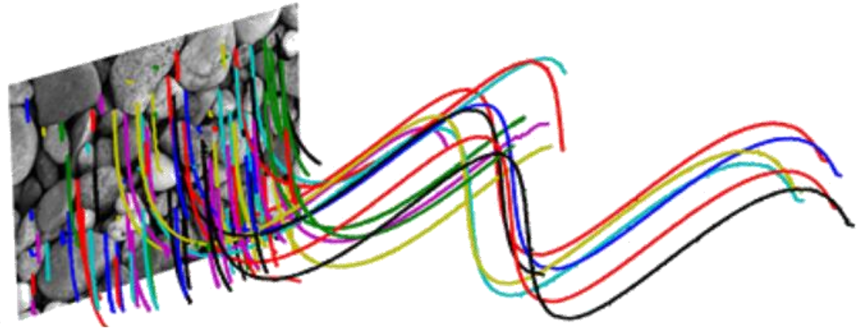
[1] Zhou et al., *Does computer vision matter for action?*, Science Robotics, 2019. PDF. Video.
[2] Sax et al., Mid-Level Visual Representations for Improving Generalization and Sample Efficiency of Visuomotor Policies, CORL'19. PDF.
[3] Chen et al., Robust Policies via Mid-Level Visual Representations: An Experimental Study in Manipulation and Navigation, CORL'20. PDF.

Which Sensory Abstraction Reduces the Sim-to-Real Gap?

Simulated

Real

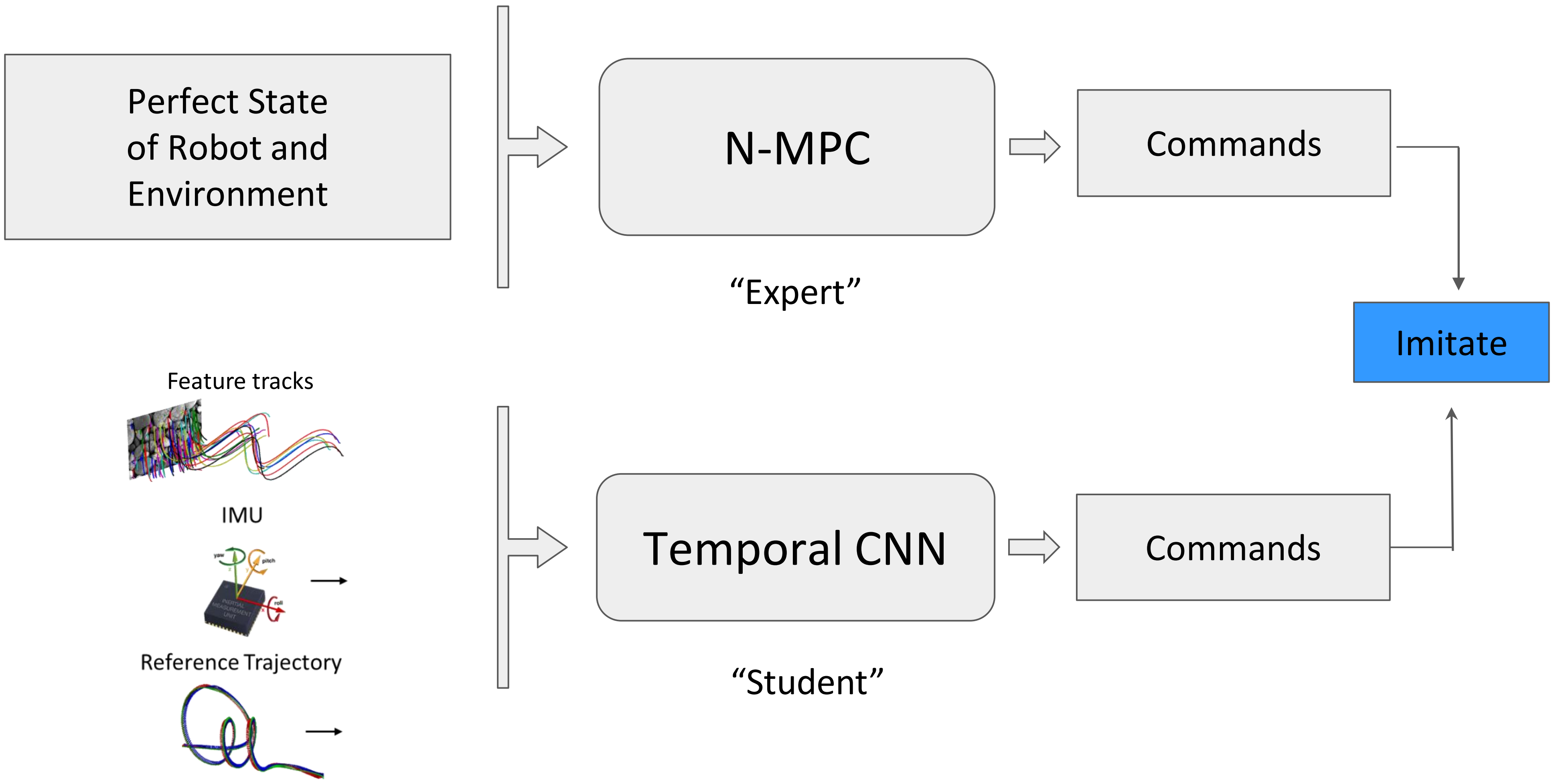
Feature tracks



We prove that sensory abstractions reduce the Wasserstein distance between observation models in simulation and reality

Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** [PDF]. [Video].[Code]

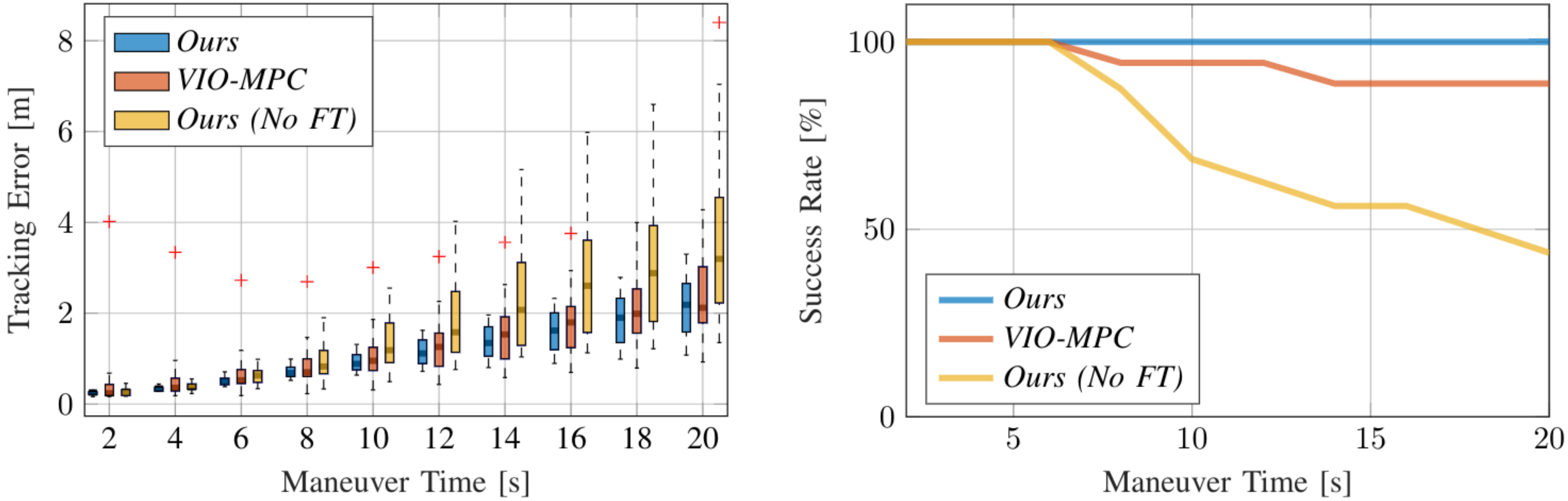
Trained via Privileged Imitation Learning



Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** PDF. Video.Code

Comparison against Traditional Baselines

The **learned end-to-end control policy outperforms by up to 25%** the traditional baseline of state-estimation and control (VIO+MPC)



Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** [PDF](). [Video]().[Code]()

Key Questions

- How do we get enough training data?
- Can we learn in simulation?
- How do we address the simulation to reality gap?

Tasks

Autonomous Drone Acrobatics

Navigation in the wild

Autonomous Drone Racing

Learning High-Speed Flight in the Wild (40 km/h)

GitHub

Code available

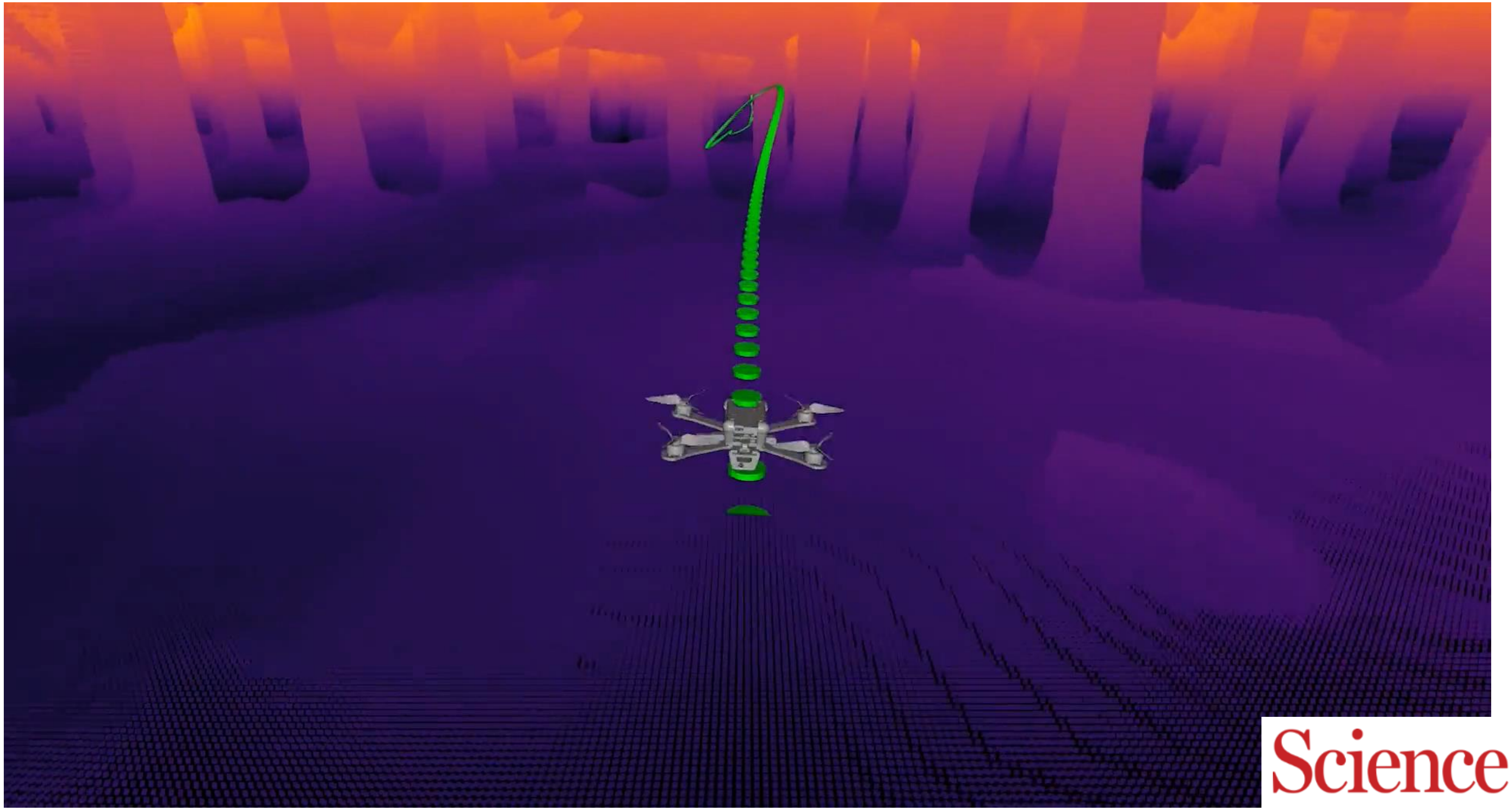
Science Robotics

This AI-controlled drone is fully autonomous and uses onboard vision and computation

Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, *Learning High Speed Flight in the Wild*, **Science Robotics, 2021**

PDF. Video. Code & Datasets

Trained only in simulation: One-Shot Sim-to-Real Transfer



GitHub

Code available

Simulator:
Flightmare

Science Robotics

[1] Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, *Learning High Speed Flight in the Wild*, **Science Robotics**. PDF. Video. Code & Data
[2] Simulator used: Song, *Flightmare: A Flexible Quadrotor Simulator*, **CORL'20**, PDF Video Website

Which Sensory Abstraction Reduces the Sim-to-Real Gap?

RGB

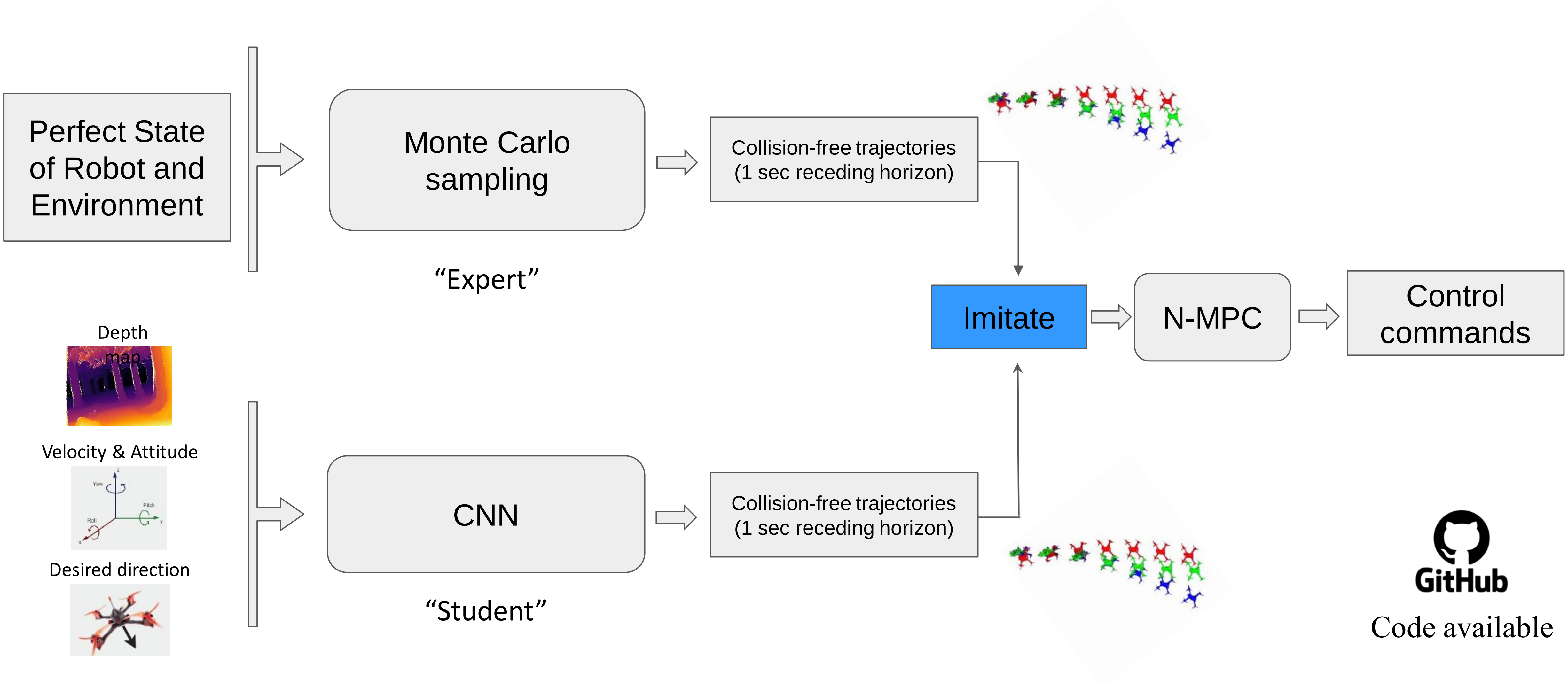
Simulation

Real-World

Depth map

Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, *Learning High Speed Flight in the Wild*, **Science Robotics., 2021**
PDF. Video. Code & Datasets

Trained via Privileged Imitation Learning



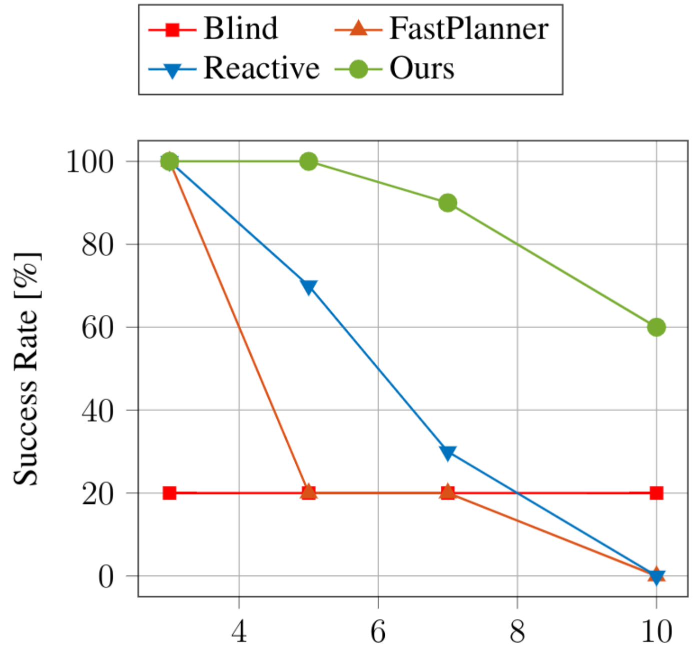
Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, *Learning High Speed Flight in the Wild*, **Science Robotics., 2021**

[PDF](). [Video](). [Code & Datasets]()

Comparison against Baselines

- Up to **6x more efficient** than baselines
- At 10 m/s the success rate only dropped to 60% → **We can fly 2x faster** than baselines

Method	Components	Perc. [%]	Total Proc. Latency [ms]
FastPlanner (18)	Pre-processing Mapping Planning	22.3 75.5 2.2	65.2
Reactive (32)	Pre-processing Planning	72.3 27.7	19.1
Ours	Pre-processing NN inference Projection	3.9 93.0 3.1	10.3 (2.6*)

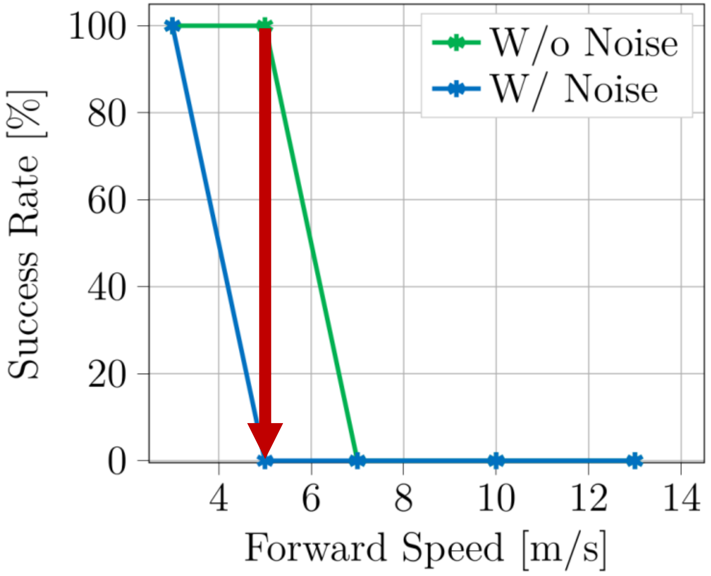


- [FastPlanner] Zhou, Gao, Wang, Liu, Shen, *Robust and efficient quadrotor trajectory generation for fast autonomous flight*, RAL'19
- [Reactive] Florence, Carter, Tedrake, *Integrated perception and control at high speed: Evaluating collision avoidance maneuvers without maps*, WAFR'20
- [Ours] Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, Learning High-Speed Flight in the Wild, Science Robotics, 2021

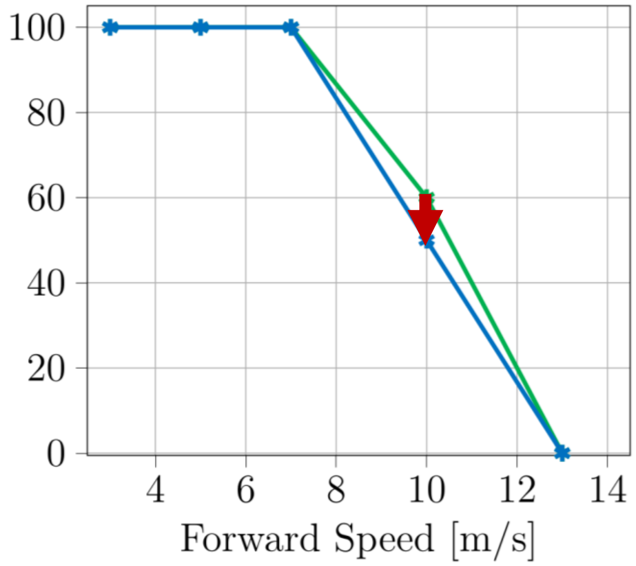
Comparison against Baselines

With sensor noise, the performance **only drops by 5%**

FastPlannerer



Ours



- [**FastPlanner**] Zhou, Gao, Wang, Liu, Shen, *Robust and efficient quadrotor trajectory generation for fast autonomous flight*, RAL'19
- [**Ours**] Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, Learning High-Speed Flight in the Wild, Science Robotics, 2021

Learning High-Speed Flight in the Wild (40 km/h)

This AI-controlled drone is fully autonomous and uses onboard vision and computation

Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, *Learning High Speed Flight in the Wild*, **Science Robotics, 2021**
[PDF](). [Video](). [Code & Datasets]()

Takeaways

- Pros:
  - **Complex navigation** strategies can be **distilled** into efficient deep sensorimotor policies
  - Learning-based policies are **more robust** against **sensing noise** and have **lower latency**
  - **Abstraction of policy inputs** enables robust transfer from simulation to reality

- Cons:
  - The presented approaches still rely on **labeled expert data**. What if we cannot create such an expert?

Key Questions

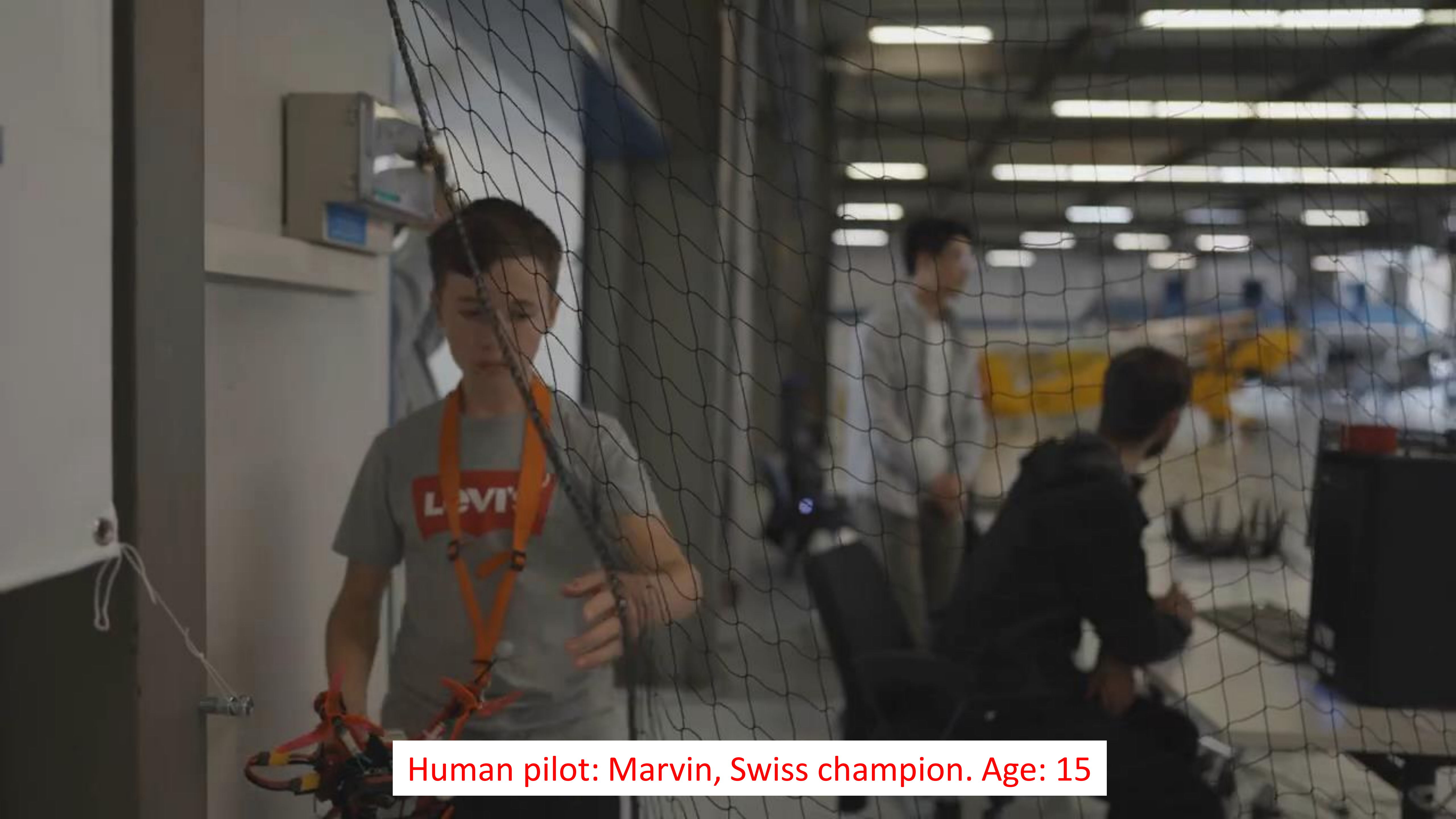
- How do we get enough training data?
- Can we learn in simulation?
- What's the expert?
- What sensory and control abstractions reduce the simulation to reality gap?

Tasks

Autonomous Drone Acrobatics

Navigation in the wild

Autonomous Drone Racing



Human pilot: Marvin, Swiss champion. Age: 15

In racing, drones are constantly pushed to their physical limits.
Any small mistake can lead to a crash!

In racing, drones are constantly pushed to their physical limits. Any small mistake can lead to a crash!

Autonomous Drone Racing: Problem Definition

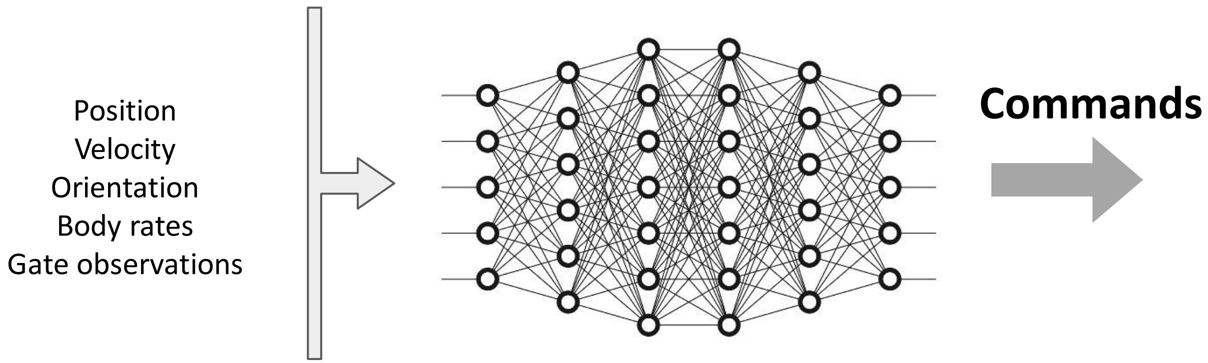
Find control inputs $u(t)$ that minimize flight time through a series of gates

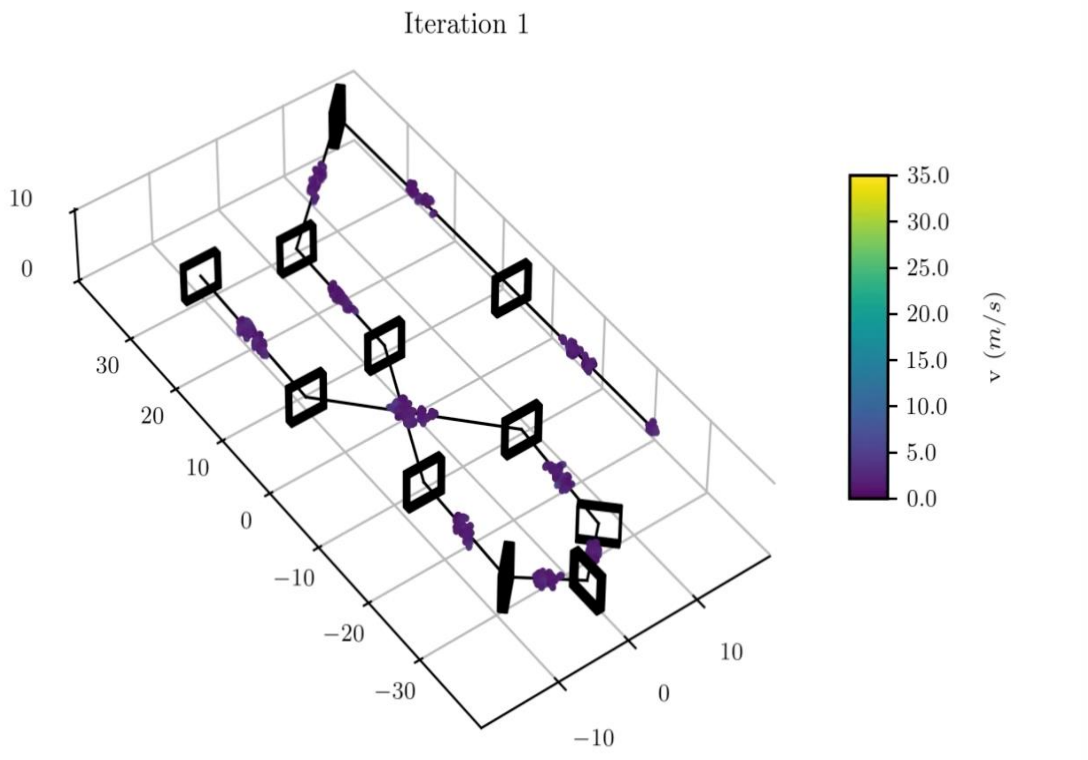
$$min_u \int_0^T dt$$

subject to:
- *System dynamics*
- *Input constraints*
- *State constraints*
- *Physical constraints*

Autonomous Drone Racing with Deep Reinforcement Learning

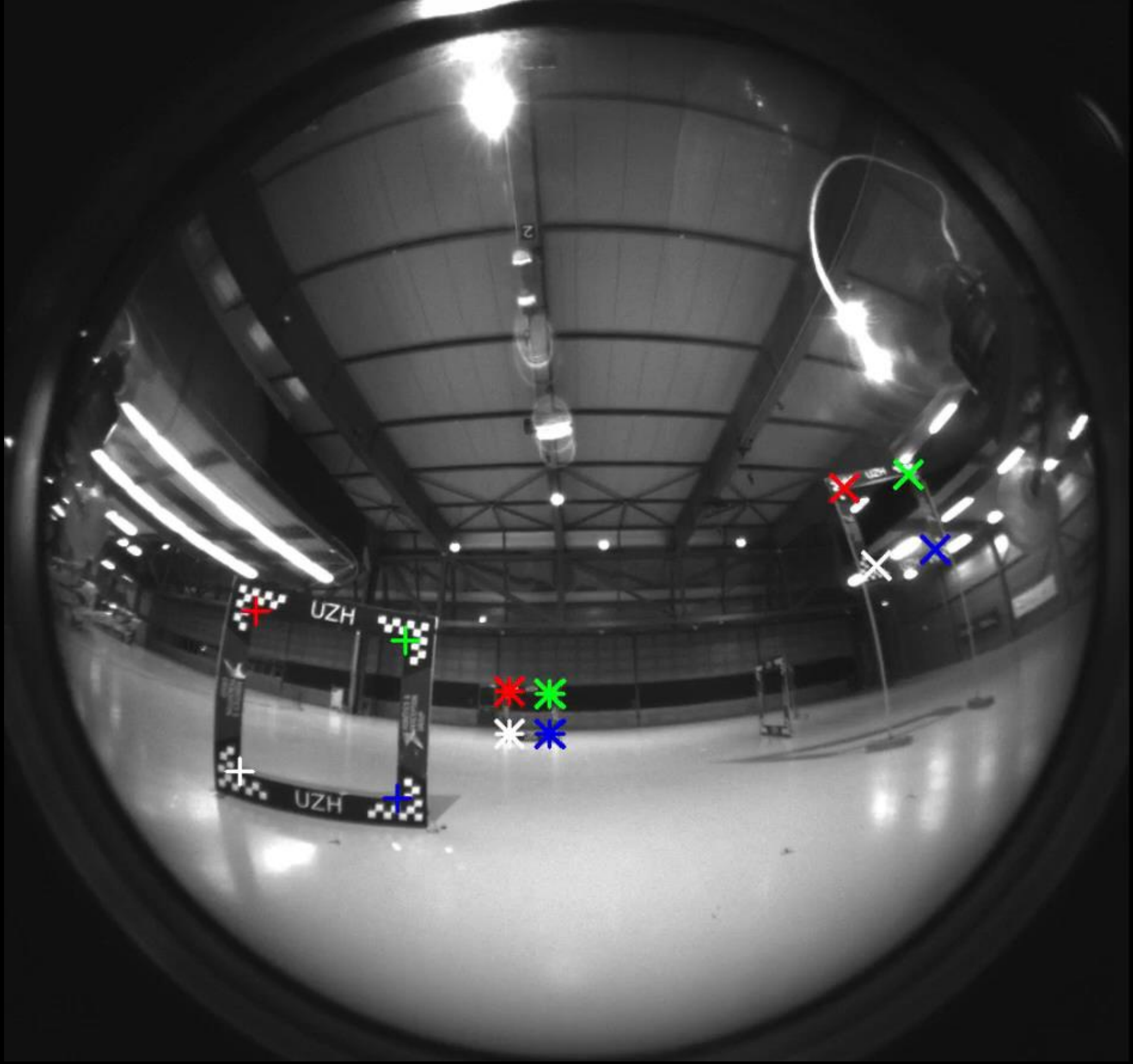
- Agent based on **PPO**
- Parallel sampling using **100 quadrotors**
- **Distributed initialization** strategy
- **500s** of flight time **per update step**
- Typical training requires **~1,400h of flight time or 2-3h** of wall clock time





Song, Steinweg, Kaufmann, Scaramuzza, *Autonomous Drone Racing with Deep Reinforcement Learning*, IROS'21. PDF. Video.
Kaufmann, Bauersfeld, Scaramuzza, *A Benchmark Comparison of Learned Control Policies for Agile Quadrotor Flight*, ICRA'22. PDF. Video

This AI-controlled drone is fully autonomous and uses onboard vision and computation


UZH
UZH

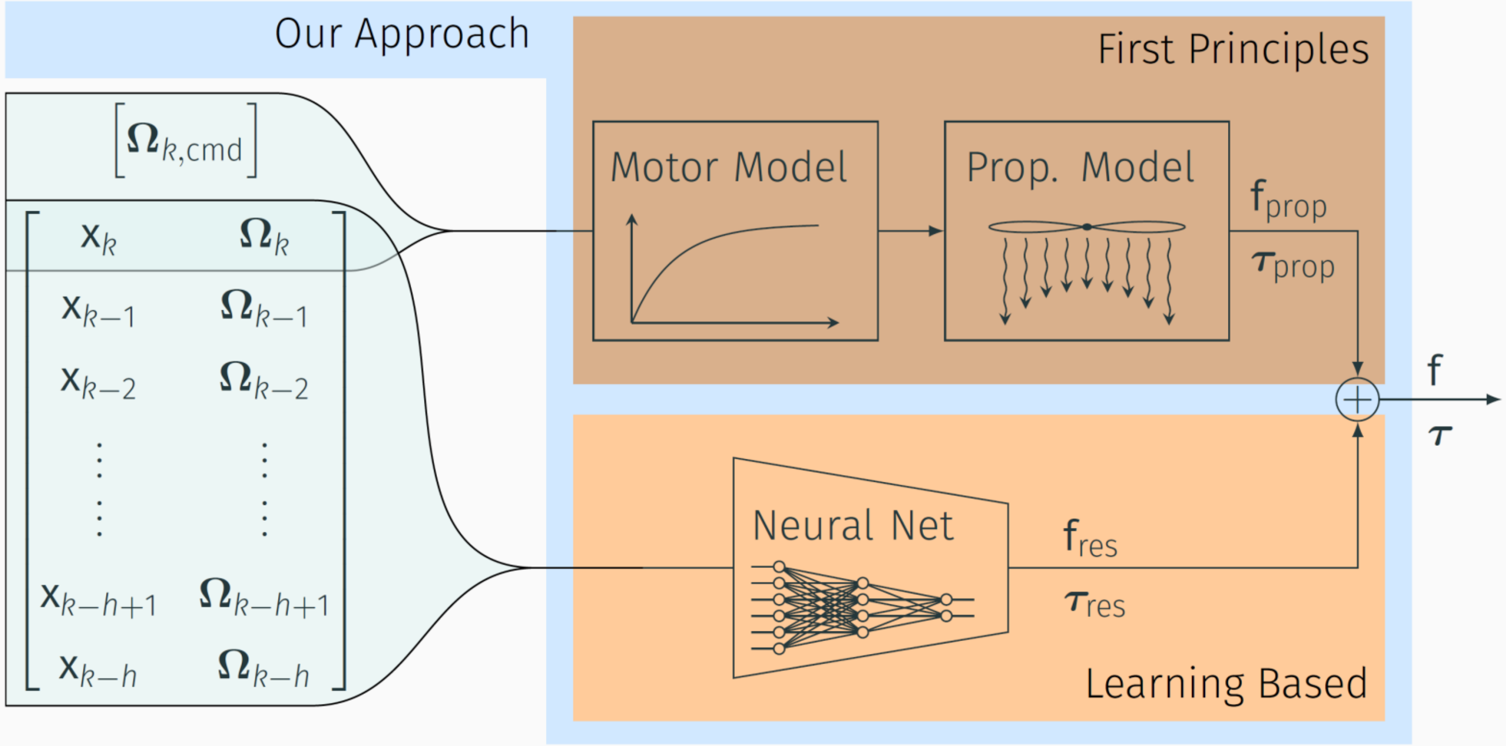
Improving Physics simulation: Aerodynamic Effects

- Rotor-to-rotor interactions
- Turbulences

[1] Bauersfeld et al. *NeuroBEM: Hybrid Aerodynamic Quadrotor Model*, **RSS'21**. PDF. Video. Code & Datasets.
[2] Torrente et al., *Data-Driven MPC for Quadrotors*, **RAL'21**. PDF. YouTube. Code

NeuroBEM: Hybrid Aerodynamic Quadrotor Model

- We use a **neural network** to **model residual forces and torques** unexplained by Blade Element Momentum Theory.
- **Improves physics realism** wrt classic drone simulators by **up to 60%**.



Bauersfeld et al. *NeuroBEM: Hybrid Aerodynamic Quadrotor Model*, **RSS'21**. PDF. Video. Code & Datasets.

Conclusions

- Deep Sensorimotor Policies allow to push robotic platforms to their limits, even with **onboard sensing** and computation.
- Simulation-to-Reality transfer can be facilitated using:
  - **Abstraction of policy inputs**
    - Feature tracks
    - Depth images
  - **Choice of control modality**
    - Collective thrust and body rates achieve high agility while being robust to domain shift [1]
  - **By improving simulation of aerodynamic effects** from real-world data

[1] Kaufmann, Bauersfeld, Scaramuzza, *A Benchmark Comparison of Learned Control Policies for Agile Quadrotor Flight*, ICRA'22. PDF. Video

Thanks!

Code, **datasets**, **videos**, and **publications**, **slides**: http://rpg.ifi.uzh.ch/

I am hiring PhD students and Postdocs in AI

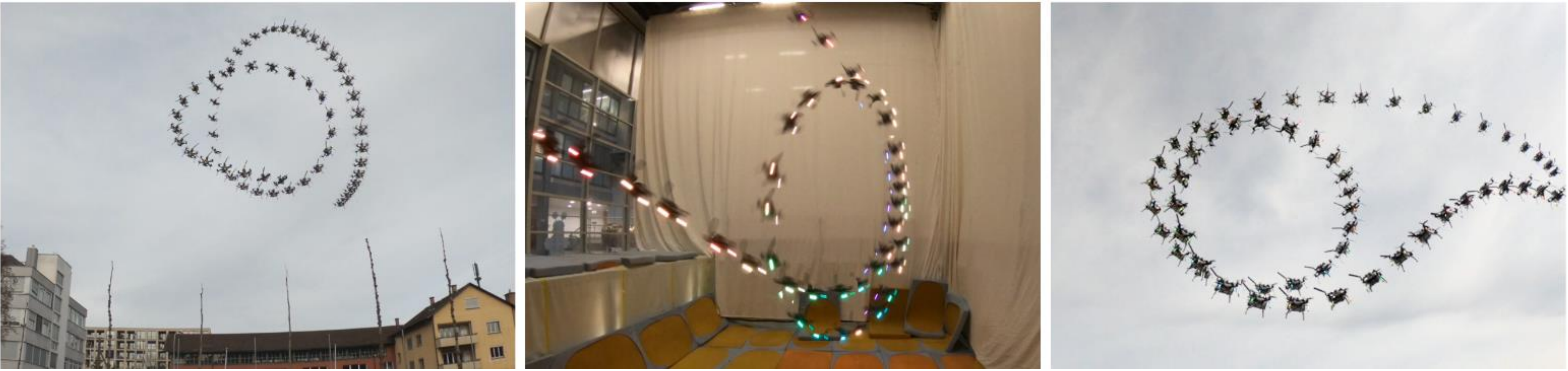
ailabRPG

@davsca1

@davidescaramuzza

Maneuvers Catalogue

- We evaluate on different acrobatic maneuvers: *Power Loop, Barrell Roll, Matty Loop, Combo*
- The maneuvers require acceleration up to *3g*.

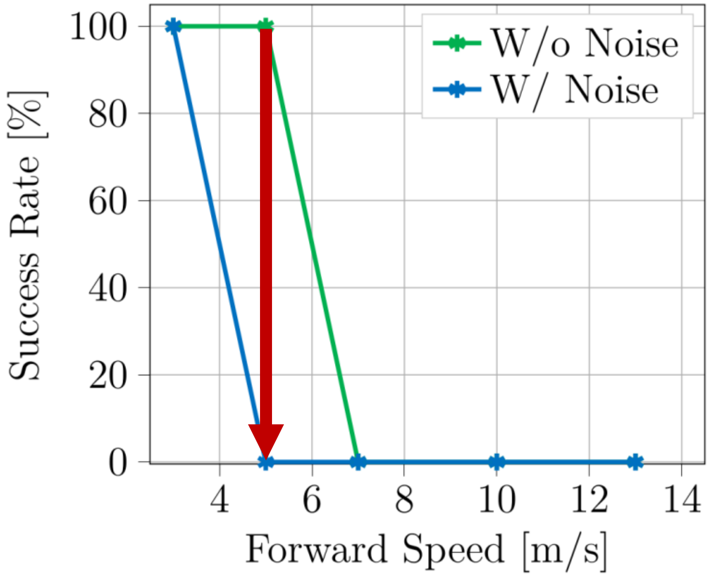


Kaufmann et al., *Deep Drone Acrobatics*, **RSS 2020. Best Paper Award Finalist.** PDF. Video.Code

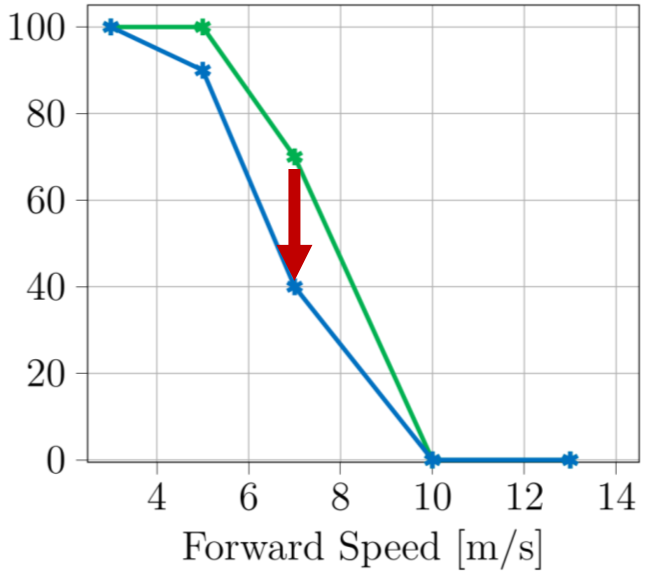
Comparison against Baselines

With sensor noise, the performance **only drops by 5%**

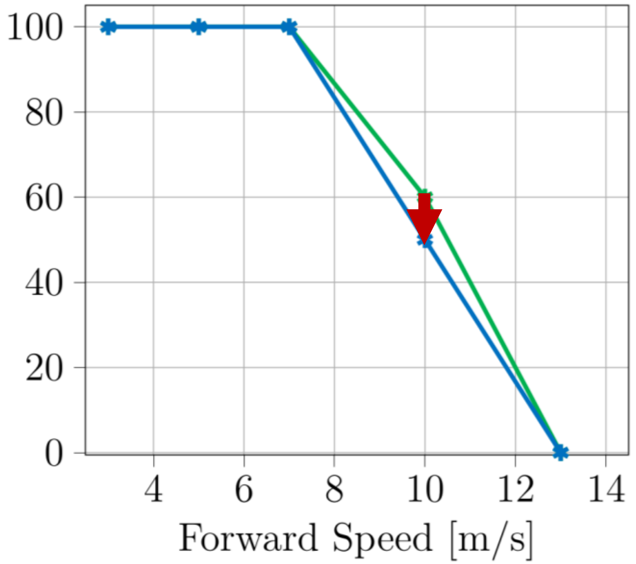
FastPlannerer



Reactive

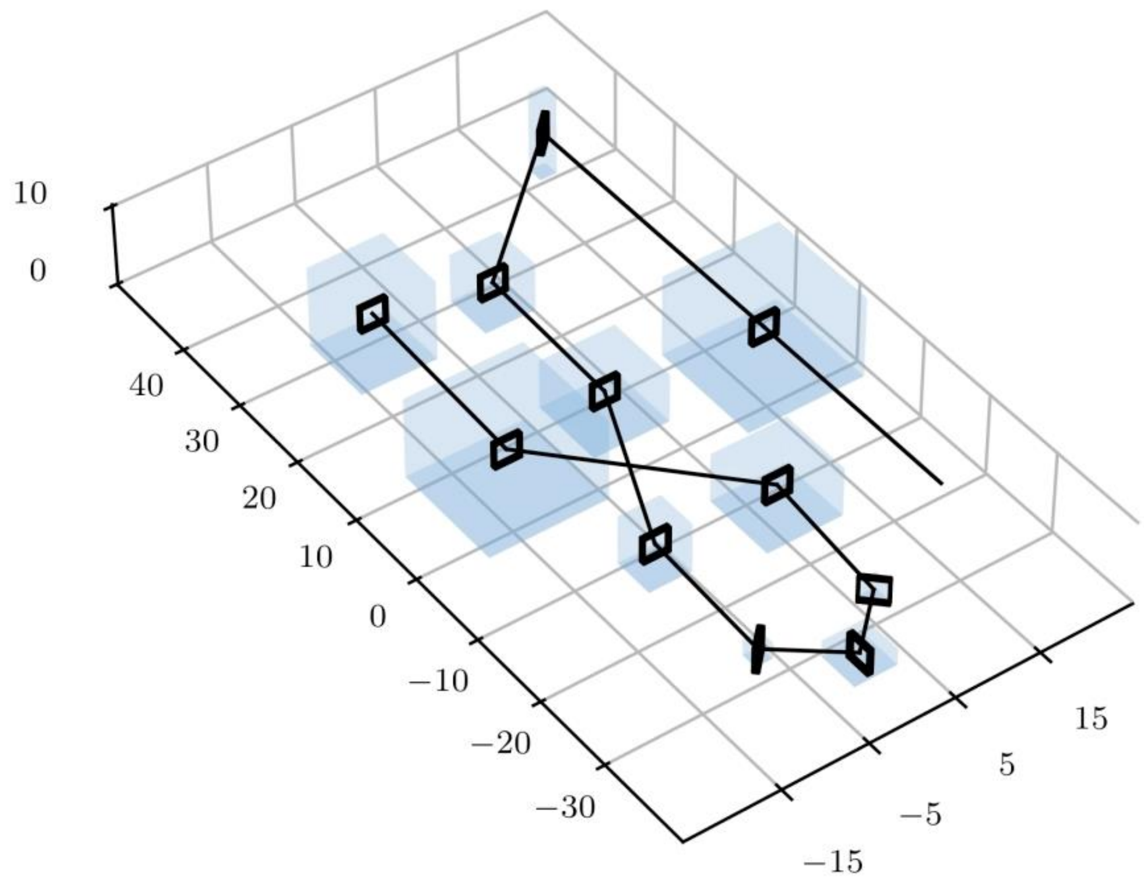


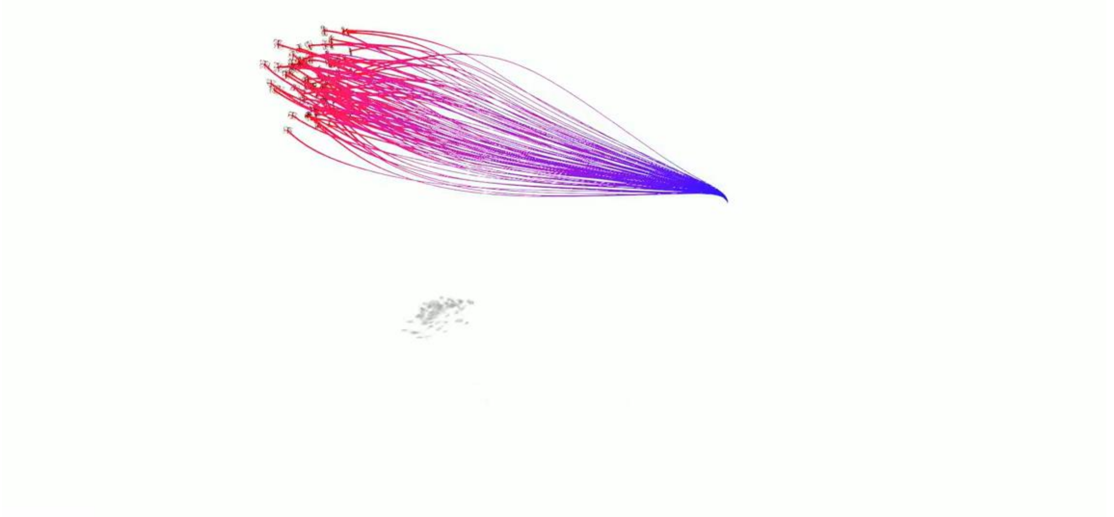
Ours



- [**FastPlanner**] Zhou, Gao, Wang, Liu, Shen, *Robust and efficient quadrotor trajectory generation for fast autonomous flight*, RAL'19
- [**Reactive**] Florence, Carter, Tedrake, *Integrated perception and control at high speed: Evaluating collision avoidance maneuvers without maps*, WAFR'20
- [**Ours**] Loquercio, Kaufmann, Ranftl, Mueller, Koltun, Scaramuzza, Learning High-Speed Flight in the Wild, Science Robotics, 2021

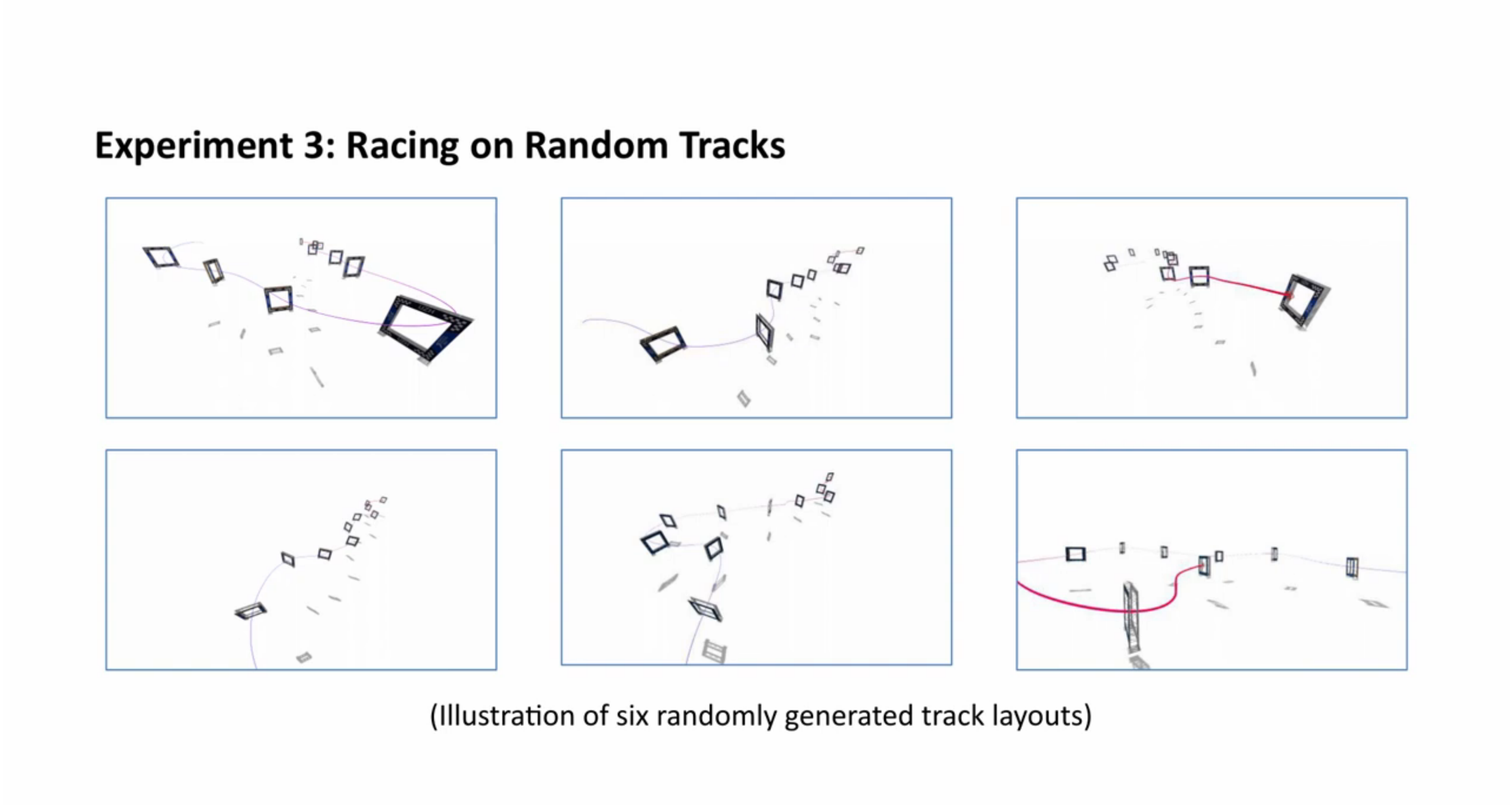
Racing in Uncertain Tracks





Song, Steinweg, Kaufmann, Scaramuzza, *Autonomous Drone Racing with Deep Reinforcement Learning*, IROS’21. PDF. Video.

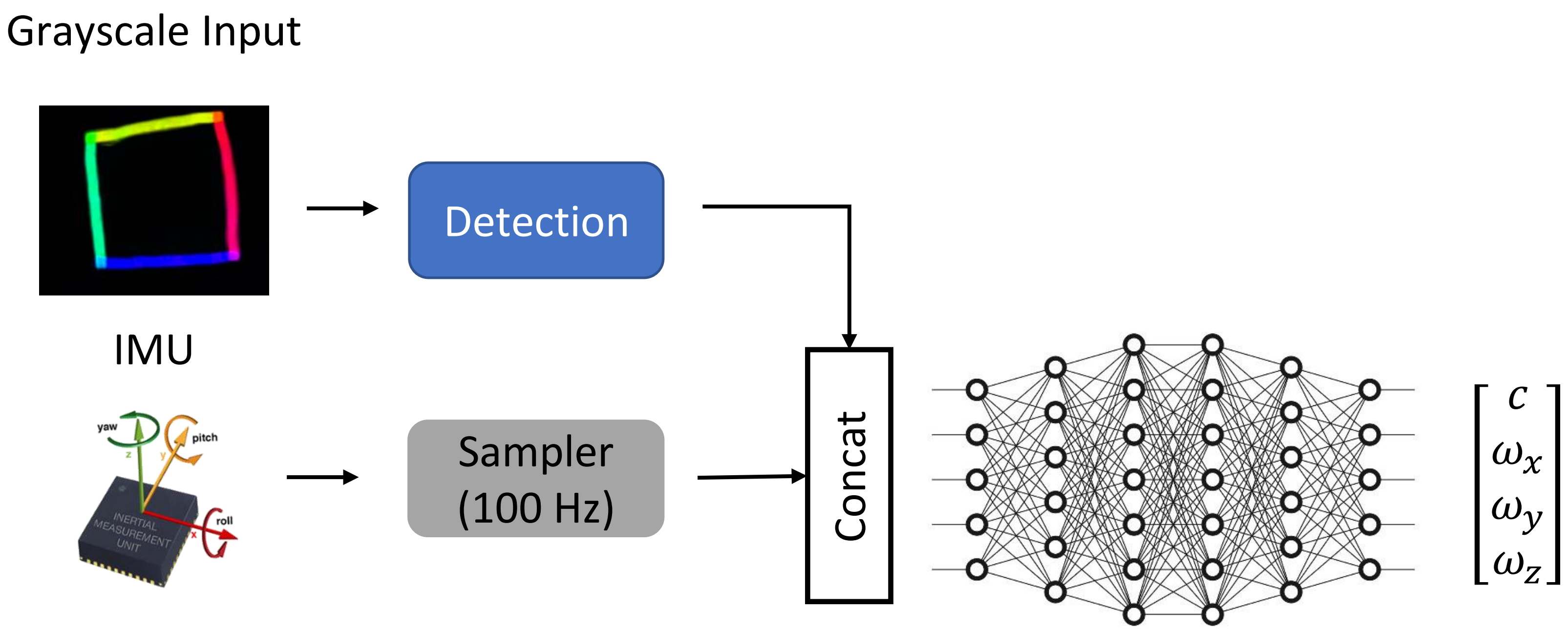
Generalization to Different Tracks



Song, Steinweg, Kaufmann, Scaramuzza, *Autonomous Drone Racing with Deep Reinforcement Learning*, IROS'21. PDF. Video.


UZH
Robotics and
Perception Group
intel robotics
Universität
Zürich
ROBOTICS &
PERCEPTION
GROUP
UZH

Vision-Based Autonomous Drone Racing



Reward: $r_t = r_t^{\text{prog}} + r_t^{\text{perc}} - r_t^{\text{crash}}$

Imitation Learning vs Model-Free RL

Imitation Learning

- Simple to train (few hyperparameters to tune)
- Relatively data-efficient

- Requires access to an expert policy to imitate

Model-Free RL

- High-Level task description might be enough
- No need for expert design

- Sensitive to hyperparameter tuning
- Less data efficient → efficient simulation

Outlook: Head-to-Head AI vs. human

AI Drone

Human

Marvin is a 15-y.o.
world class drone-racing pilot


UZH
UZH
UZH
UZH
UZH
UZH
UZH
ROBOTICS & PERCEPTION GROUP
ROBOTICS & PERCEPTION GROUP

Outlook: Learning to Race Autonomously

This AI-controlled drone is fully autonomous and uses onboard vision and computation

[1] Foehn et al., *Time-Optimal Planning for Quadrotor Waypoint Flight*, **Science Robotics**, 2021. PDF. Video. Code. **Featured on Forbes magazine.**
[2] Foehn et al., *AlphaPilot: Autonomous Drone Racing*, **RSS** 2020, **Best Systems Paper Award**. PDF Video. **2nd place at AlphaPilot Challenge**
[3] Kaufmann et al., *Beauty and the Beast: Optimal Methods Meet Learning for Drone Racing*, ICRA'19. PDF. Video **1st place at IROS'18 Drone race**. Video.
[4] Loquercio, et al., *Deep Drone Racing,* **IEEE Transactions on Robotics** 2020. **Best Paper Award finalist.** PDF. Video
[5] Song et al, *Autonomous Drone Racing with Deep Reinforcement Learning*, IROS'21. PDF. Video
[6] **Simulator used**: Song et al., *Flightmare: A Flexible Quadrotor Simulator*, CORL'20, PDF Video Website

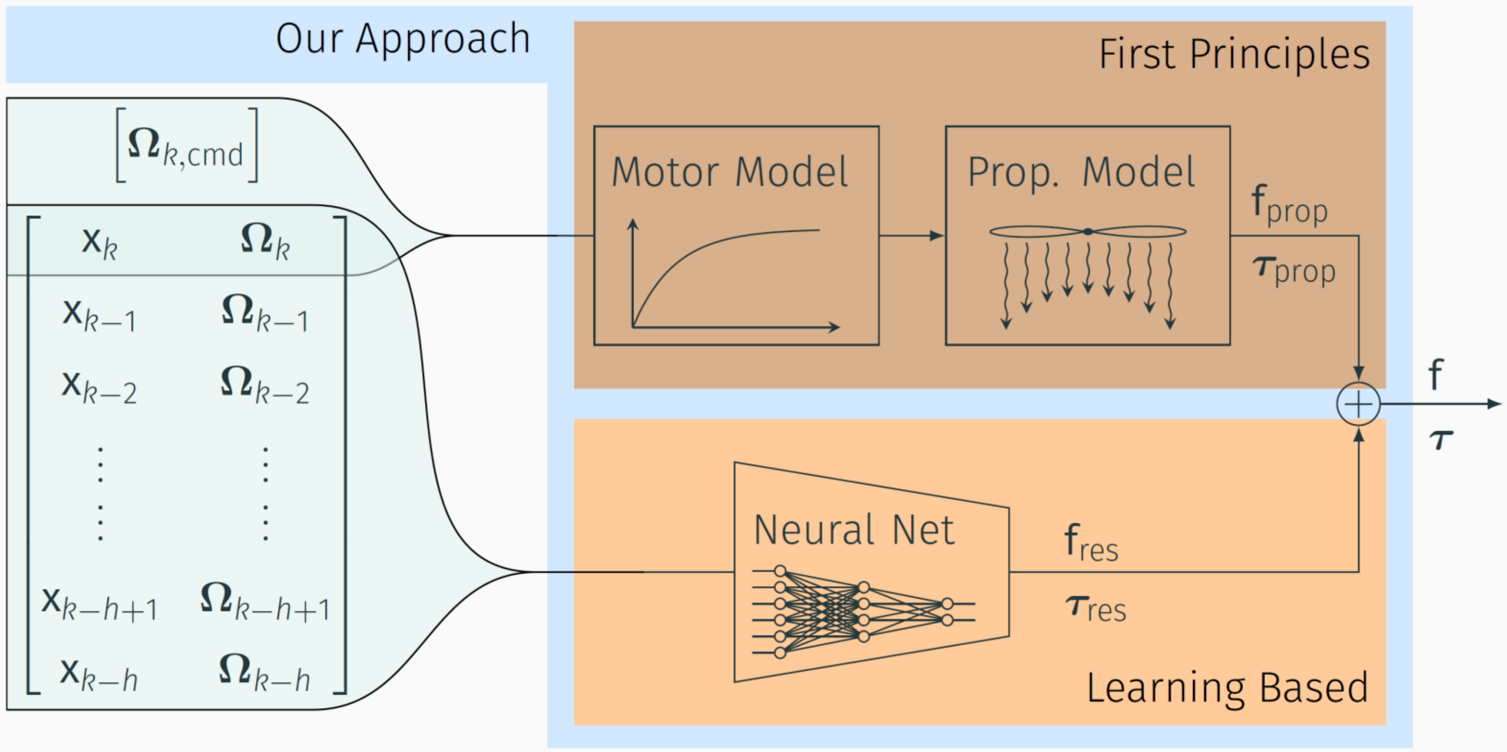
Improving Physics simulation: Aerodynamic Effects

- Rotor-to-rotor interactions
- Turbulences

[1] Bauersfeld et al. *NeuroBEM: Hybrid Aerodynamic Quadrotor Model*, **RSS'21**. PDF. Video. Code & Datasets.
[2] Torrente et al., *Data-Driven MPC for Quadrotors*, **RAL'21**. PDF. YouTube. Code

NeuroBEM: Hybrid Aerodynamic Quadrotor Model

- We use a **neural network** to **model residual forces and torques** unexplained by Blade Element Momentum Theory.
- **Outperforms state of the art** models by **up to 60%**.

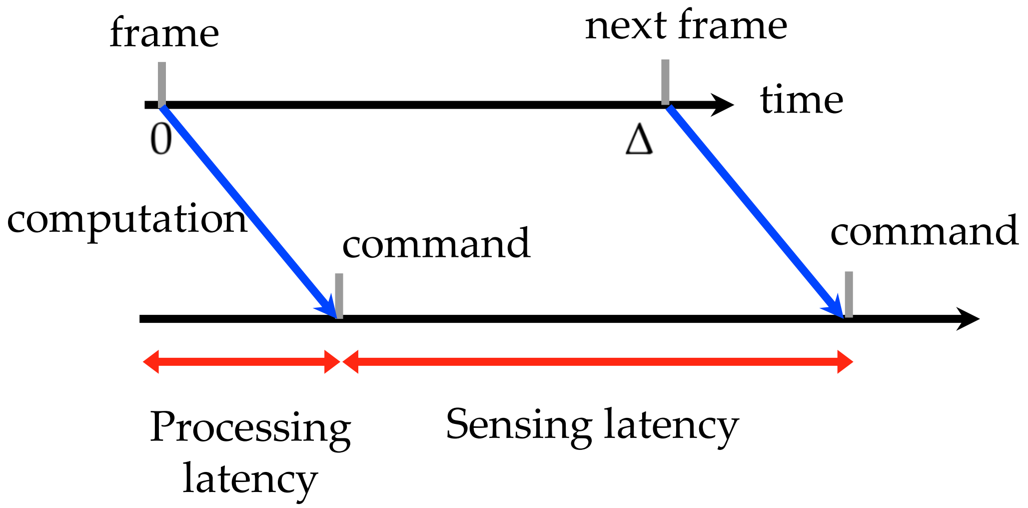


Bauersfeld et al. *NeuroBEM: Hybrid Aerodynamic Quadrotor Model*, **RSS'21**. PDF. Video. Code & Datasets.

BBC: “How sparrowhawks catch garden birds”
https://youtu.be/Ra6I6svXQPg

To go faster, we need faster sensors and algorithms

The agility of a robot is limited by the **latency sensing pipeline perception-action pipeline**

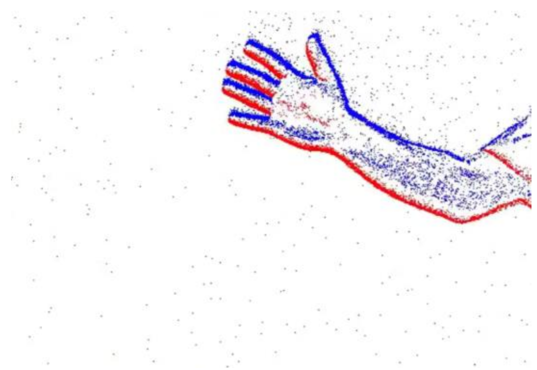


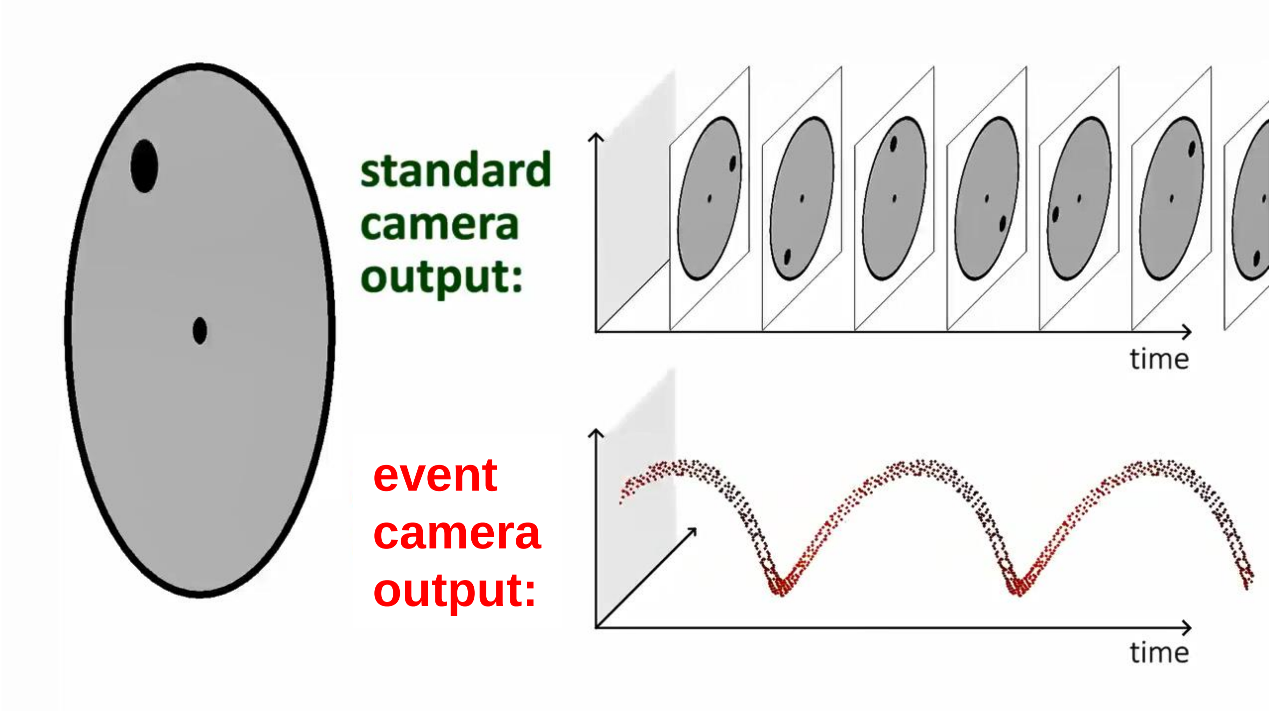
Can we create a low-latency, low-discretization perception pipeline?

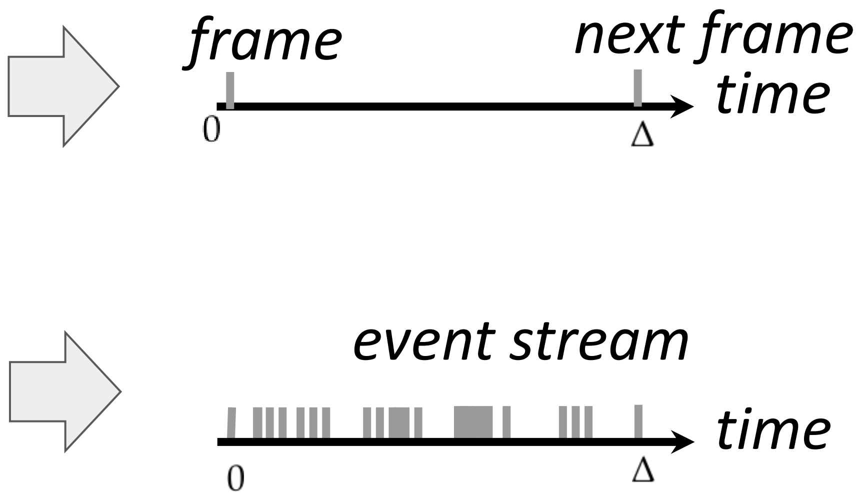
Yes, using event cameras

What is an Event Camera?

- It is camera that measures only **motion in the scene**
- **Key advantages:**
  1. **Low-latency** (~ 1 μs)
  2. **No motion blur**



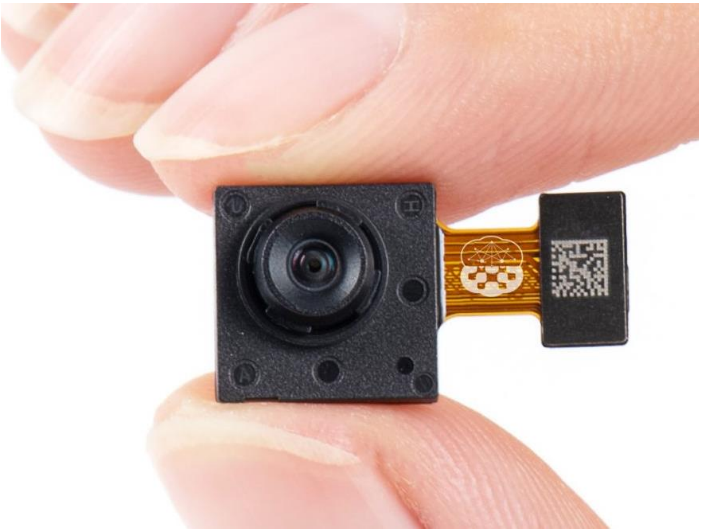




[1] Lichtsteiner, Posch, Delbruck, *A 128x128 120 dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor*, IEEE Journal of Solid-State Circuits, 2008. PDF
[2] Gallego et al., *Event-based Vision: A Survey*, T-PAMI, 2020. PDF.

Opportunities

- **Low latency**: AR/VR, automotive, robotics (<10ms)
- **Low energy**: AR/VR, always-on devices (see Synsense)
- **Low memory storage**: AR/VR, automotive, robotics
- **HDR & No motion blur**

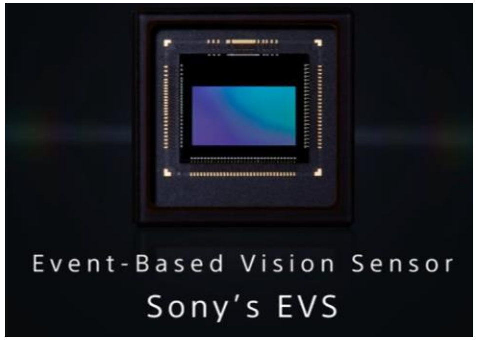


Event camera + Speck spiking-network neuromorphic processor from Synsense: can recognize faces at 50Hz while consuming < 1mW (also demoed at CVPR 2019)

Who sells event cameras and how much are they?

- Prophesee & SONY:
  - **ATIS sensor: events**, **IMU**, absolute intensity at the event pixel
  - Resolution: **1M pixels**
  - **Cost: ~5,000 USD**
- Inivation & Samsung
  - **DAVIS sensor**: **frames**, **events**, **IMU**.
  - Resolution: **VGA** (640x480 pixels)
  - **Cost: ~5,000 USD**
- CelePixel Technology & Omnivision:
  - **Celex One:** events, IMU, absolute intensity at the event pixel
  - Resolution: **1M pixels**
  - **Cost: ~1,000 USD**
- **Cost to sink to <5$** when killer application found (recall first ToF camera (>10,000 USD) today <5 USD), e.g., Samsung SmartThings Vision sensor

SONY

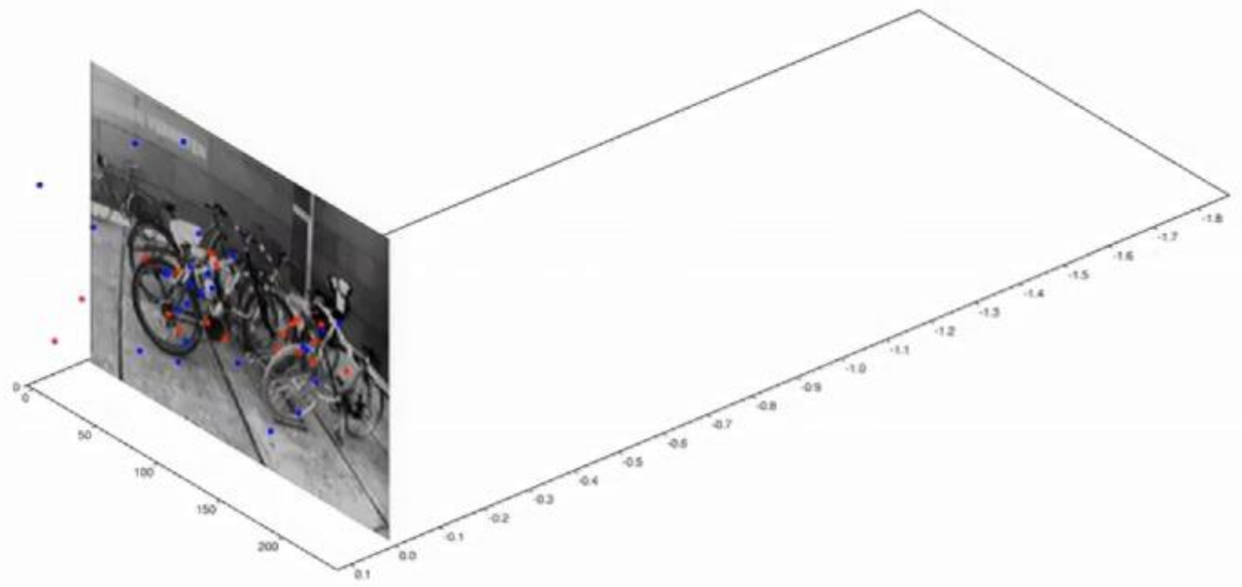


SAMSUNG

SAMSUNG

SmartThings Vision

Event Cameras enable Continuous-Time Visual Measurements in the Blind Time between frames

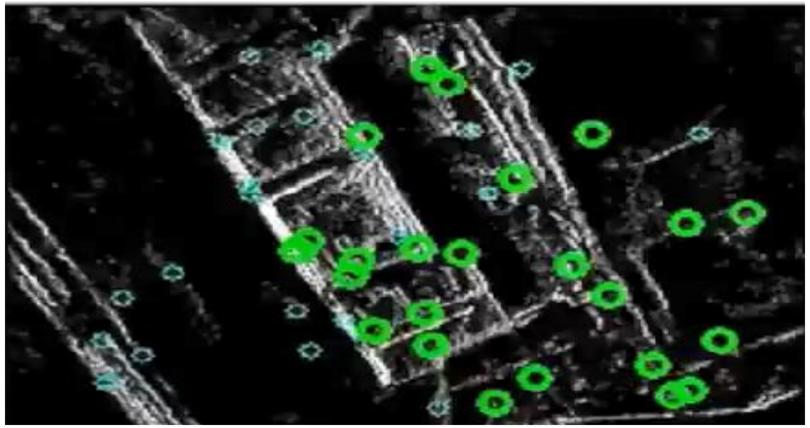


- [1] Mueggler et al., *Continuous-Time Visual-Inertial Odometry for Event Cameras*, **TRO'18**. PDF
- [2] Rosinol et al., *Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM in HDR and High Speed Scenarios,* **RAL'18 Best Paper Award Honorable Mention** PDF. Video. IEEE Spectrum.
- [3] Gehrig et al., EKLT: Asynchronous, Photometric Feature Tracking using Events and Frames, IJCV 2019. PDF, YouTube, Evaluation Code, Tracking Code

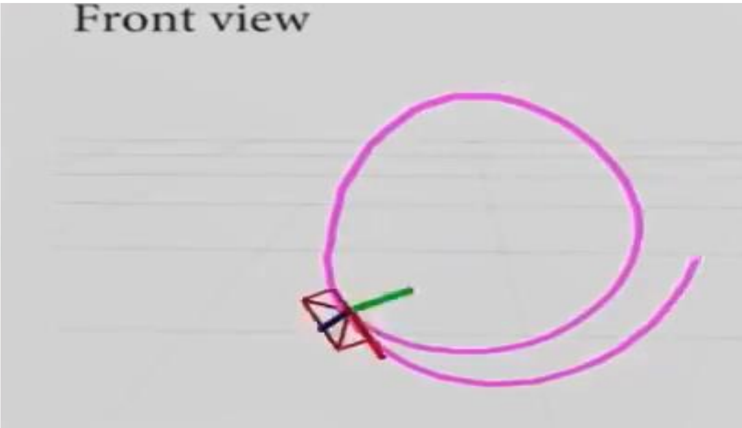
Application 1: SLAM in high-speed scenarios

Standard camera

Event camera



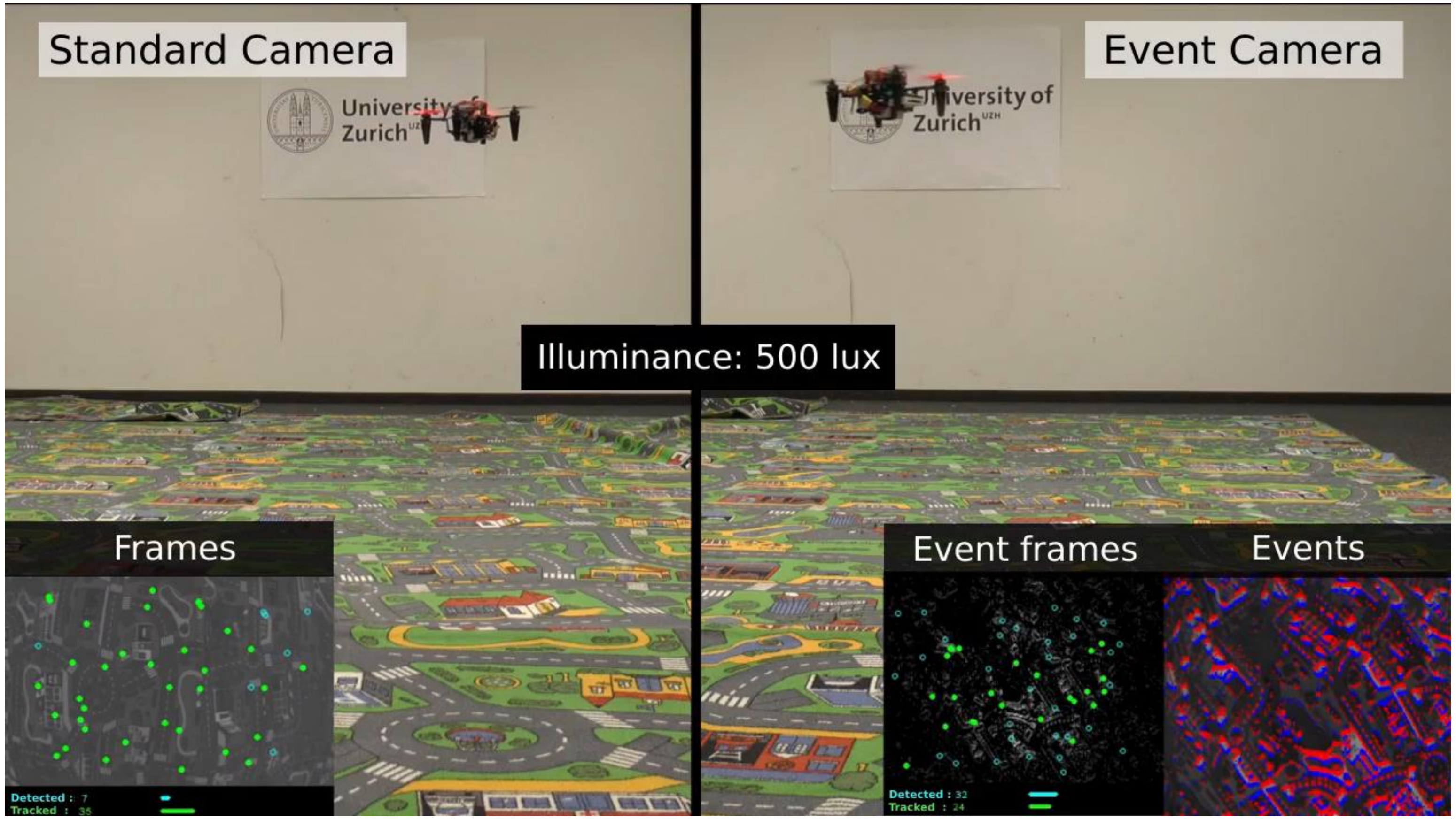
Estimated trajectory

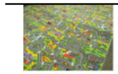
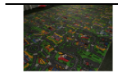
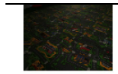


[1] Rosinol et al., *Ultimate SLAM? Combining Events, Images, and IMU for Robust Visual SLAM,* **RAL'18 Best Paper Award Hon. Mention.** PDF. Video. IEEE Spectrum.
[2] Mueggler et al., *Continuous-Time Visual-Inertial Odometry for Event Cameras*, **T-RO'18**. PDF

Application 2: Keeping drones Flying when a **Rotor Fails**

- Quadrotors subject to full rotor failure require accurate position estimates to avoid crashing
- SOTA systems used external position tracking systems (e.g., GPS, Vicon, UWB)
- We achieve this with **only onboard cameras. With event cameras**, we can make it work in very low light!

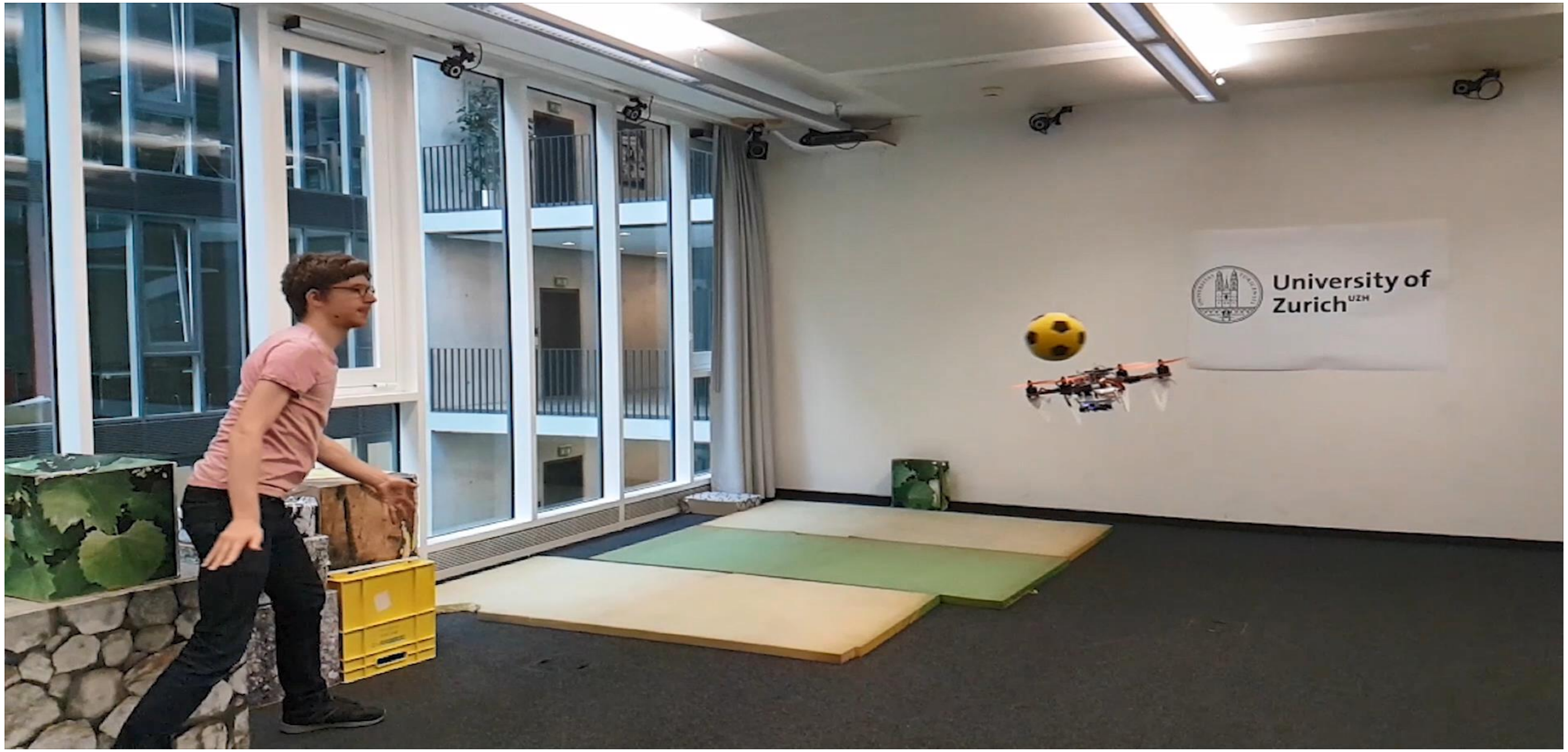


Environment Illuminance			Standard Frames		Events	
	Env lux	Cam lux	Exposure (ms)	RMSE (m)	Num of events	RMSE (m)
	500	71.5	2	**0.50**	15000	**0.48**
	100	18.0	8	**0.93**	15000	**0.42**
	50	7.3	12	∞	6000	**0.58**
	10	2.3	12	∞	6000	**0.89**
	1	0.2	12	∞	6000	∞

Sun, Cioffi, de Visser, Scaramuzza, *Autonomous Quadrotor Flight despite Rotor Failure with Onboard Vision Sensors: Frames vs. Events*, IEEE RAL'2021.
PDF. Video. Code. **1st place winner of the NASA TechBrief Award** (out of 700 participants)

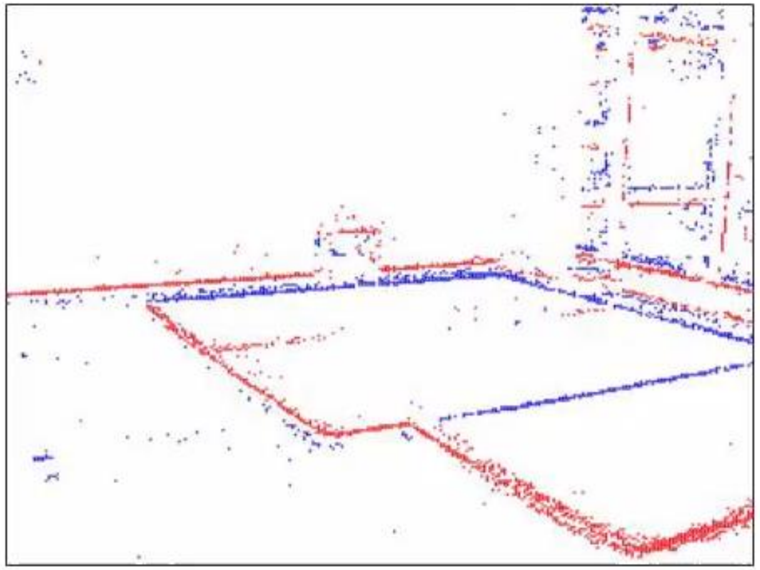
Application 3: Dodging Dynamic Objects

- Perception latency: **3.5 ms**
- Works with relative speeds of up to **10 m/s**

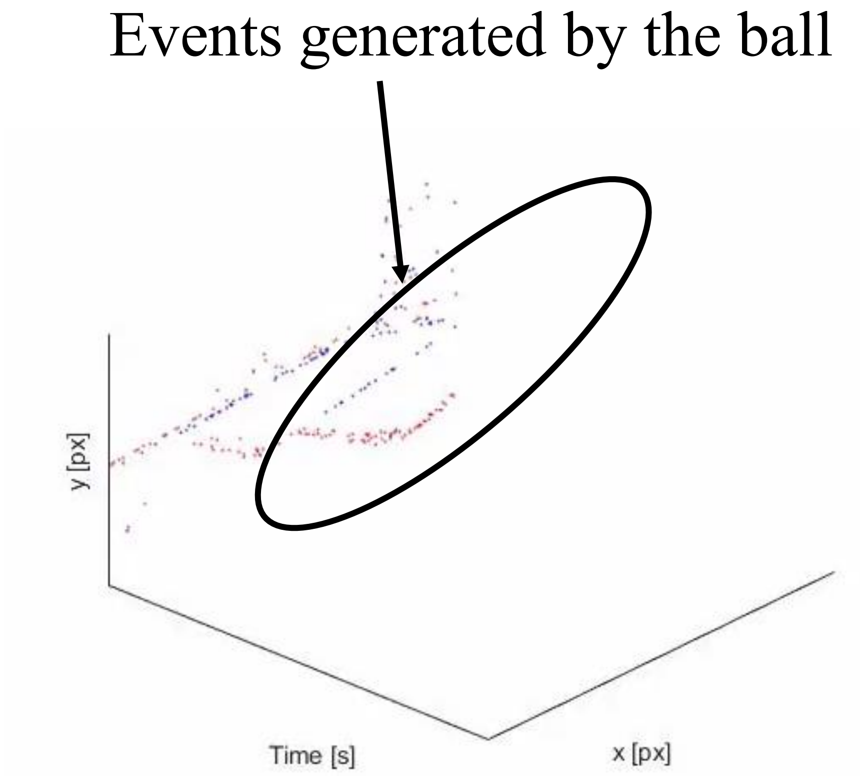


[1] Falanga et al., *Dynamic Obstacle Avoidance for Quadrotors with Event Cameras*, **Science Robotics,** 2020. PDF. Video. Featured in IEEE Spectrum
[2] Sanket et al., *EVDodgeNet: Deep Dynamic Obstacle Dodging with event cameras*, ICRA'20, PDF. Video. Code.
[3] Falanga et al. *How Fast is too fast? The Role of Perception Latency in High-Speed Sense and Avoid*, IEEE RAL'19. PDF. Video.

Events generate surfaces in the **space-time** domain



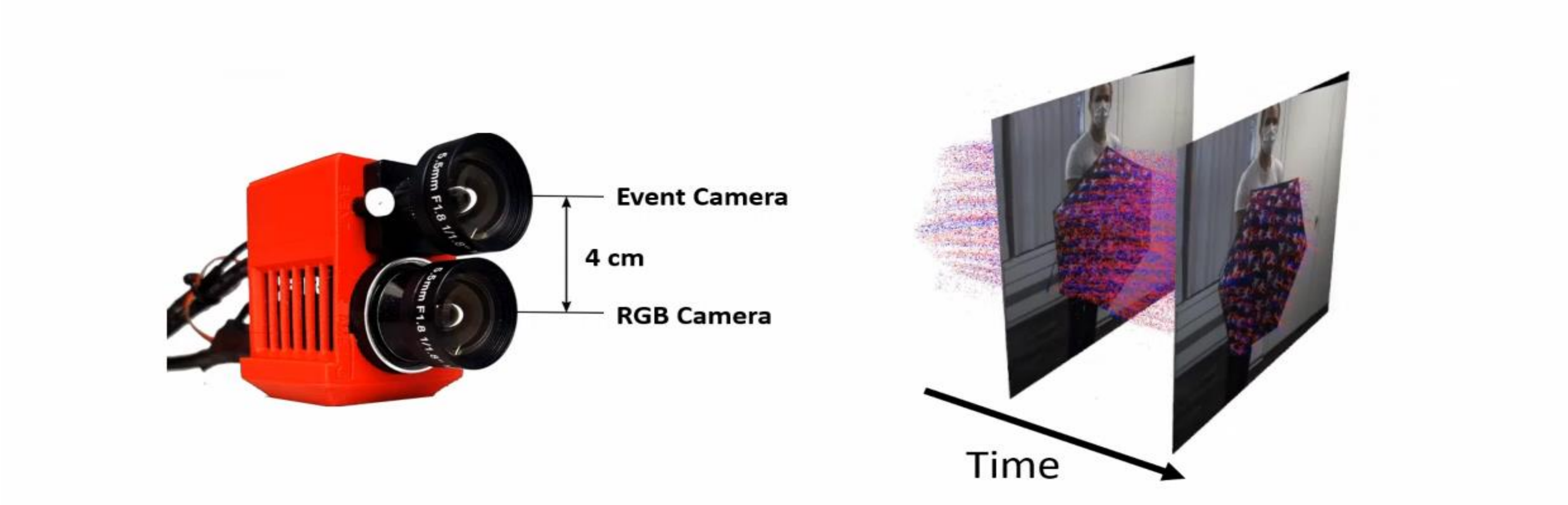
Events in the **image domain** (x, y)



Events in the **space-time** domain $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$

Application 4: Slow Motion Video

- Goal: **Upsample low-framerate RGB video using events**
- Results: we achieve **50-times upsampling** with **only 1/40th of the memory** footprint!

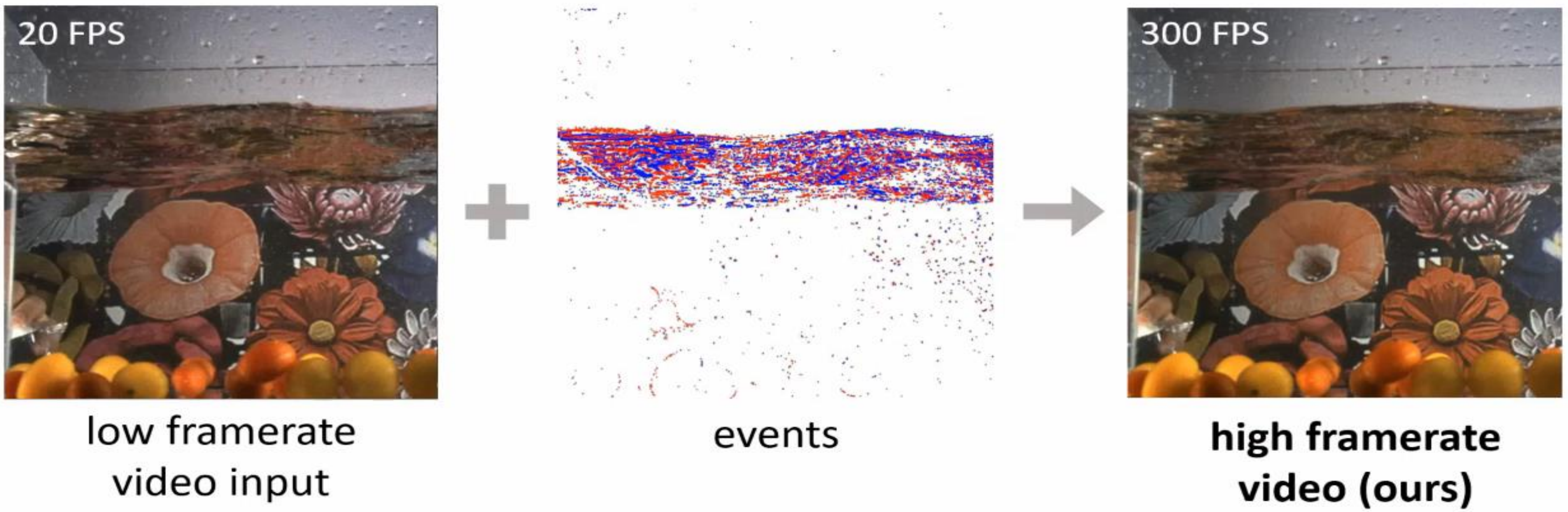


Code & Datasets: http://rpg.ifi.uzh.ch/timelens

Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21. PDF. Video. Code. Featured on Two-Minute-Papers: Video.

Application 4: Slow Motion Video

- Goal: **Upsample low-framerate RGB video using events**
- Results: we achieve **50-times upsampling** with **only 1/40th of the memory** footprint!

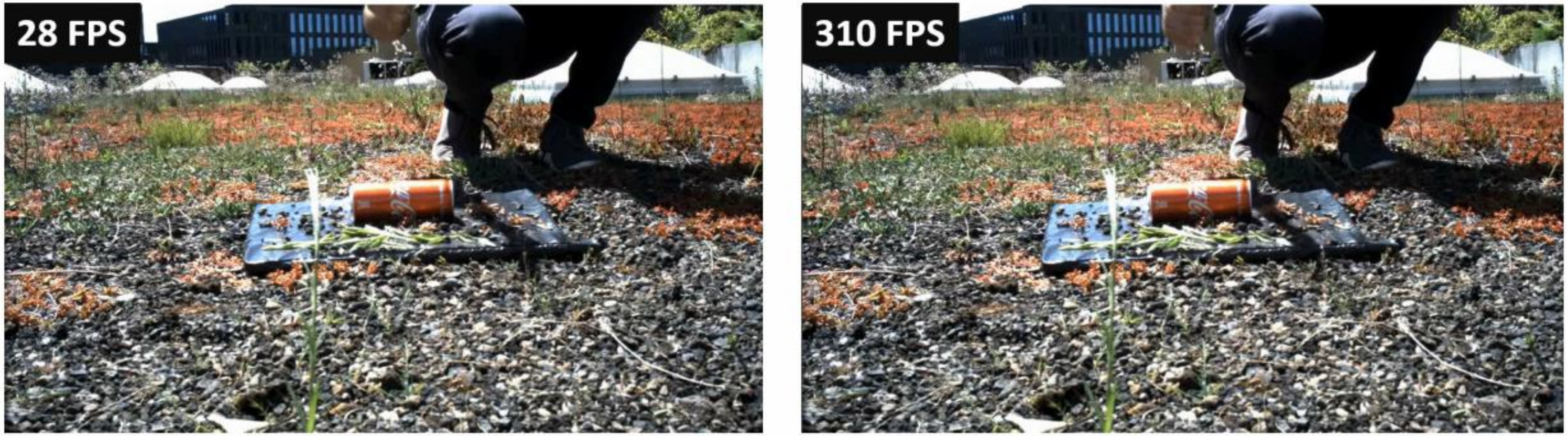


Code & Datasets: http://rpg.ifi.uzh.ch/timelens

Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21. PDF. Video. Code. Featured on Two-Minute-Papers: Video.

Application 4: Slow Motion Video

- Goal: **Upsample low-framerate RGB video using events**
- Results: we achieve **50-times upsampling** with **only 1/40th of the memory** footprint!
- **Outperforms SOTA methods using standard cameras** (e.g., DAIN)

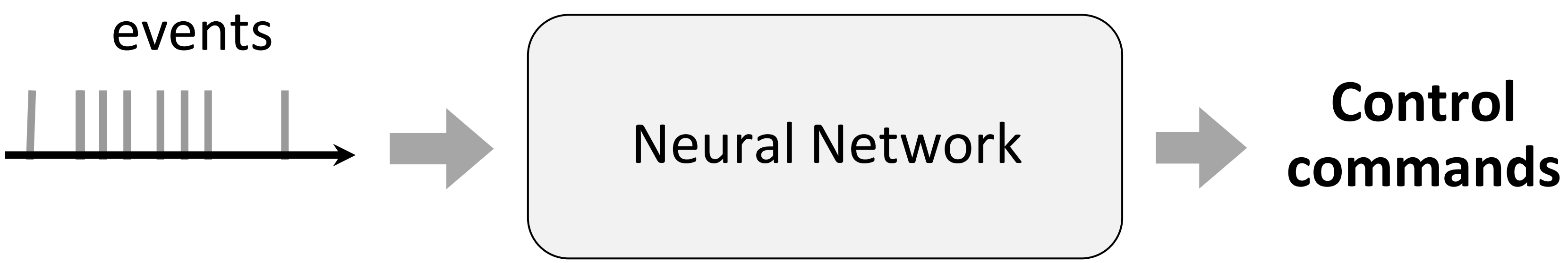


low framerate video input | **Time Lens (this work)**

Code & Datasets: http://rpg.ifi.uzh.ch/timelens

Tulyakov et al., TimeLens: *Event-based Video Frame Interpolation*, CVPR'21. PDF. Video. Code. Featured on Two-Minute-Papers: Video.

Outlook: Event-driven Control on Neuromorphic Processors



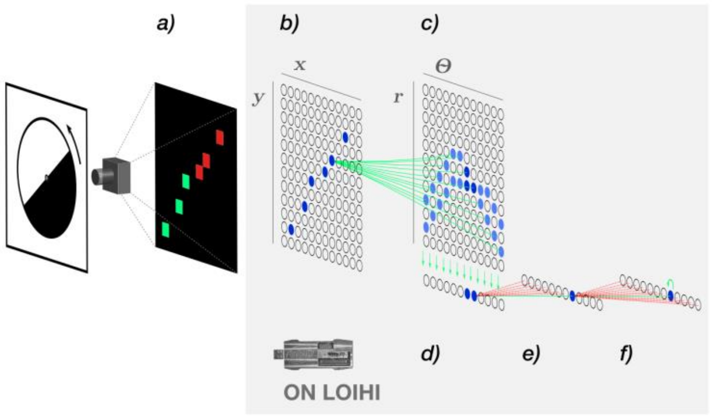
- Motivation: agile maneuvers require **low perception latency** and **high controller bandwidth**
- Goal: **map single events directly to control commands**
- Method: **Spiking Network** (SNN) running on **Intel Loihi neuromorphic chip**
- Advantage: **low-latency perception, high-bandwidth control**

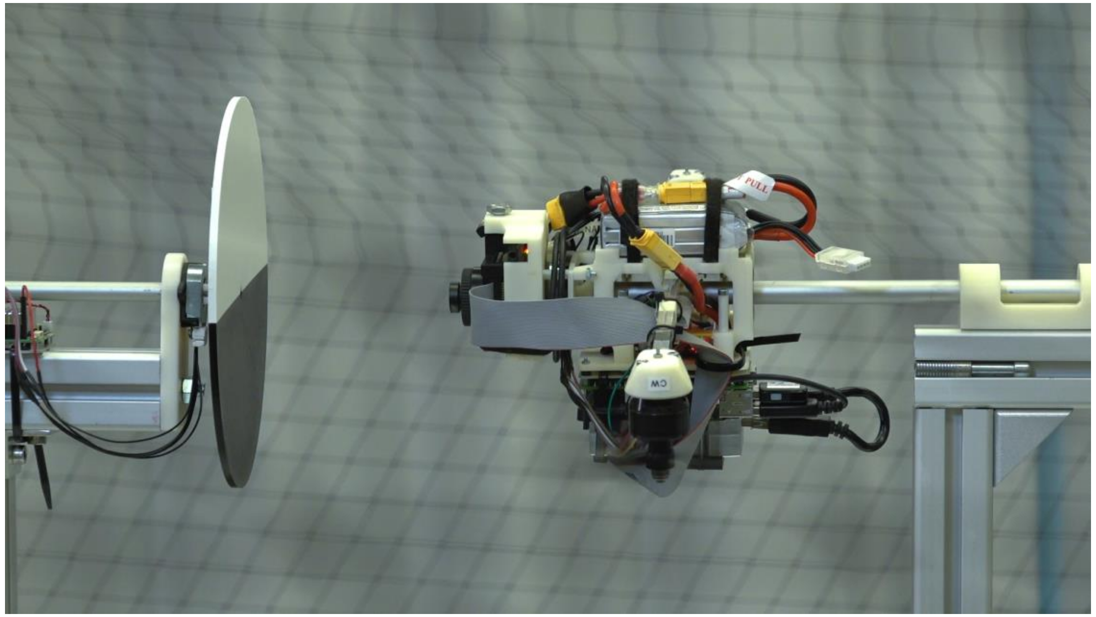
[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA'21. PDF. Video.
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA'20, PDF YouTube

Outlook: Event-driven Control on Neuromorphic Processors

events → Neural Network → **Control commands**

Reflex control task: mimic 1D rotation of a disc





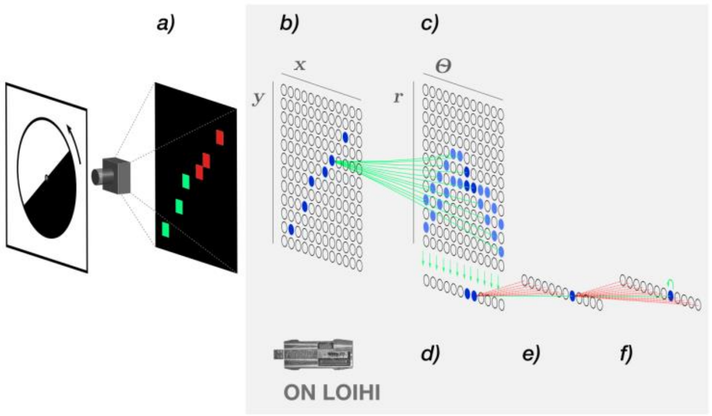
[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA'21. PDF. Video.

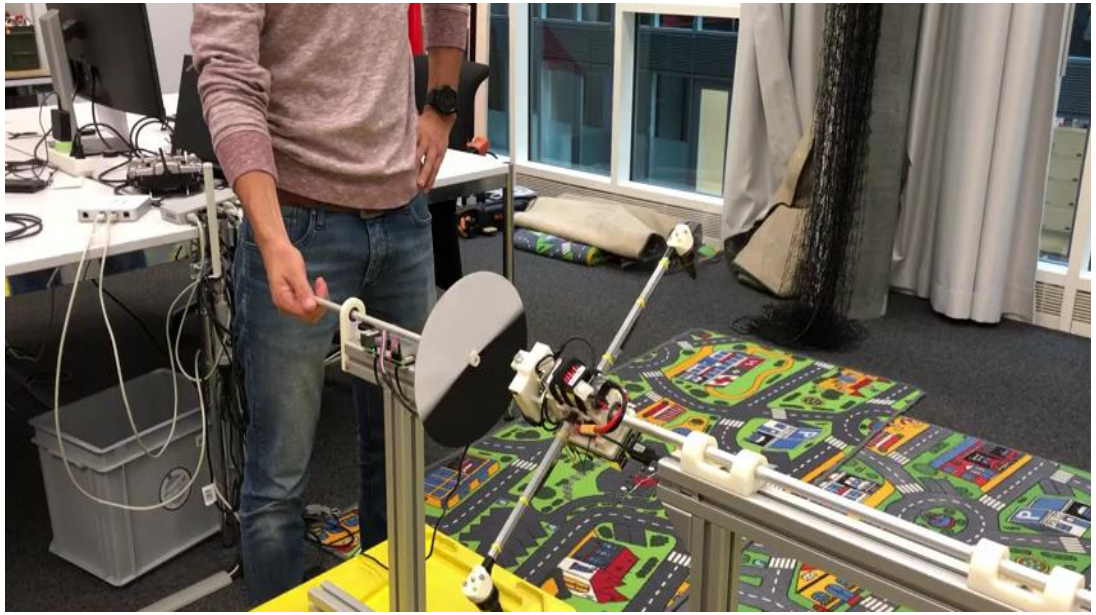
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA'20, PDF YouTube

Outlook: Event-driven Control on Neuromorphic Processors

events → Neural Network → **Control commands**

Reflex control task: mimic 1D rotation of a disc



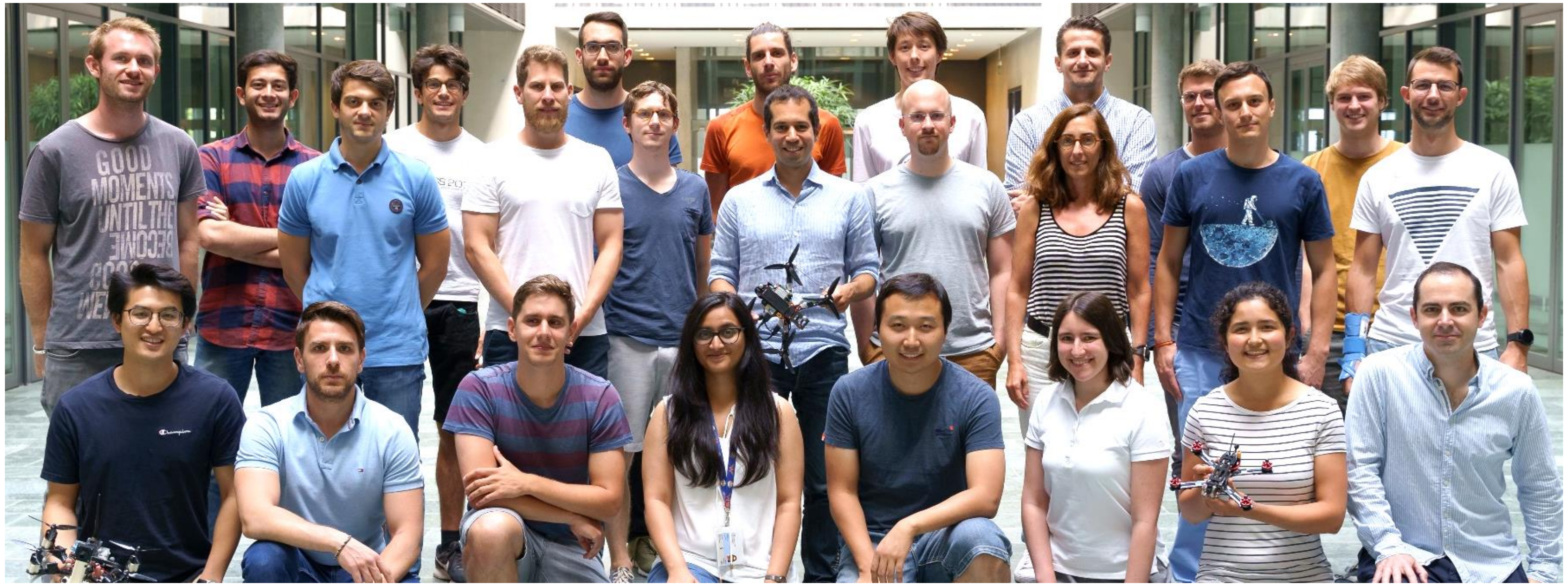


[1] Vitale et al., *Event-driven Vision and Control for UAVs on a Neuromorphic Chip*, ICRA'21. PDF. Video.
[2] Sugimoto et al., *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, ICRA'20, PDF YouTube

Conclusions and Takeaways

- Autonomous **vision-based agile flight** as a new research topic (at least 10 years to solve it)
  - **Pushes the limit of existing algorithms** in extreme situations
  - Raises **fundamental problems** for robotics **research**
- **Learning**-based **methods** are **more robust to imperfect perception** and exhibit **lower latency** than traditional ones
- **Simulation to Reality** transfer possible with **appropriate input/output abstraction.**
- **Event cameras** significantly **reduce perception latency** and **motion blur** and, if coupled with **neuromorphic chips**, allow **low-latency, high-bandwidth control**

Thanks!



Code, **datasets**, **videos**, and **publications**, **slides**: http://rpg.ifi.uzh.ch/

I am hiring PhD students and Postdocs in AI

ailabRPG @davsca1 @davidescaramuzza

This drone uses AI to race against a human

Goal: beat a human in a time race by passing through a sequence of gates in a given order in the least possible time.

AI Drone

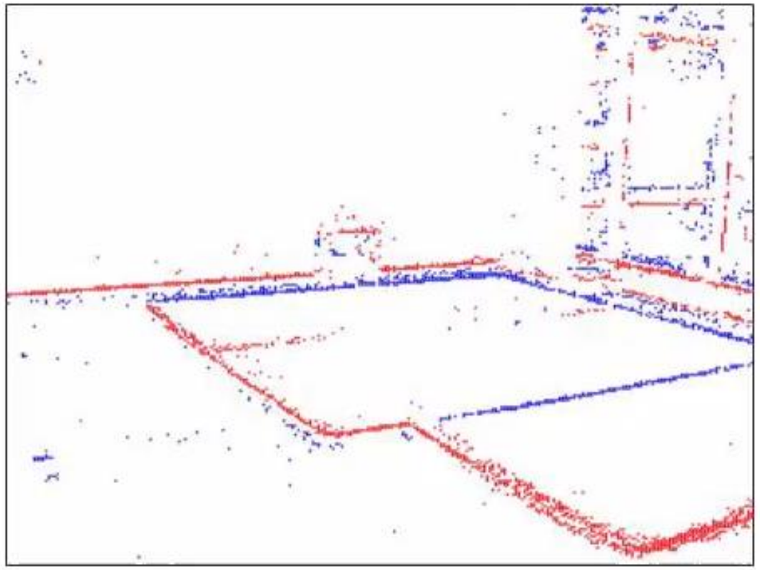
Human

Marvin is a 15-y.o.
world class drone-racing pilot

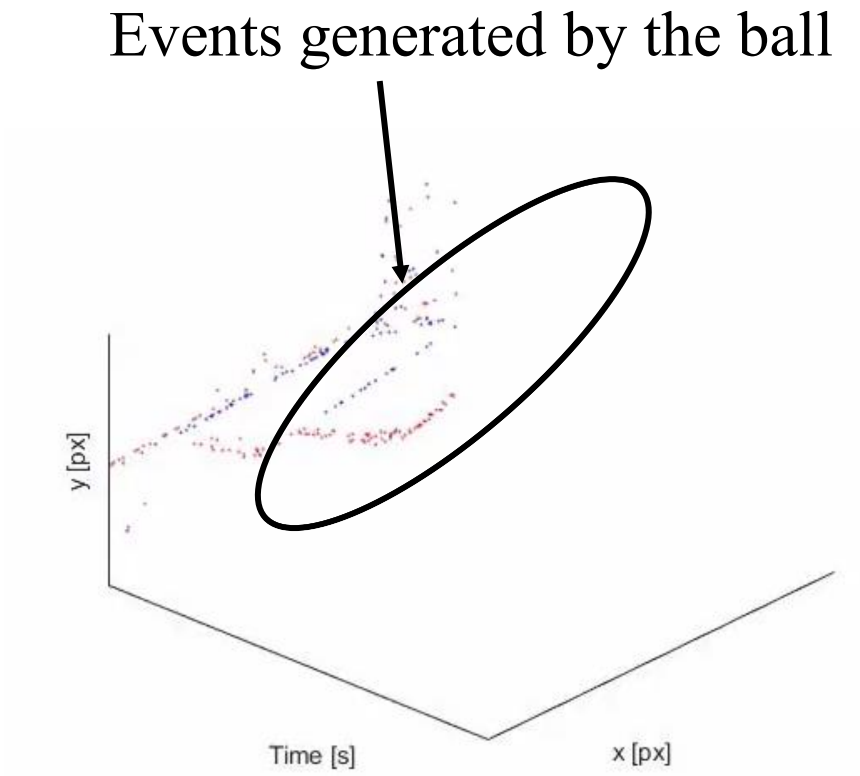
Foehn et al., *Time-Optimal Planning for Quadrotor Waypoint Flight*, **Science Robotics**, 2021. PDF. Video. Code. **Featured on Forbes magazine**.


Levi's®

Events generate surfaces in the **space-time** domain



Events in the **image domain** (x, y)



Events in the **space-time** domain $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t})$

64mW, 27gram, 8cm size, GAP8-powered Autonomous Drone

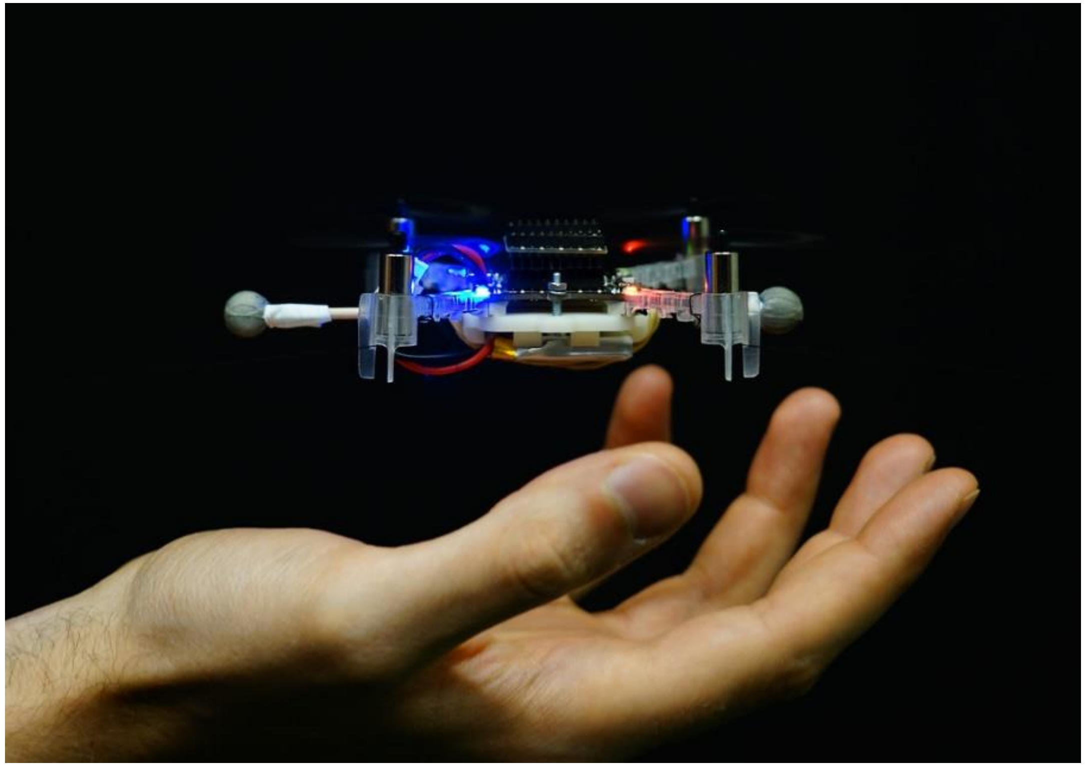
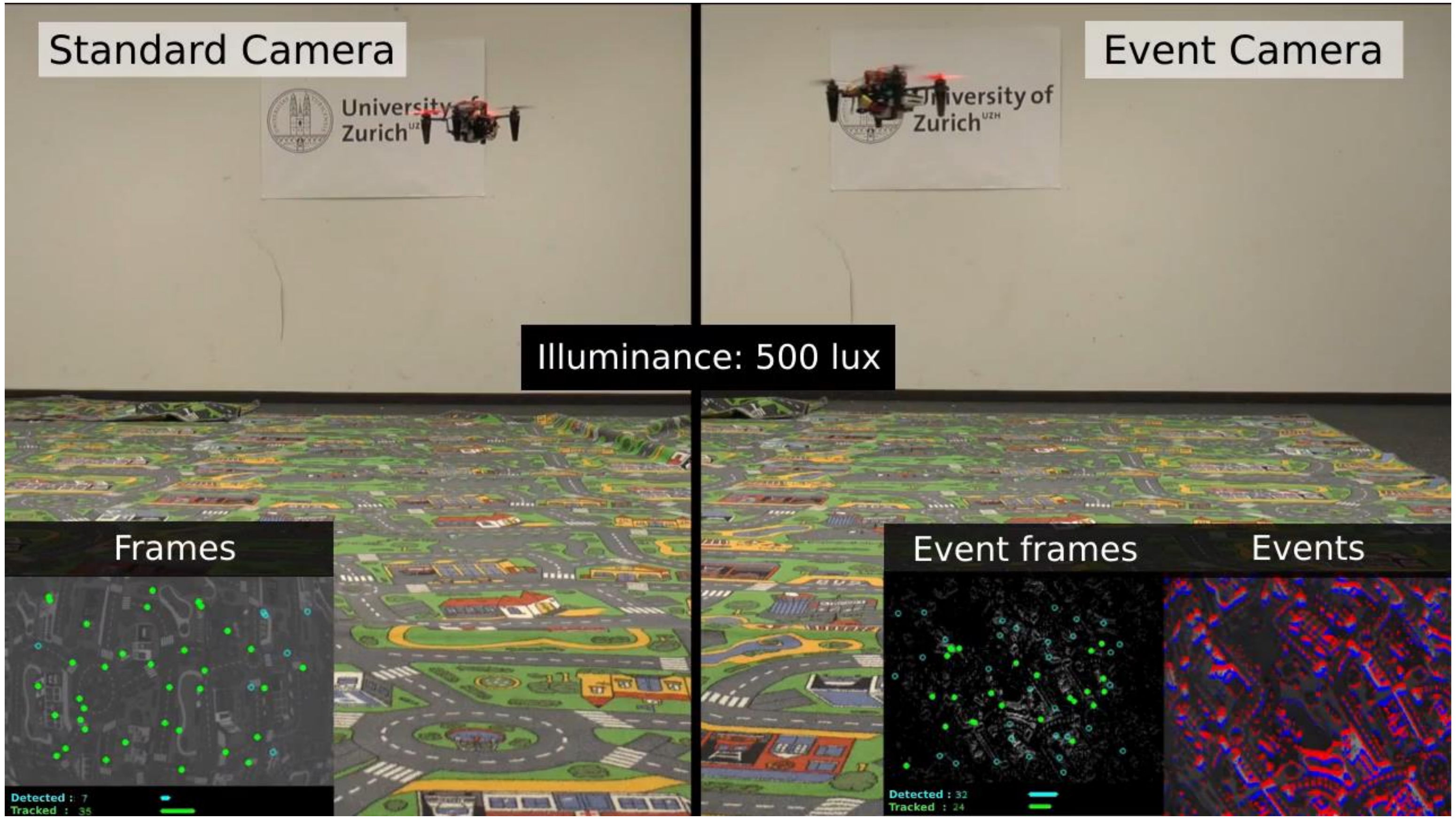


Image credit: Paul Beuchat

Palossi, Loquercio, Conti, Flamand, Scaramuzza, Benini, A 64mW DNN-based Visual Navigation Engine for Autonomous Nano-Drones
IEEE Internet of Things Journal, 2019. PDF. Video. Code.

Application 2: Autonomous Flight despite **Rotor Failure**

- Quadrotors subject to full rotor failure require accurate position estimates to avoid crashing
- SOTA systems used external position tracking systems (e.g., GPS, Vicon, UWB)
- We achieve this with **only onboard cameras. With event cameras**, we can make it work in very low light!



Environment Illuminance			Standard Frames		Events	
	Env lux	Cam lux	Exposure (ms)	RMSE (m)	Num of events	RMSE (m)
	500	71.5	2	**0.50**	15000	**0.48**
	100	18.0	8	**0.93**	15000	**0.42**
	50	7.3	12	∞	6000	**0.58**
	10	2.3	12	∞	6000	**0.89**
	1	0.2	12	∞	6000	∞

Sun, Cioffi, de Visser, Scaramuzza, *Autonomous Quadrotor Flight despite Rotor Failure with Onboard Vision Sensors: Frames vs. Events*, IEEE RAL'2021.
PDF. Video. Code. **1st place winner of the NASA TechBrief Award** (out of 700 participants)