University of Zurich[UZH] | ETH zürich

Institute of Informatics – Institute of Neuroinformatics

ROBOTICS & PERCEPTION GROUP

Lecture 10
Multiple View Geometry 4

Davide Scaramuzza

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

This afternoon: Intermediate VO Integration

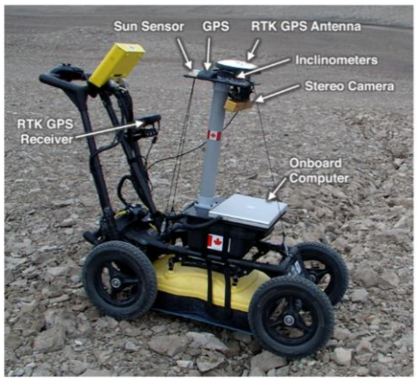
19.09.2019	**Lecture 01 - Introduction to Computer Vision and Visual Odometry**	Davide Scaramuzza
26.09.2019	**Lecture 02 - Image Formation 1: perspective projection and camera models** **Exercise 01 - Augmented reality wireframe cube**	Davide Scaramuzza Daniel & Mathias Gehrig
03.10.2019	**Lecture 03 - Image Formation 2: camera calibration algorithms** **Exercise 02 - PnP problem**	Davide Scaramuzza Daniel & Mathias Gehrig
10.10.2019	**Lecture 04 - Filtering & Edge detection**	Davide Scaramuzza
17.10.2019	**Lecture 05 - Point Feature Detectors, Part 1** **Exercise 03 - Harris detector + descriptor + matching**	Davide Scaramuzza Daniel & Mathias Gehrig
24.10.2019	**Lecture 06 - Point Feature Detectors, Part 2** **Exercise 04 - SIFT detector + descriptor + matching**	Davide Scaramuzza Daniel & Mathias Gehrig
31.10.2019	**Lecture 07 - Multiple-view geometry** **Exercise 05 - Stereo vision: rectification, epipolar matching, disparity, triangulation**	Davide Scaramuzza Daniel & Mathias Gehrig
07.11.2019	**Lecture 08 - Multiple-view geometry 2** **Exercise 06 - Eight-Point Algorithm**	Antonio Loquercio Daniel & Mathias Gehrig
14.11.2019	**Lecture 09 - Multiple-view geometry 3 (Part 1)**	Davide Scaramuzza
21.11.2019	**Lecture 10 - Multiple-view geometry 3 (Part 2)** **Exercise session: Intermediate VO Integration**	Davide Scaramuzza Daniel & Mathias Gehrig
28.11.2019	**Lecture 11 - Optical Flow and Tracking (Lucas-Kanade)** **Exercise 08 - Lucas-Kanade tracker**	Davide Scaramuzza Daniel & Mathias Gehrig
05.12.2019	**Lecture 12 - Place recognition and 3D Reconstruction** **Exercise session: Deep Learning Tutorial**	Davide Scaramuzza Daniel & Mathias Gehrig
12.12.2019	**Lecture 13 - Visual inertial fusion** **Exercise 09 - Bundle Adjustment**	Davide Scaramuzza Daniel & Mathias Gehrig
19.12.2019	**Lecture 14 - Event based vision** After the lecture, we will Scaramuzza's lab. Departure from lecture room at 12:00 via tram 10. **Exercise session: Final VO Integration**	Davide Scaramuzza Daniel & Mathias Gehrig

How can we evaluate VO/SLAM algorithms?

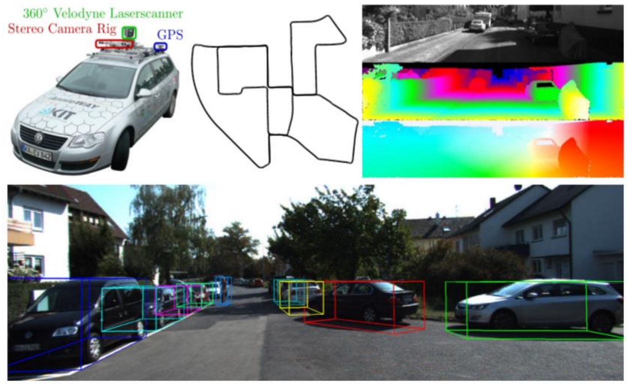
This problem is known as “Benchmarking”

Popular Datasets for VO/SLAM Benchmarking

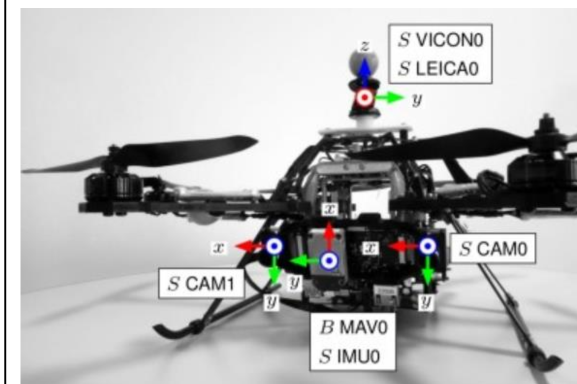
Devon Island [Furgale'11]
Stereo + D-GPS + inclinometer + sun sensor



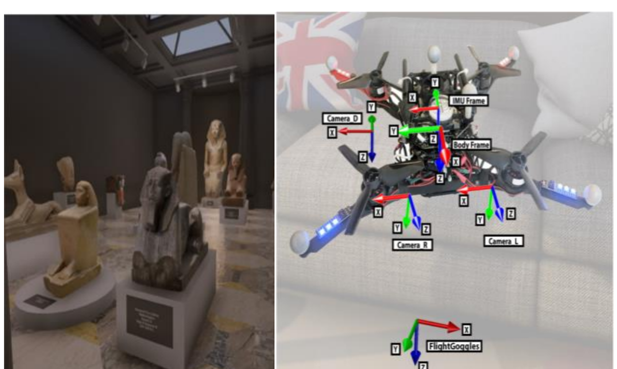
KITTI [Geiger'12]
Automobile, Laser + stereo + GPS, multiple tasks



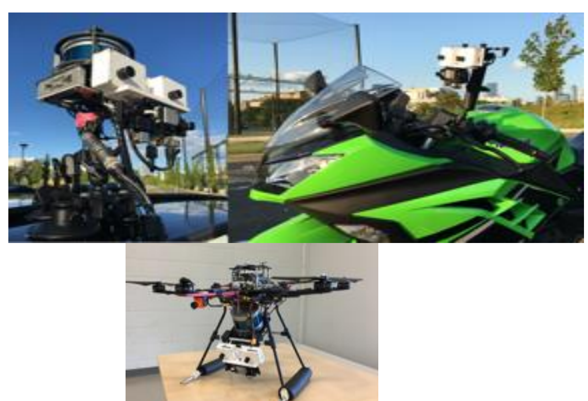
EuRoC [Burri'16]
MAV with synchronized IMU and stereo



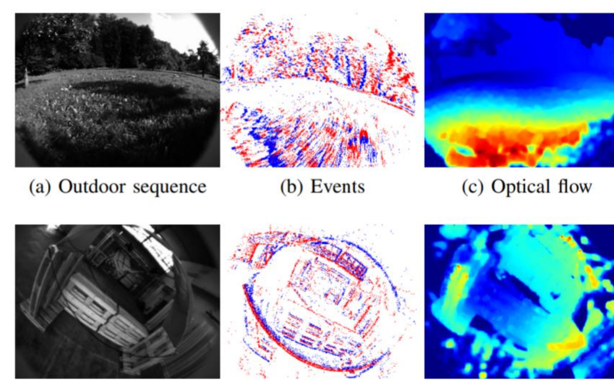
Blackbird [Antonini'18]
MAV indoor aggressive flight with rendered images and real dynamics + IMU



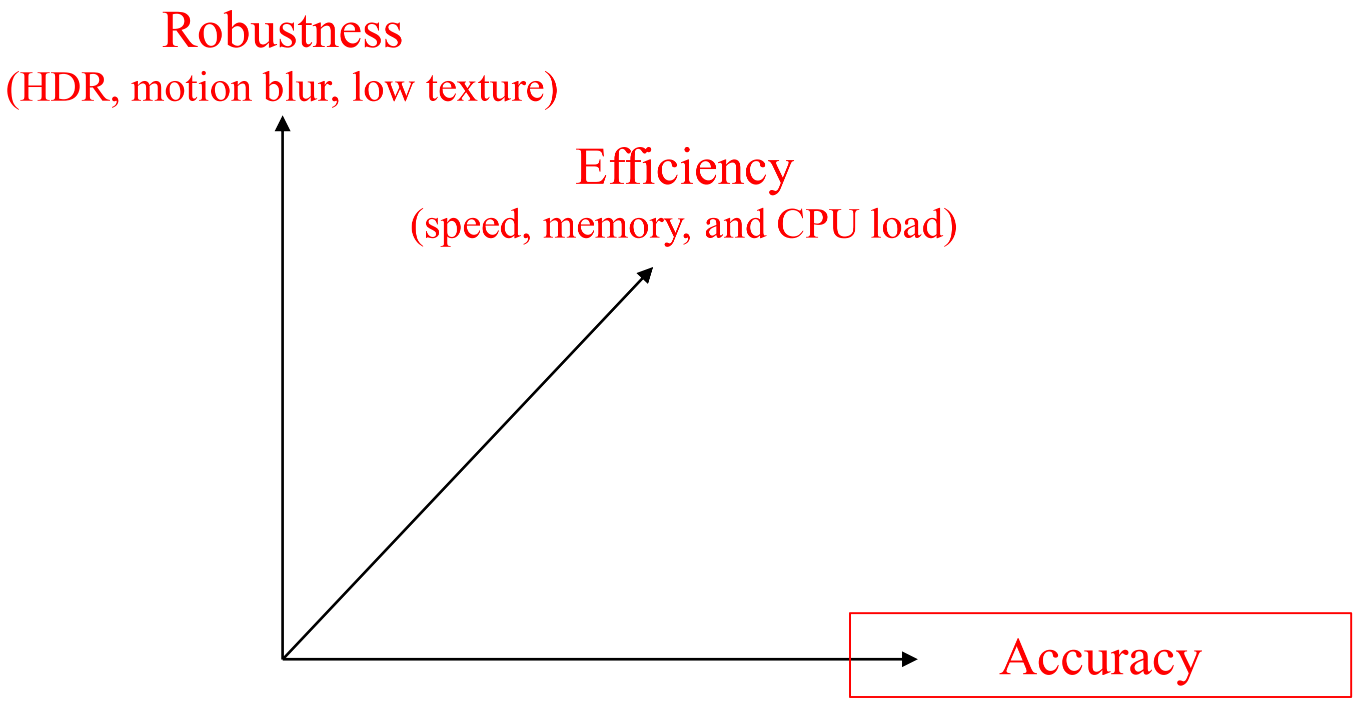
MVSEC [Zhu'18]
Events, frames, lidar, GPS, IMU from cars, drones, and motorcycles



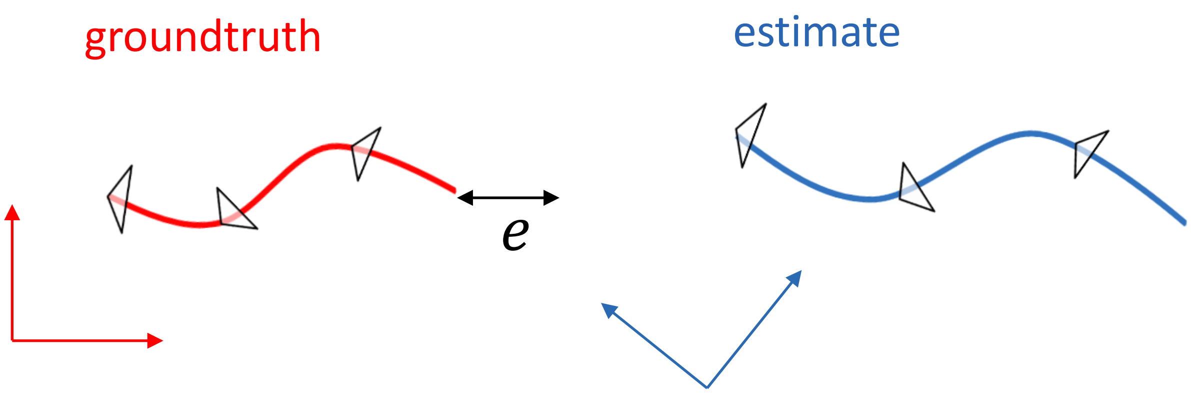
UZH Drone Racing [Delmerico'19]
MAV aggressive flight, standard + event cameras, IMU, indoors and outdoors


(a) Outdoor sequence (b) Events (c) Optical flow
(d) Indoor sequence (e) Events (f) Optical flow

What metrics should be used?



Evaluation is a non-trivial task...



Direct difference?

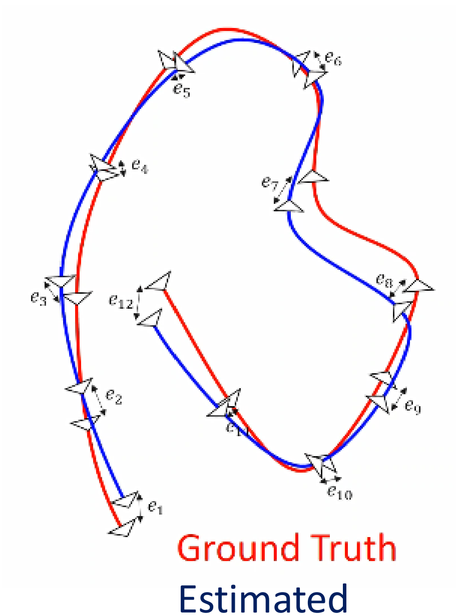
- Different reference frame
- Different scale
- Different times stamps
- ...

Maybe align the first poses and measure the **end-pose error**?

- **How many poses** should be used for the alignment?
- **Not robust:**
  - **Most VIOs are non-deterministic** (e.g., RANSAC, multithreading) → every time you run your VIO on the same dataset, you get different results
  - **Not meaningful:**
    - too sensitive to the trajectory shape
    - does not capture the error statistics

Metric 1: Absolute Trajectory Error (ATE)

Absolute Trajectory Error
RMSE of the aligned estimate and the groundtruth.



✓ Single number metric
✗ Many parameters to specify

Step 1: Align the trajectory

groundtruth positions

$$\underset{R,T,s}{\operatorname{argmin}} \sum_{i=0}^{N} \|\hat{t}_i - sRt_i - T\|^2$$

estimated positions

Alignment parameters

Step 2: Root mean squared errors between the aligned estimate and the groundtruth.

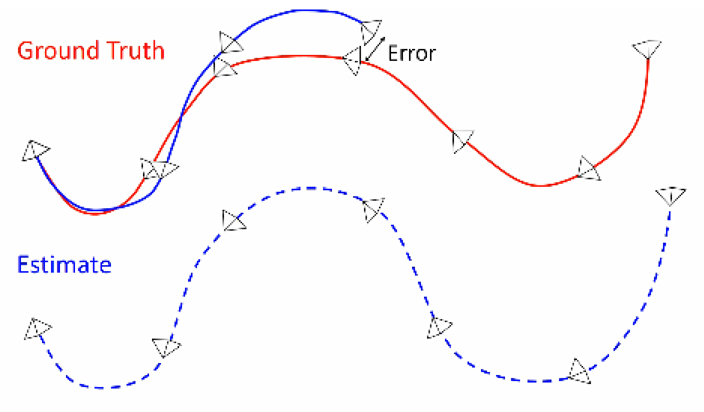
$$\sqrt{\frac{\sum_{i=1}^{N} \|\hat{t}_i - sRt_i - T\|^2}{N}}$$

- Sturm et al., "A benchmark for the evaluation of RGB-D SLAM systems." IROS 2012.
- Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS’18. PDF

Metric 2: Relative Trajectory Error (RTE)

Relative Error (Odometry Error)
Statistics of sub-trajectories of specified lengths.

➢ Calculate errors for all the subtrajectories of certain lengths.



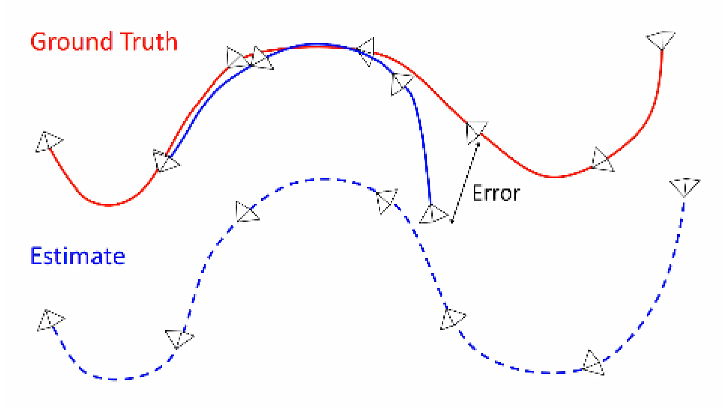
✓ Informative statistics
✗ Complicated to compute and rank

- Geiger et al. "Are we ready for autonomous driving? the KITTI vision benchmark suite." CVPR 2012.
- Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS'18. PDF

Metric 2: Relative Trajectory Error (RTE)

Relative Error (Odometry Error)
Statistics of sub-trajectories of specified lengths.

- Calculate errors for all the subtrajectories of certain lengths.



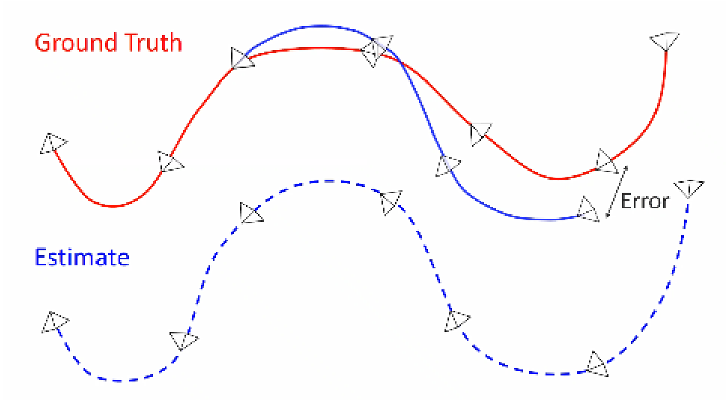
✓ Informative statistics
✗ Complicated to compute and rank

- Geiger et al. "Are we ready for autonomous driving? the KITTI vision benchmark suite." CVPR 2012.
- Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS'18. PDF

Metric 2: Relative Trajectory Error (RTE)

Relative Error (Odometry Error)
Statistics of sub-trajectories of specified lengths.

- Calculate errors for all the subtrajectories of certain lengths.

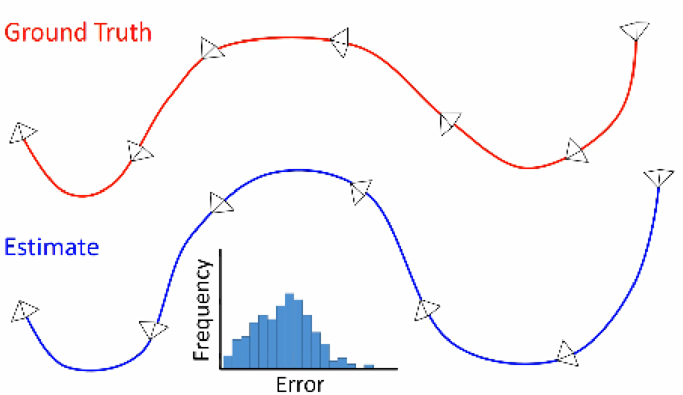


✓ Informative statistics
✗ Complicated to compute and rank

- Geiger et al. "Are we ready for autonomous driving? the KITTI vision benchmark suite." CVPR 2012.
- Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS'18. PDF

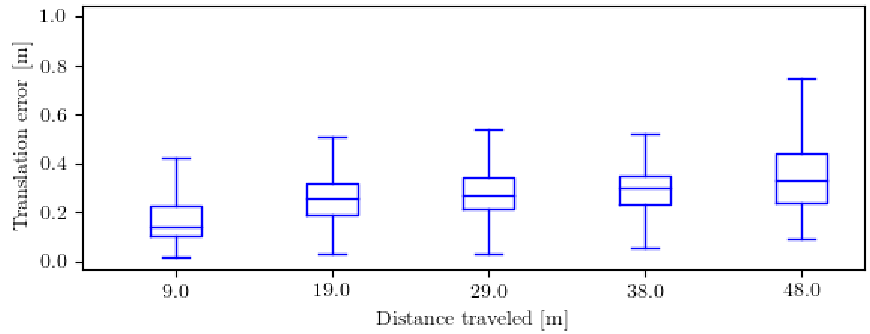
Metric 2: Relative Trajectory Error (RTE)

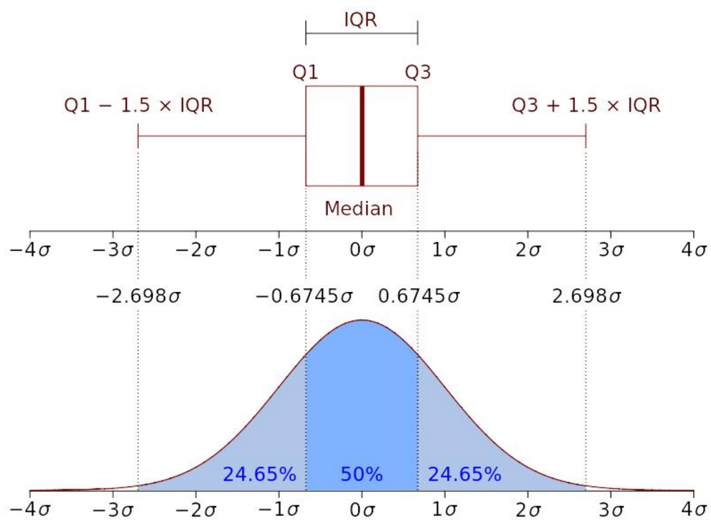
Relative Error (Odometry Error)
Statistics of sub-trajectories of specified lengths.



- ✓ Informative statistics
- ✘ Complicated to compute and rank

➢ Calculate errors for all the subtrajectories of certain lengths.

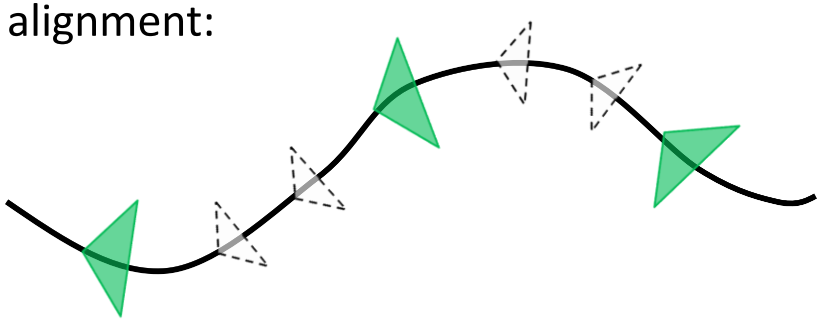




- Geiger et al. "Are we ready for autonomous driving? the KITTI vision benchmark suite." CVPR 2012.
- Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS'18. PDF

Trajectory Accuracy: Error Metrics

- Both ATE and RTE are widely used in practice, but:
  - **Many details** need to be specified which are **often omitted in papers**
    - **Number of poses** used for the alignment (also, **frames or keyframes**?)
    - **Type of transformation used** for the alignment:
      - **SE(3)** for stereo VO
      - **Sim(3)** for monocular VO
      - **4DOF** for VIO
    - **Sub-trajectory lengths** in RTE

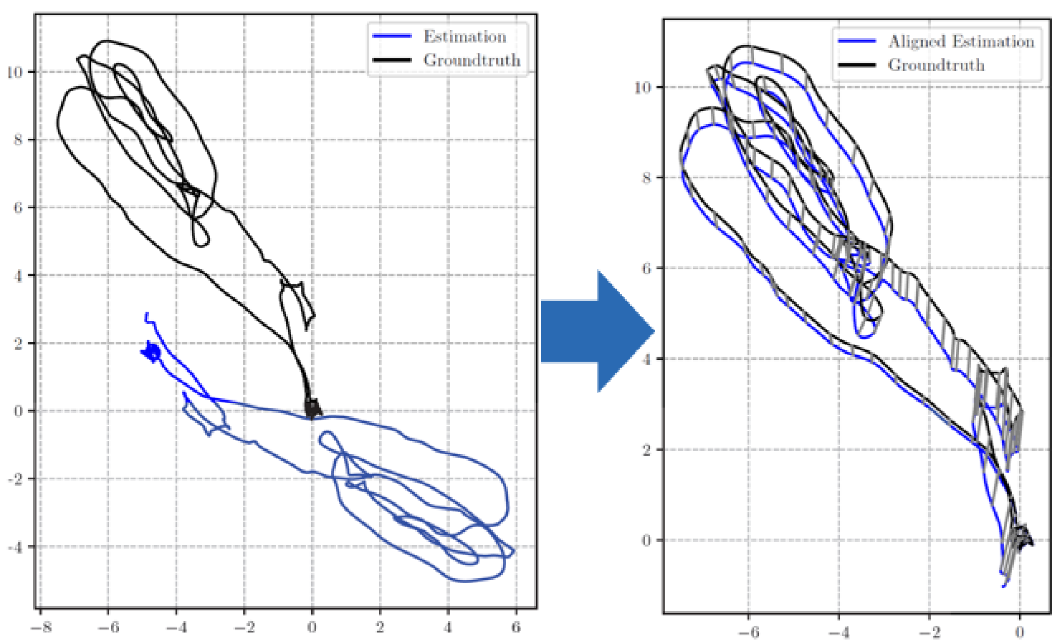


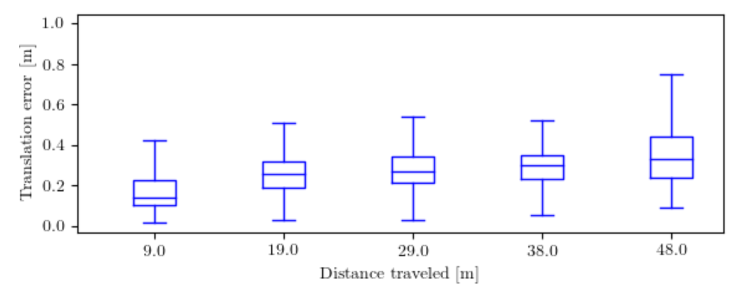
- White: **Normal frames** (used for **real time pose update**)
- Green: **Keyframes** (usually updated **after BA**)

- Results are not directly comparable with different settings
  - **Report** the **evaluation settings** in detail.
  - Is there a publicly available evaluation tool to facilitate reproducible evaluation? Yes: **Trajectory Evaluation Toolbox**

Trajectory Evaluation Toolbox

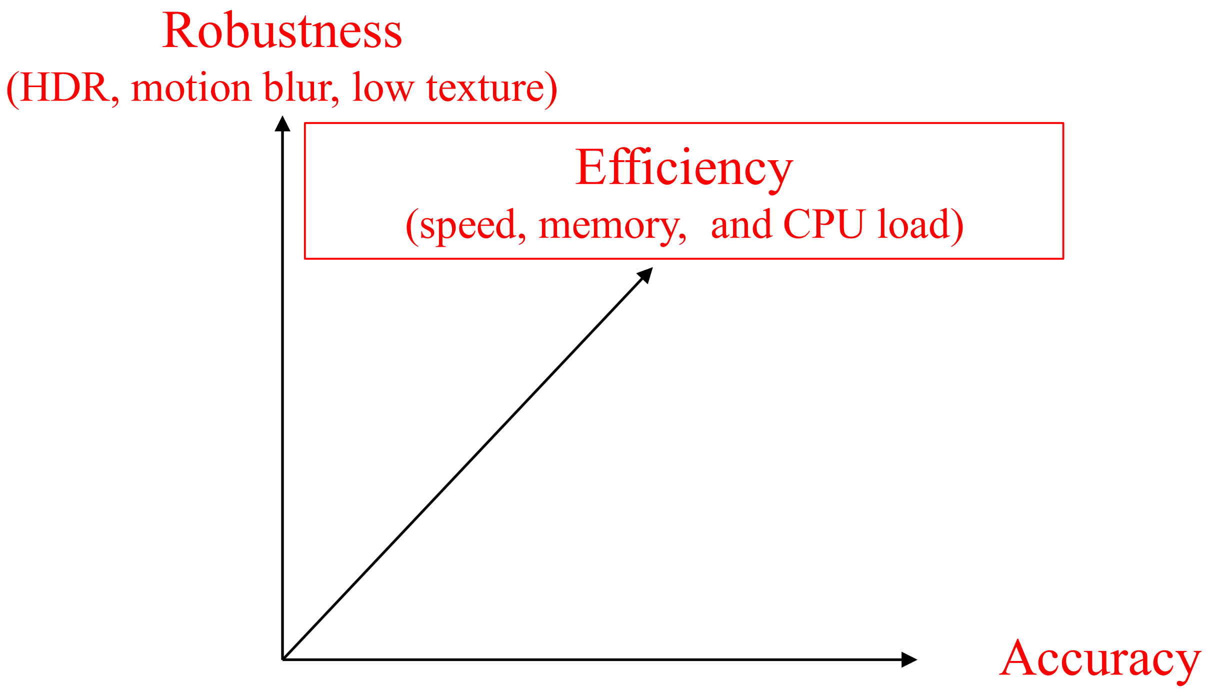
- **Designed to make trajectory evaluation easy!**
  - Implements **different alignment methods** depending on the sensing modalities: **SE(3)** for stereo, **sim(3)** for monocular, **4DOF** for VIO.
  - Implements **Absolute Trajectory Error** and **Relative Error**.
  - Automated evaluation of different algorithms on multiple datasets (for N runs).
- Code: https://github.com/uzh-rpg/rpg_trajectory_evaluation [Zhang, IROS'18]





Zhang et al., "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry." IROS'18. [PDF](PDF)

What metrics should be used?

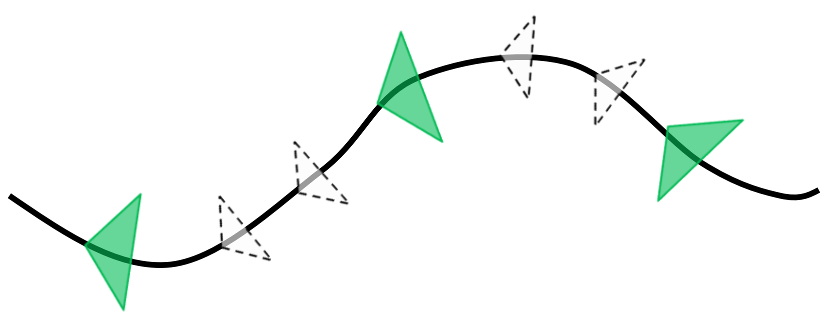


Benchmarking Efficiency

- Different computational resources
  - Memory
  - CPU load
  - Processing time

Depends not only on algorithm design, but also implementation, platforms, etc.

- There are **different definitions of processing time** in SLAM systems.



- White: Normal frames (used for real time pose update)
- Green: Keyframes (usually updated after BA)

- **Processing time for real-time pose**:
  $$t_{pose\ output} - t_{image\ arrival}$$
- **Processing time for asynchronously executed threads** (e.g., bundle adjustment)
- ……

Case study: VIO for Flying Robots [ICRA'18]

- Algorithms: **MSCKF**, **OKVIS**, **ROVIO**, **VINS-Mono**, **SVO+MSF**, **SVO+GTSAM**, VINS-Mono w/ and w/o loop closure

- Hardware: consider the limitation of flying robots

Intel NUC

Odroid XU4

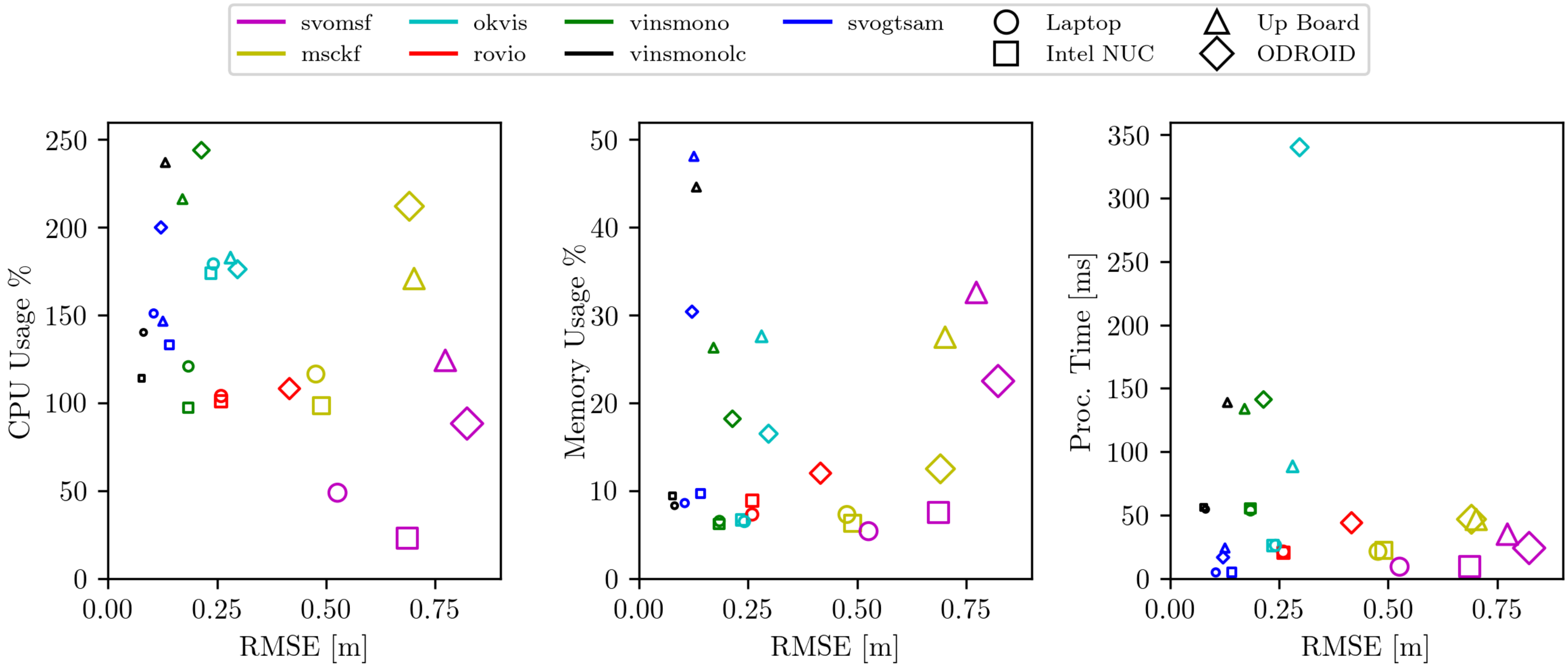
Up Board

Intel Lenovo W540 i7laptop

- Evaluation
  - **Absolute Trajectory Error (ATE)**– RMSE after sim(3) trajectory alignment (7DoF)
  - **Relative Trajectory Error (RTE)**– Error distribution of the subtrajectories
  - **CPU usage** – total load of CPU
  - **Memory usage** – total percentage of available RAM
  - **Time per frame** – from input until pose is updated

Delmerico, Scaramuzza, A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots, ICRA'18. PDF. Video.

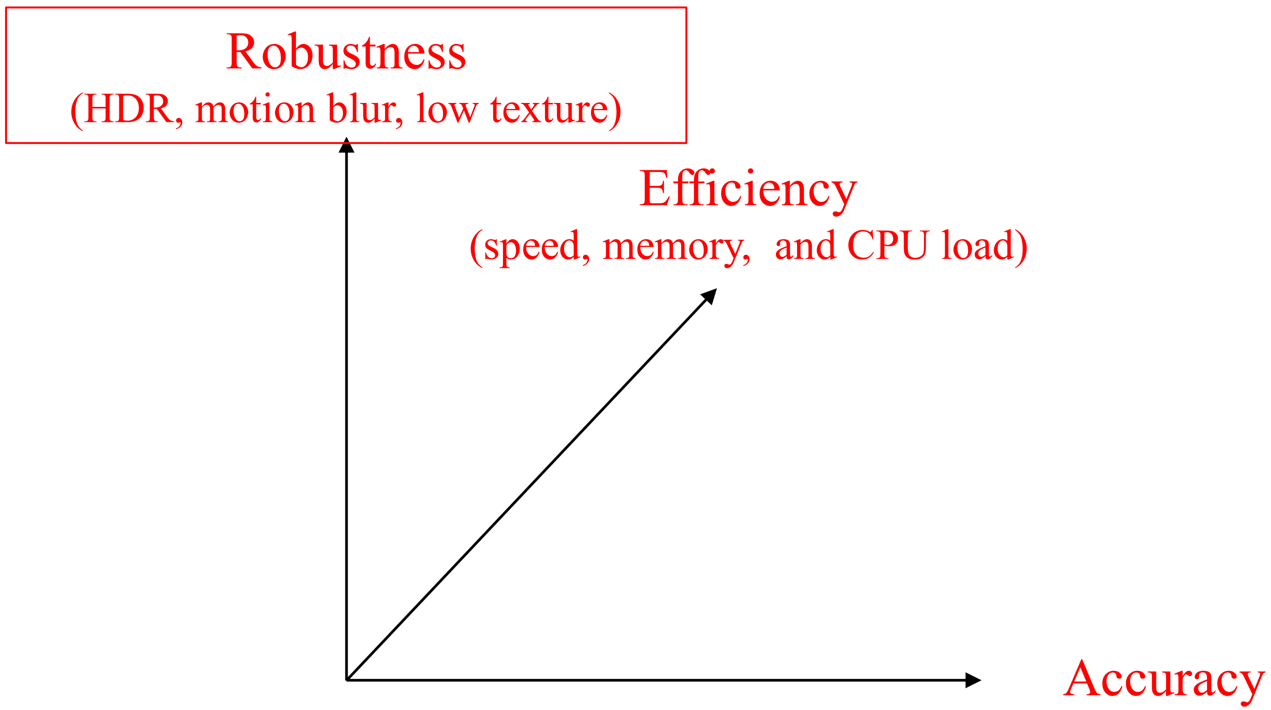
Case study: VIO for Flying Robots [ICRA'18]



No free lunch: more computation → better accuracy

Delmerico, Scaramuzza, A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots, ICRA'18. PDF. Video

What metrics should be used?



Robustness is the greatest challenge for SLAM today!

How to cope & quantify robustness to:

- Low texture
- High Dynamic Range (HDR) scenes
- Motion blur
- Dynamically changing environments
- Algorithmic randomness

How can we quantify the robustness of algorithms to such situations?

High Dynamic Range

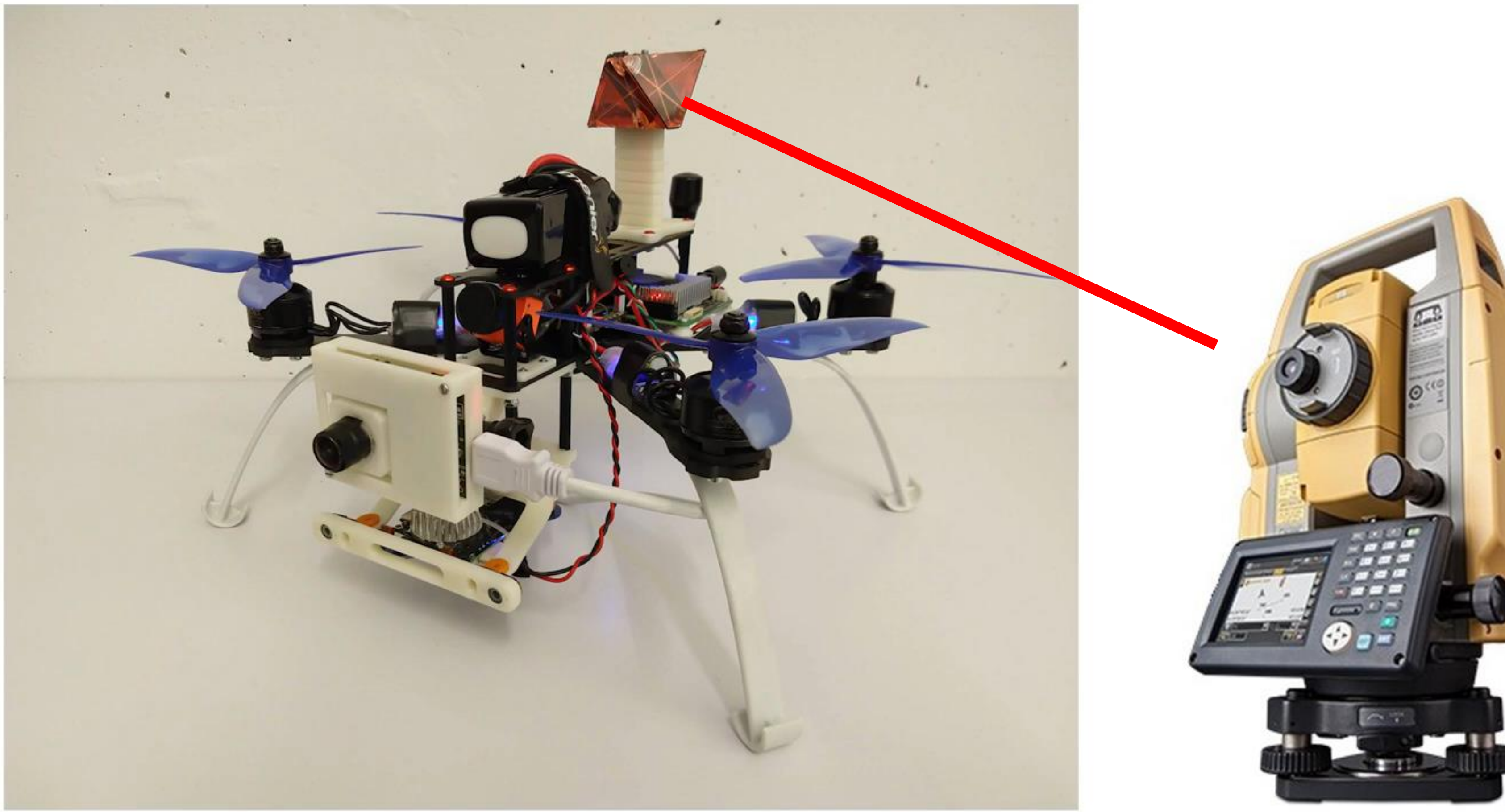
Motion blur

Latency

Cadena, Carlone, Carrillo, Latif, Scaramuzza, Neira, Reid, Leonard, *Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age,* IEEE Transactions on Robotics, 2016. PDF

UZH-FPV Drone Racing Dataset

Contains data recorded by a drone flying up to over 20m/s indoors and outdoors frown by a professional pilot. Contains frames, events, IMU, and Ground Truth from a Robotic Total Station: http://rpg.ifi.uzh.ch/uzh-fpv.html



Delmerico et al. "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset" ICRA'19
PDF. Video. Datasets.

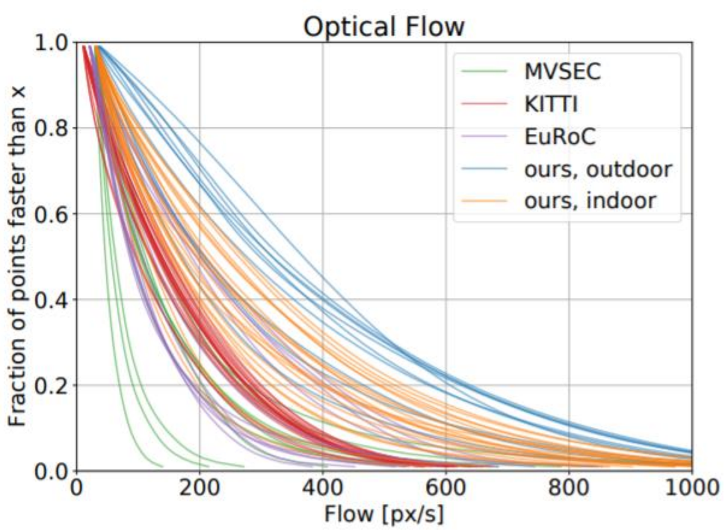
UZH-FPV Drone Racing Dataset

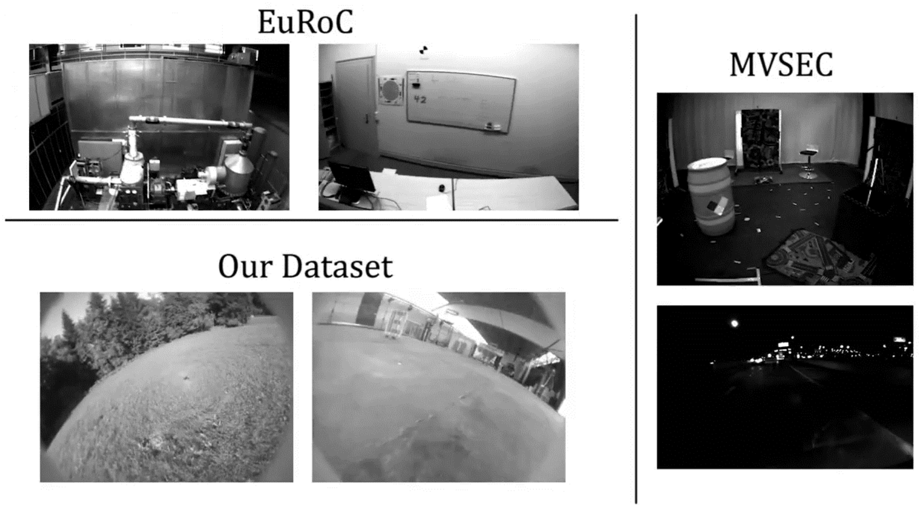
- Recorded with a drone flown by a **professional pilot up to over 20m/s**
- Contains **images, events, IMU**, and **ground truth from a robotic total station**: http://rpg.ifi.uzh.ch/uzh-fpv.html

Delmerico et al. "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset" ICRA'19
PDF. Video. Datasets.

Robustness to high speed motion

Can be quantified in terms of "optical flow" (see Lecture 10 for def. of optical flow)

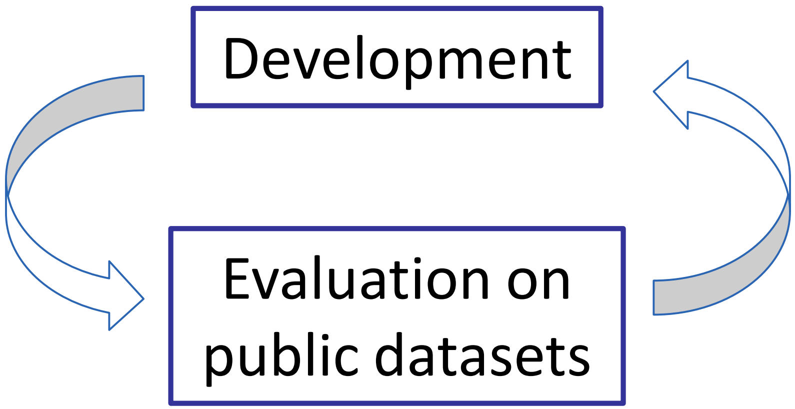




Delmerico et al. "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset" ICRA'19
PDF. Video. Datasets.

Dataset Bias

Typical workflow of developing VO/VIO/SLAM algorithms:



As a community, we are overfitting the public dataset.

Potential problems:

- **Generalizability:** Performance on one does not guarantee to generalize to others
  - E.g., KITTI → low frame rate, not friendly for direct methods

- **Old datasets (e.g., KITTI) are already saturated:**
  - It becomes more and **more difficult to tell whether we are making real progress** or just overfitting the datasets.
  - E.g., **does 1 or 2 cm improvement in RMSE** over a 100 meter trajectory **really mean something**?

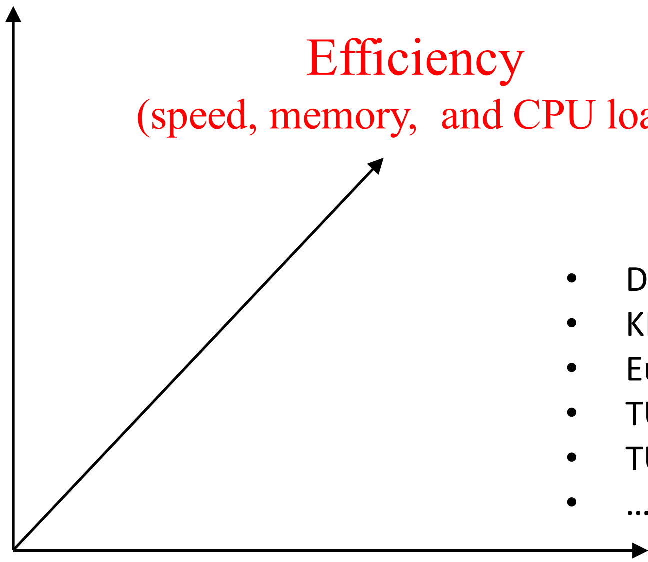
Dataset Bias

We need more datasets to evaluate the performance of SLAM algorithms along different axes

Robustness
(HDR, motion blur, low texture)

- BlackBird [Antonini’18]
- UZH-FPV dataset [Delmerico’18]
- Event Camera [Mueggler’17]
- MVSEC [Zhu’18]
- …

Efficiency
(speed, memory, and CPU load)



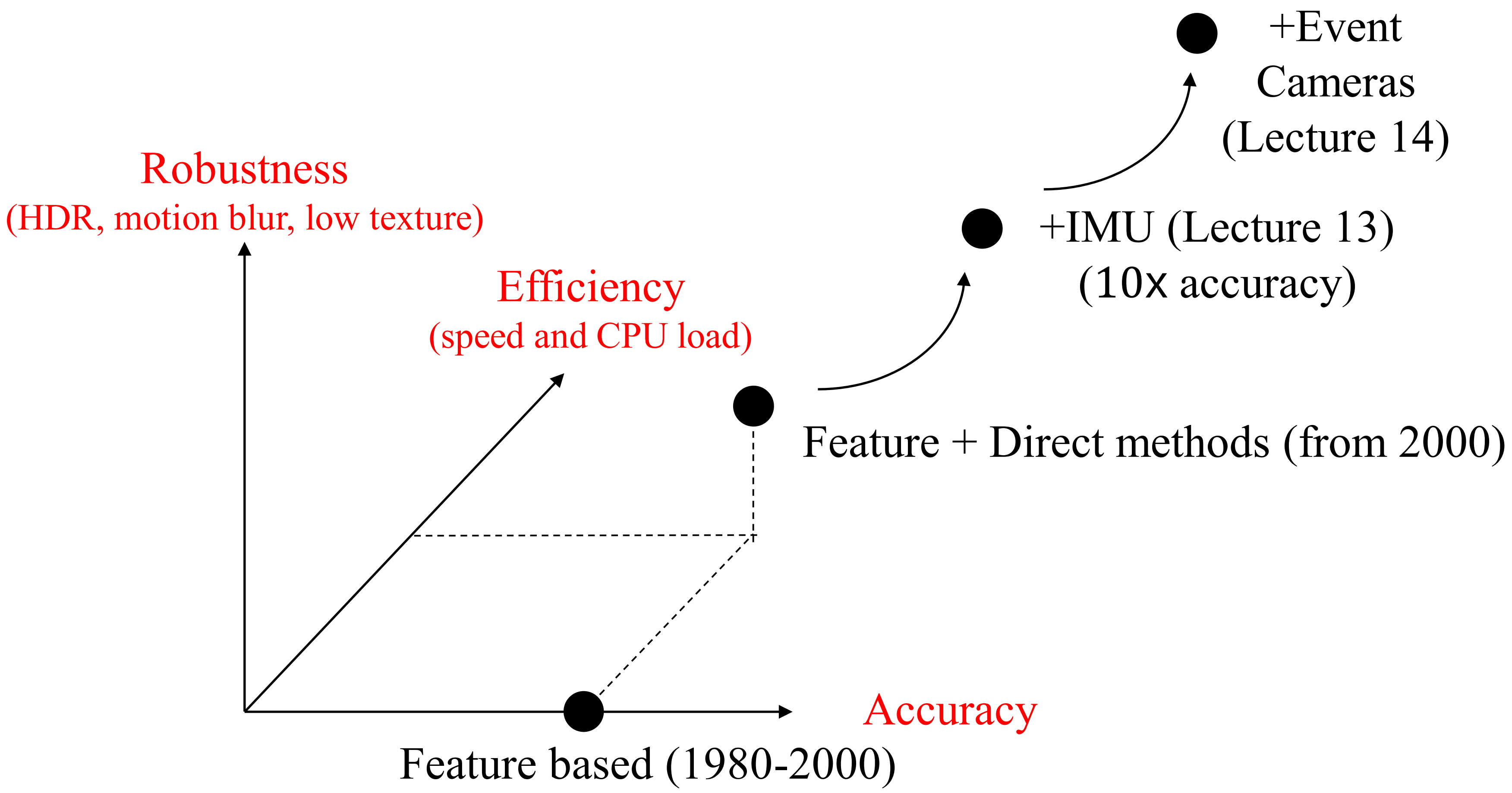
- Devon Island [Furgale’13]
- KITTI [CVPR’12]
- EuRoC [Burri’16]
- TUM-RGBD [Sturm’12]
- TUM VI Benchmark [Schubert’18]
- …

Accuracy

Realistic simulators:

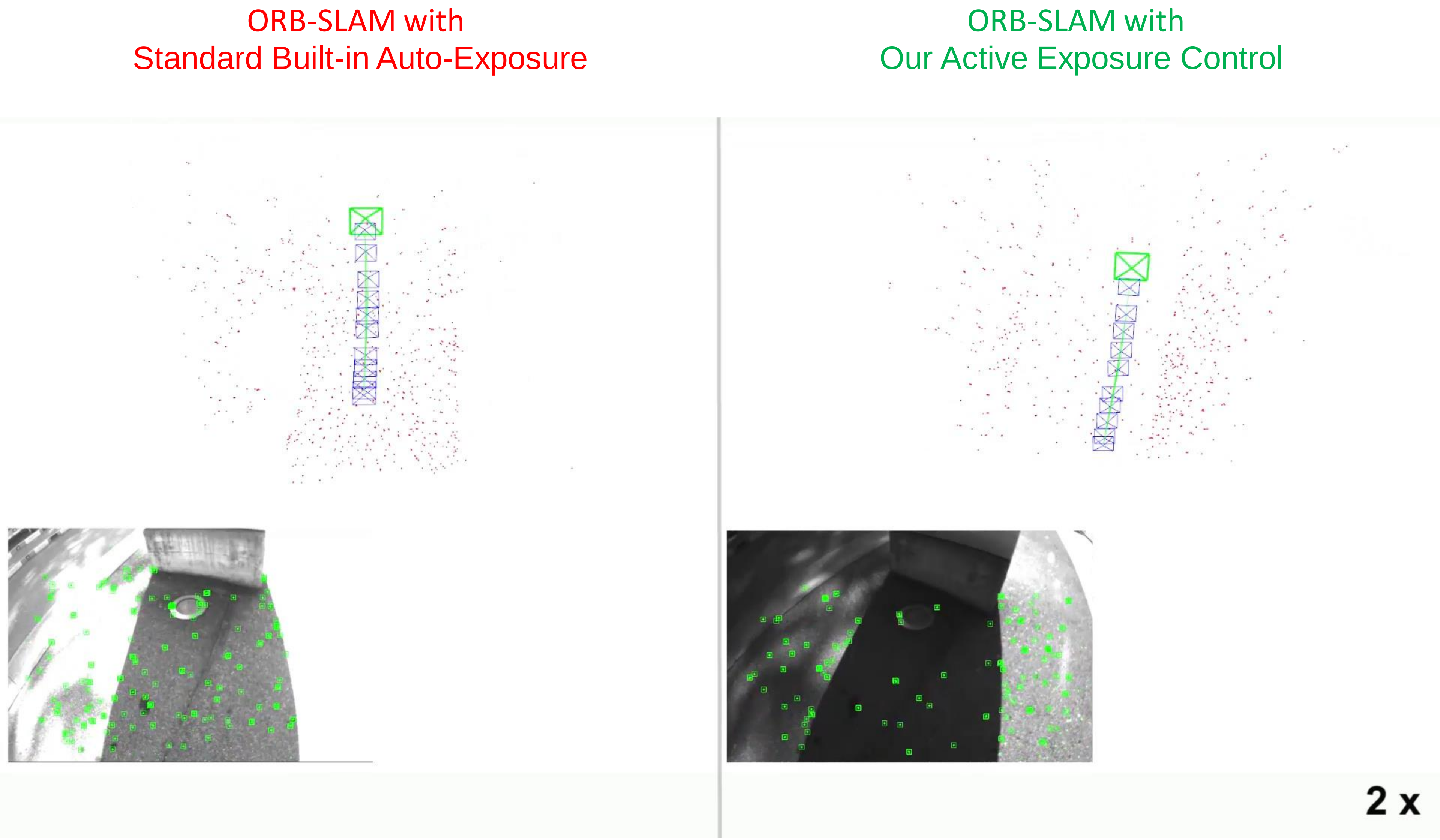
- AirSim
- FlightGoggles [Guerra’19]
- ESIM [Rebecq’18]
- ……

Overview of the last 30 years of Visual Inertial Odometry & SLAM



Open
Research Opportunities

Actively Control Camera Exposure Time to achieve Robustness to High Dynamic Range (HDR) scenes



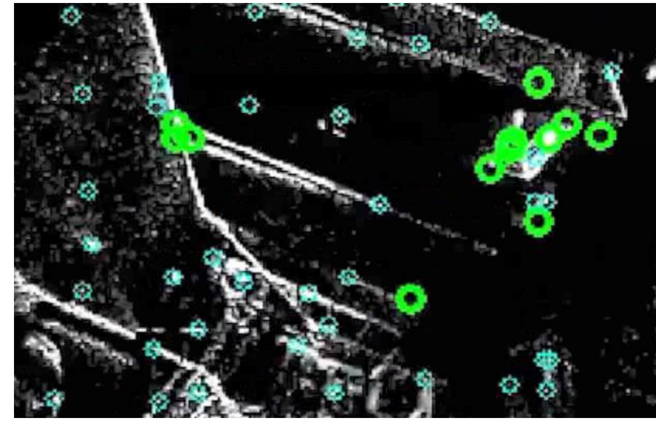
Zhang, et al., Active Exposure Control for Robust Visual Odometry in HDR Environments, ICRA'17. [PDF](). [Video]()

“UltimateSLAM”: Frames + Events + IMU (Lecture 14)

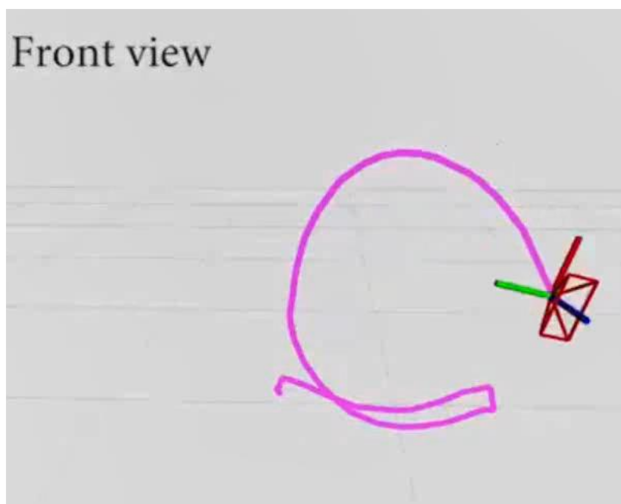
85% accuracy gain over standard visual-inertial SLAM in HDR and high speed scenes

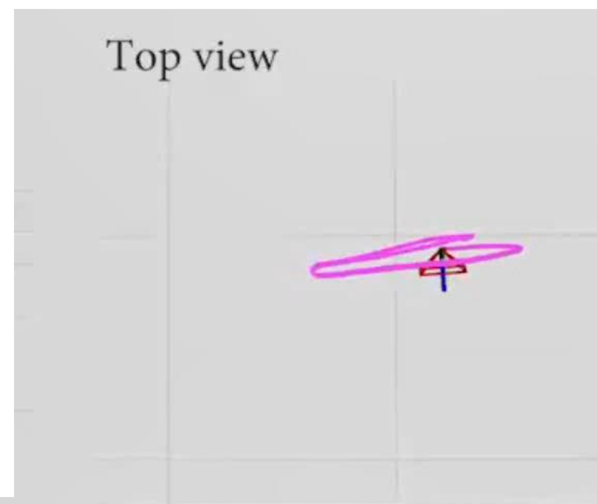
Standard camera

Event camera



Estimated trajectory





Rosinol et al., *Ultimate SLAM?* **IEEE RAL’18 best Paper Award Honorable Mention** PDF. Video. IEEE Spectrum.

Understanding Check

Are you able to answer the following questions:

- How do we benchmark VO/SLAM algorithms?
- Along which axes can we evaluate them?
- Benchmarking accuracy: Can we use the end pose error? What are ATE and RTE?
- How can we quantify Efficiency? And Robustness?
- What are open research opportunities?