State of the Art in Visual Localisation for UAVs in GNSS-denied environments

Michael Schleiss

SET-275

Split, 29 September 2021

Fraunhofer FKIE

Motivation for Vision-based Navigation

- Independent from external localization sources
  - Safe against jamming and spoofing
  - Works indoors and outdoors
  - No additional radio based infrastructure needed (e.g. UWB)
- Accuracy
  - Can reduce localization error if precise navigation required (e.g. lane keeping in autonomous driving, landing on airport)
- Low energy and weight
  - Can be implemented even in small UAVs
- Passive
  - Reducing exposure compared to radar/lidar based systems

Structure-from-motion based navigation.

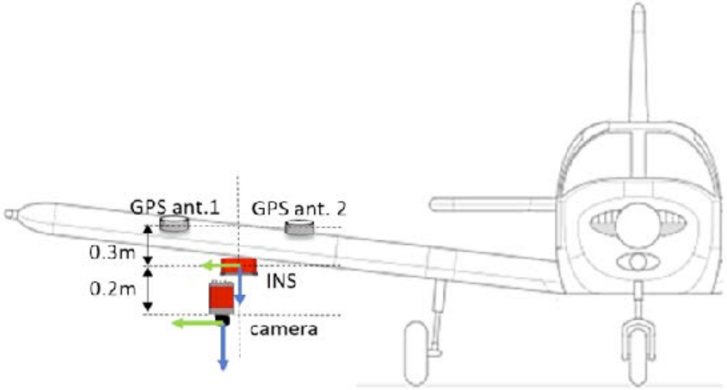
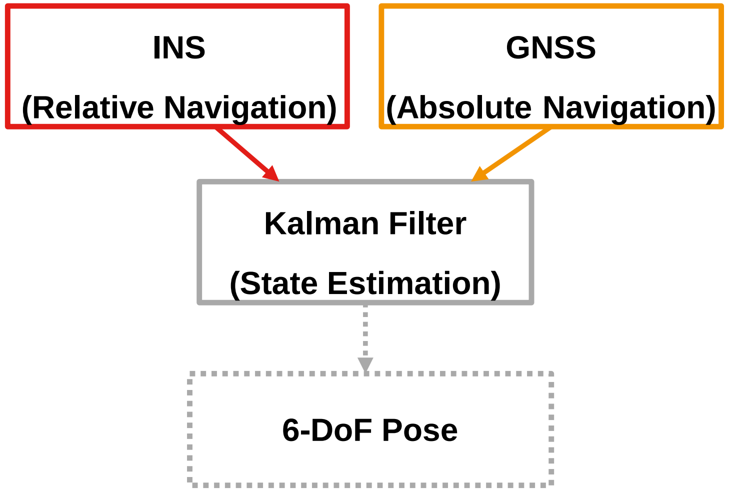
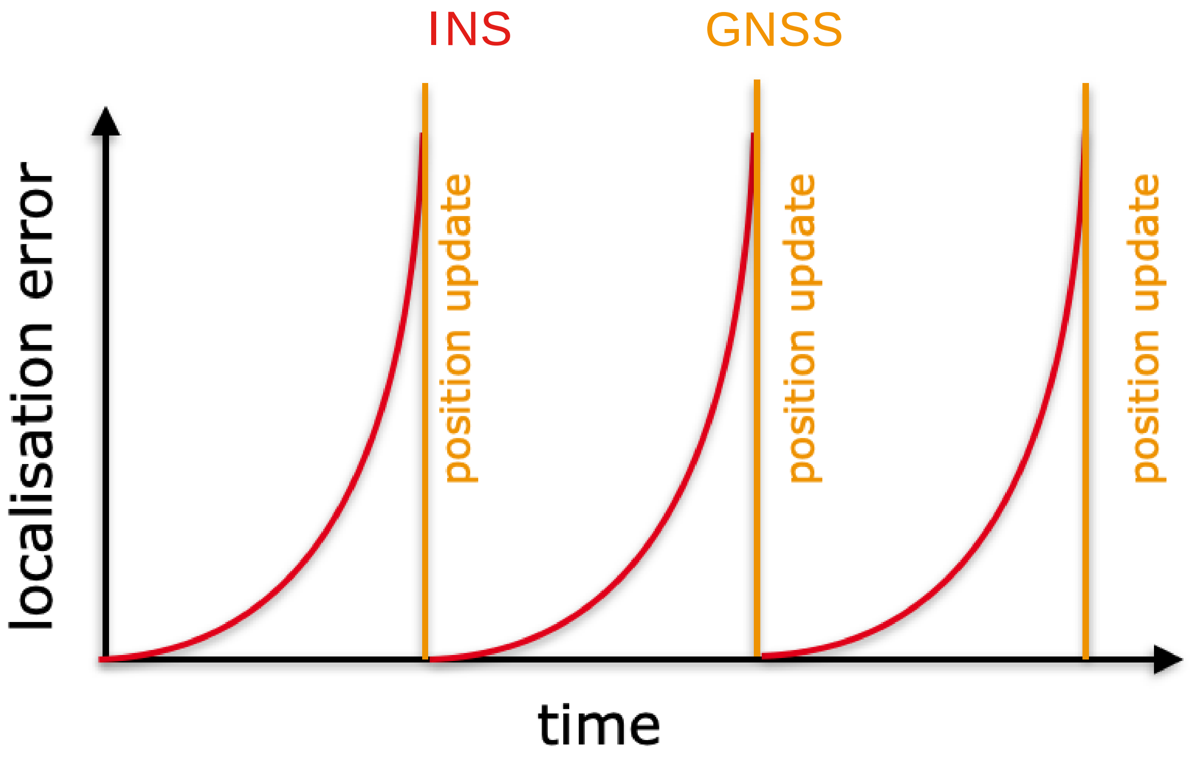


Diagram of our aerial survey payload.

Fraunhofer FKIE

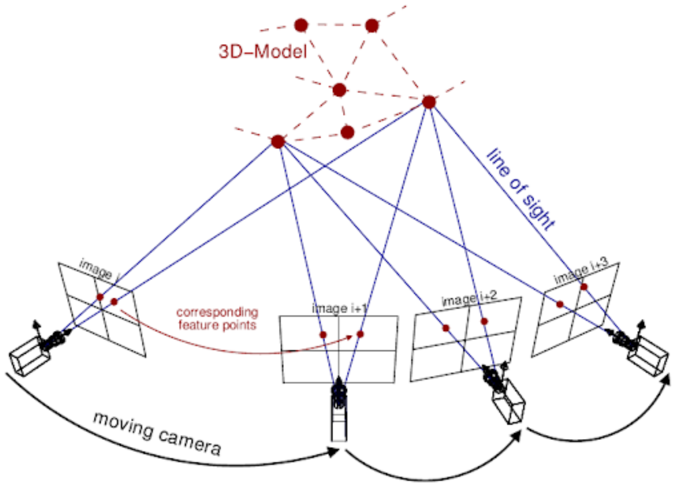
„Traditional“ Satellite-based Navigation





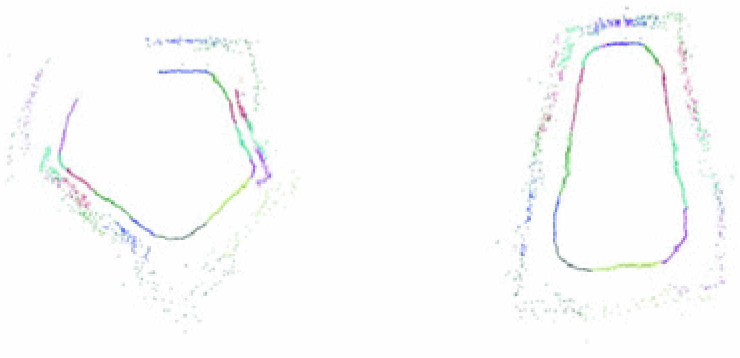
Vision-based Navigation

Visual-(Inertial)-Odometry



From: Theia Vision Library. http://theia-sfm.org/sfm.html

Visual SLAM

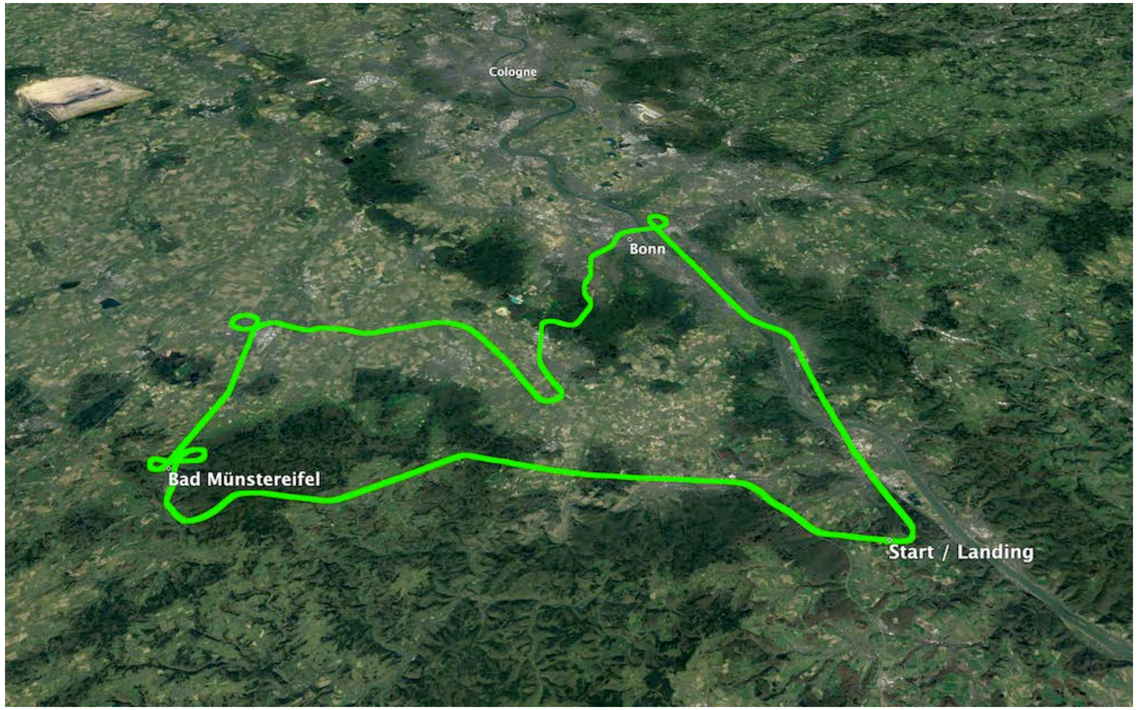


From: F. Bokovoy et al. (2017): “Original Loop-Closure Detection Algorithm for Monocular vSLAM“, Springer

Scene Matching / Global Place Recognition

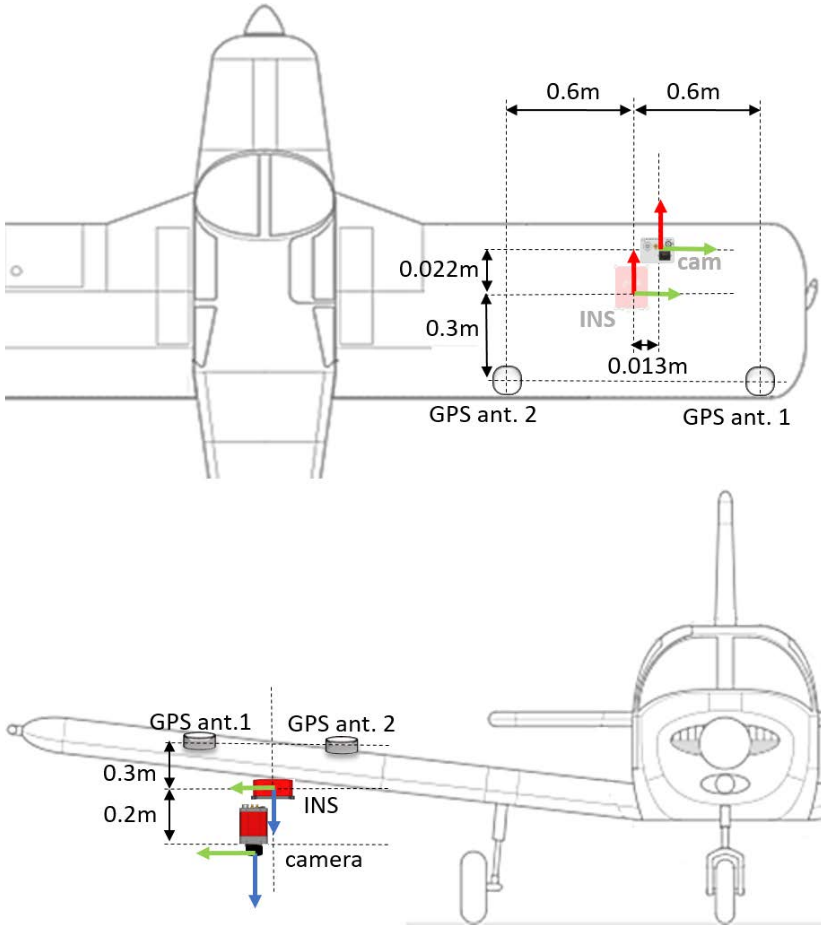
From: F. Warburg et al. (2020): “Mapillary Street-Level Sequences: A Dataset for Lifelong Place Recognition“, CVPR 2020

AerVisLoc: Dataset for Aerial Visual Localization



Sensor setup

- <u>Mono-RGB Camera</u> 1600 × 1200, 25Hz, global shutter, 57° HFoV
- <u>INS/GNSS navigation system</u>, 6-axis IMU, 200 Hz, dual antenna, resolution: <1 m (SBAS) / 0.05°



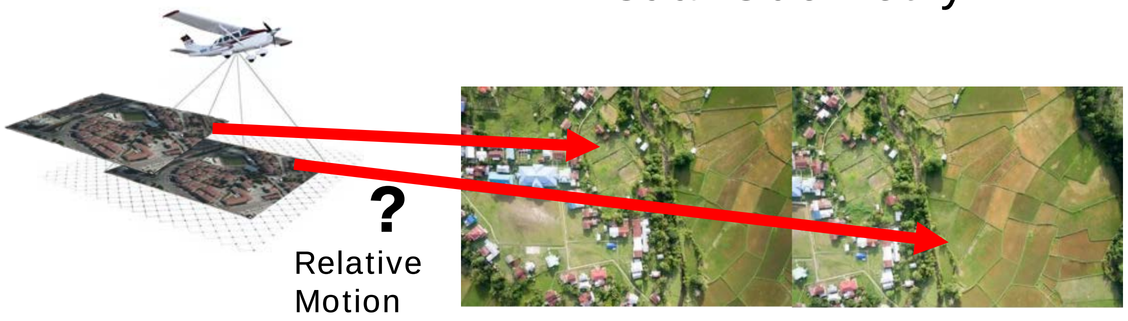
Experiment 1: Combined VO and VPR

- Visual Navigation Framework
  - Visual Odometry: ORB-SLAM3

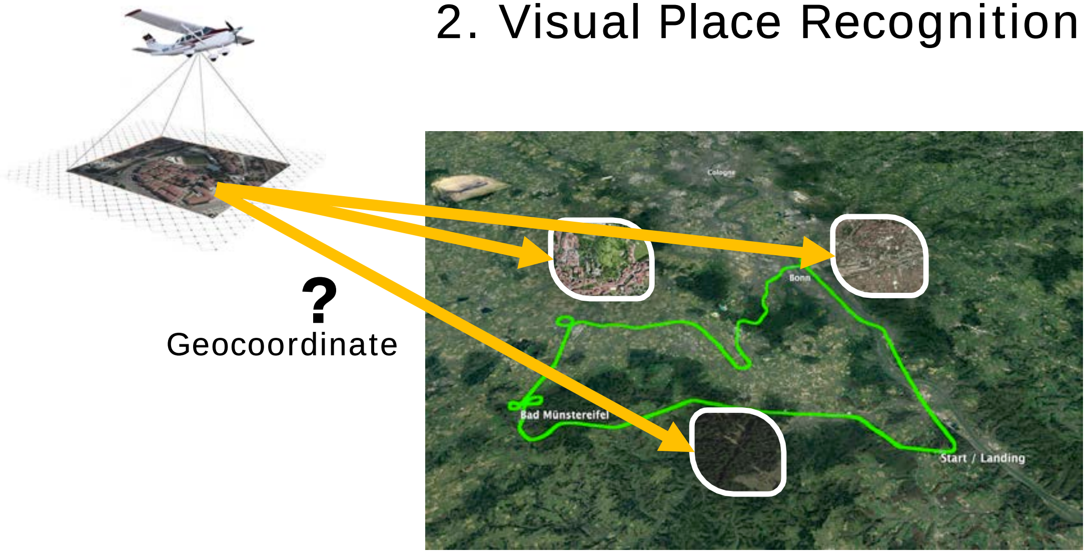
  -> Relative Position
  - Visual Place Recognition:
    - Algorithm: NetVLAD
    - Geotagged Image Database

  -> Absolute Position

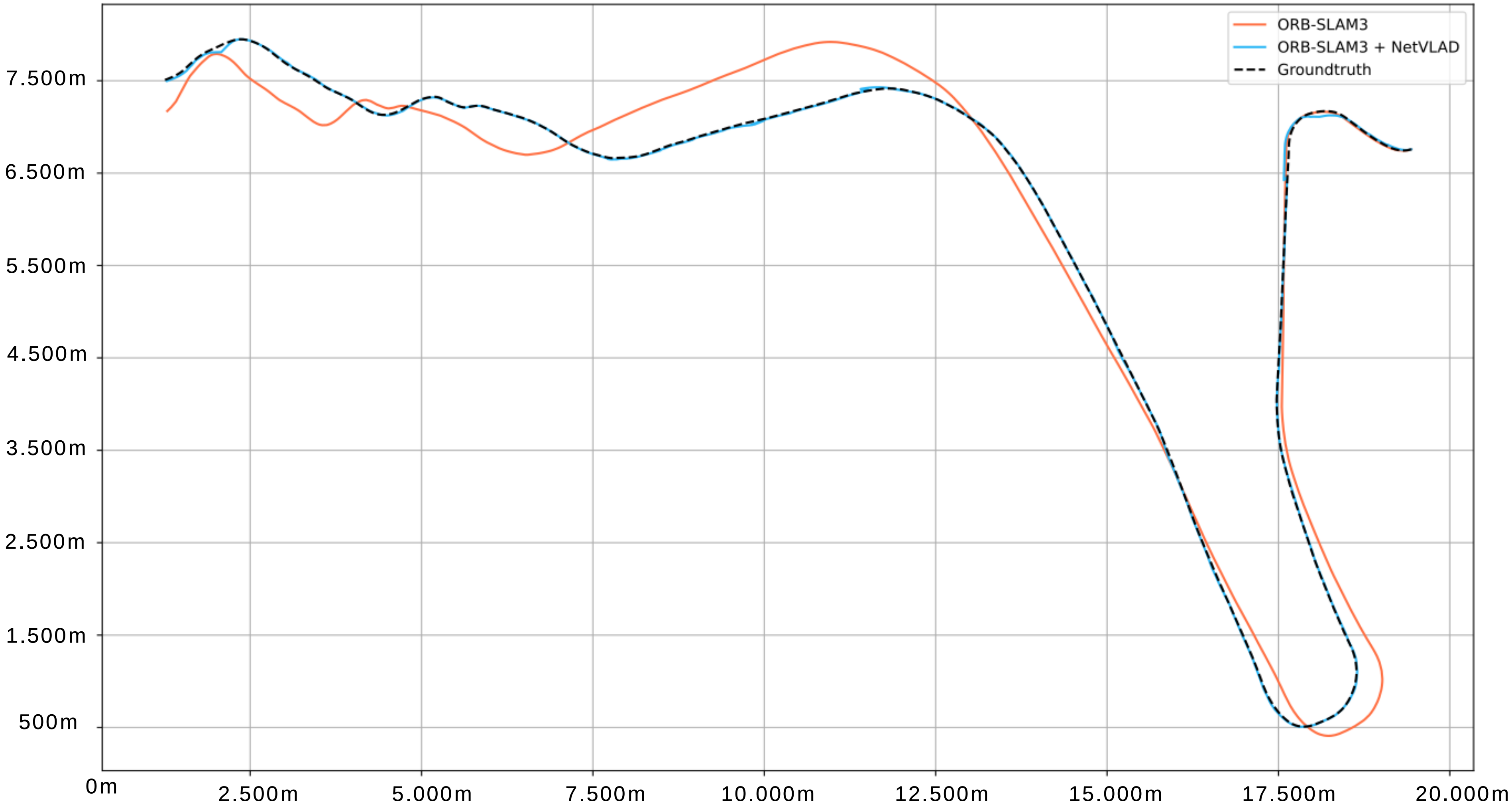
1. Visual Odometry



2. Visual Place Recognition



Experiment 1: Evaluation



Experiment 2: Rotation-Invariant Global Place Recognition

= ?

Our Dataset	Q	Q_m	Q_f
NetVLAD	47.5	41.2	30.0
NetVLAD*	51.1	47.2	35.3

Table 2. Global place recognition results in terms of Recall@1 on our dataset. $*$ denotes that the network was trained with flip and mirror augmentations.

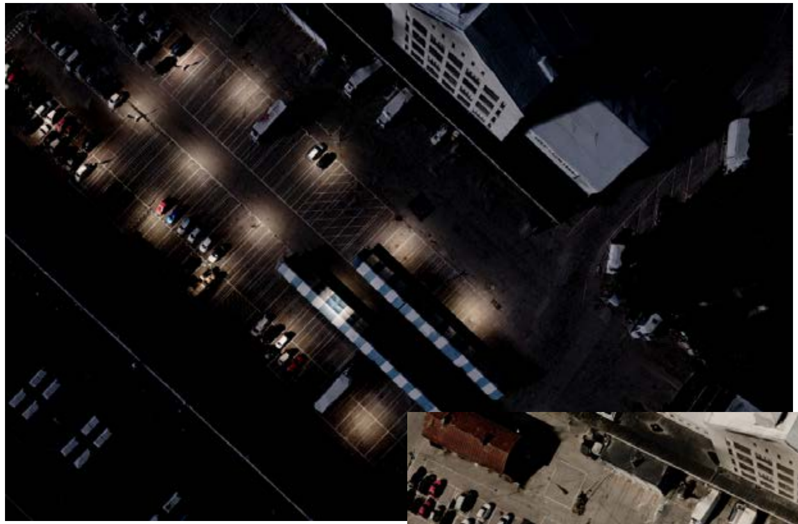
Pits30k-val	Q
NetVLAD	85.2
NetVLAD*	84.6

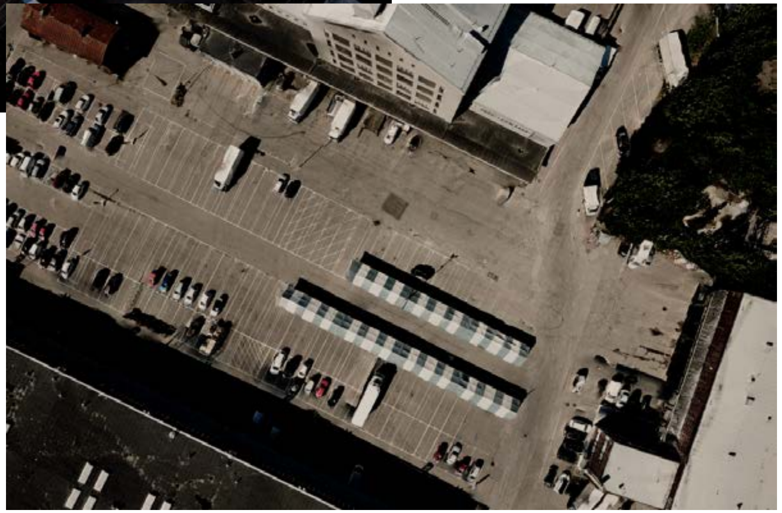
Table 3. Global place recognition results in terms of Recall@1 on Pits30k-val.

Fraunhofer FKIE

Future Work

- Simulation of photorealistic environments
  - Day/Night Simulation
  - Weather effects
  - Seasonal effects
- Thermal-Image based methods
  - For day/night navigation
- Rotation-invariant neural network
  - For viewpoint independent visual place recognition





Conclusion

- Vision-based navigation demonstrated on real, challenging flight
- Motivated why further research for aerial visual localization is needed
- Release of public benchmark dataset for measuring state-of-the art
- Open to reasearchers and partners under aervisloc.github.io