University of Zurich[UZH] | ETH zürich

Institute of Informatics – Institute of Neuroinformatics

ROBOTICS & PERCEPTION GROUP

Decentralized Visual SLAM with Minimal Data Exchange

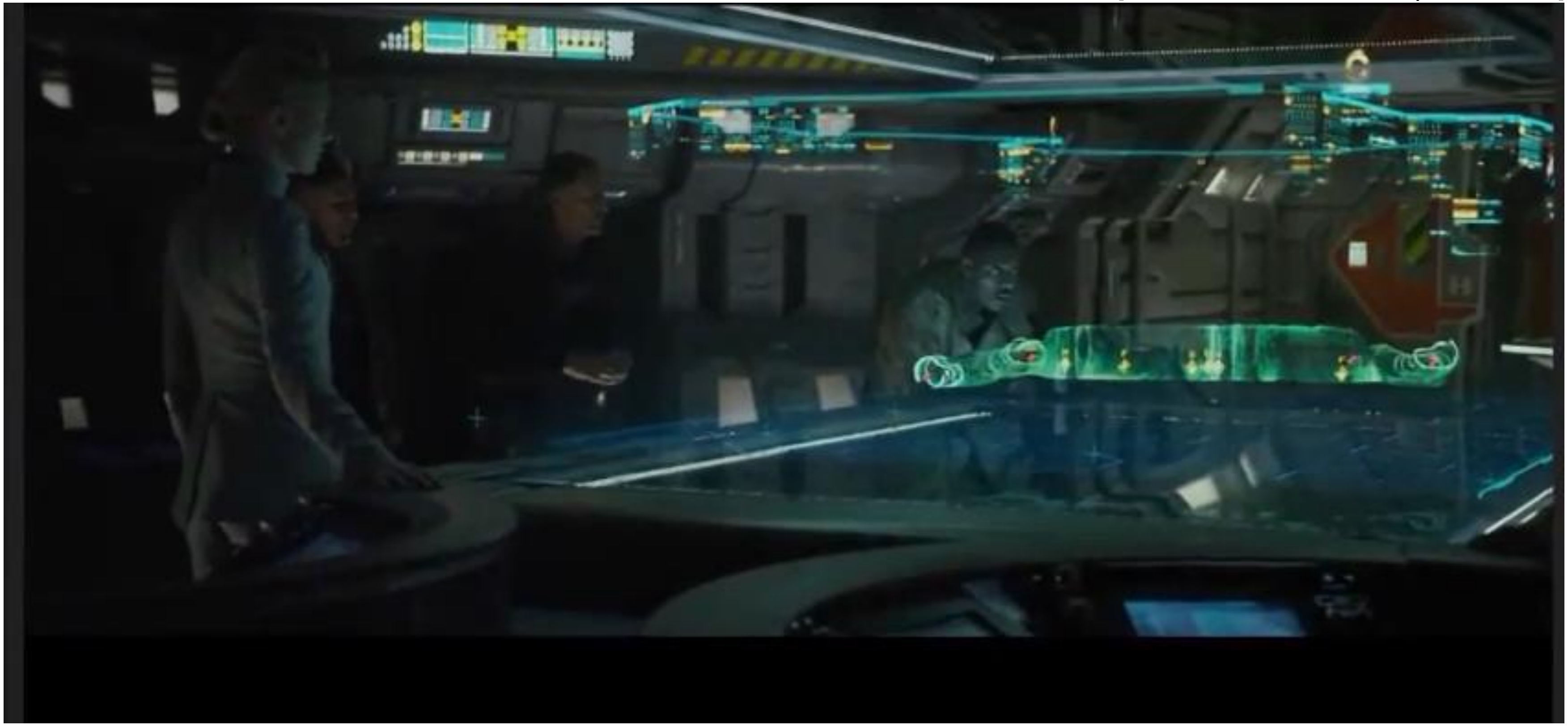
Titus Cieslewski, Davide Scaramuzza
University of Zurich and ETH Zurich

T. Cieslewski, D. Scaramuzza: *Efficient Decentralized Visual Place Recognition…,* **MRS 2017.** PDF
T. Cieslewski, S. Choudhary, D. Scaramuzza: *Data-Efficient Decentralized Visual SLAM* **ICRA 2018.** PDF Video Code
T. Cieslewski, M. Bloesch, D. Scaramuzza: *Matching Features without Descriptors…,* **BMVC 2019** PDF Code
T. Cieslewski, A. Ziegler, D. Scaramuzza: *Exploration Without Global Consistency…,* **ISRR 2019** PDF Video

Why care about Multi-Robot SLAM?

- Multi-robot teaming
- Esp. unknown environments
- Tasks involving several skills
- Collaborative transport
- Collaborative exploration
- Min. data exchange: Scalability

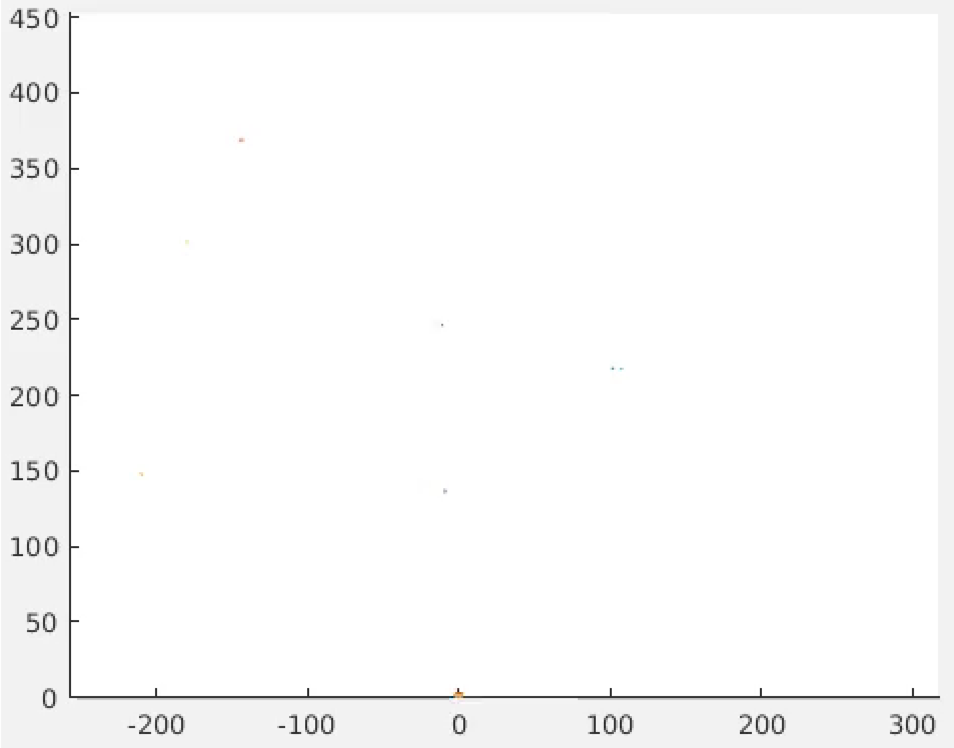
[©2012 Scott Free, Brandywine, Dune]



T. Cieslewski, S. Choudhary, D. Scaramuzza: *Data-Efficient Decentralized Visual SLAM* **ICRA 2018.** PDF Video Code

Components of multi-robot SLAM

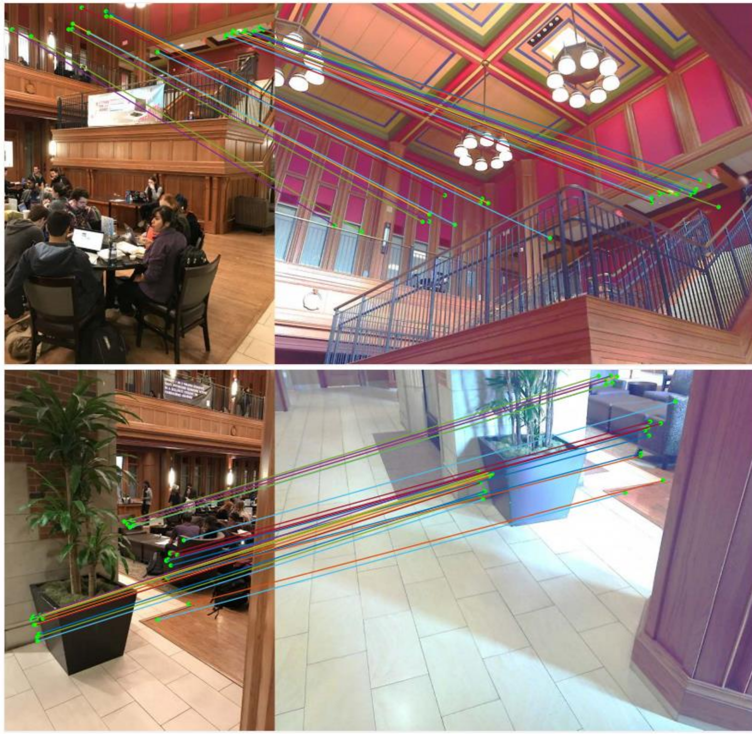
- ~~Single-robot SLAM running on each robot (no data exchange)~~
- Place recognition between the robots (yes data exchange)
- Optimization of the joint map (yes data exchange)



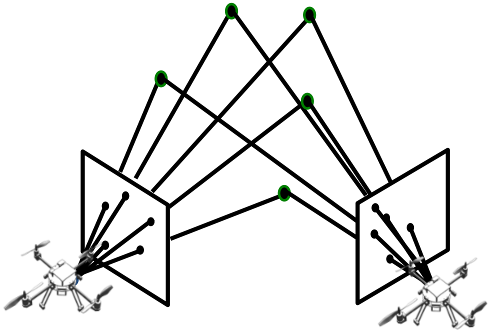
T. Cieslewski, S. Choudhary, D. Scaramuzza: *Data-Efficient Decentralized Visual SLAM* **ICRA 2018.** PDF Video Code

Pose Estimation

- Extract **2D features**, match them, do P3P and RANSAC
- Traditional features
  - SIFT, SURF, ORB, BRISK, FREAK, …
- Learned features
  - LF-NET, SuperPoint, D2-Net, …
- Typical focus: Matching score
- **Our focus: Minimal representation**
- Representation:
  - **Point count x (descriptor size + point location)**
  - E.g. ORB: 500 x (32 + 3) Bytes = 17.5 kB



[Dusmanu 2019 D2-NET]

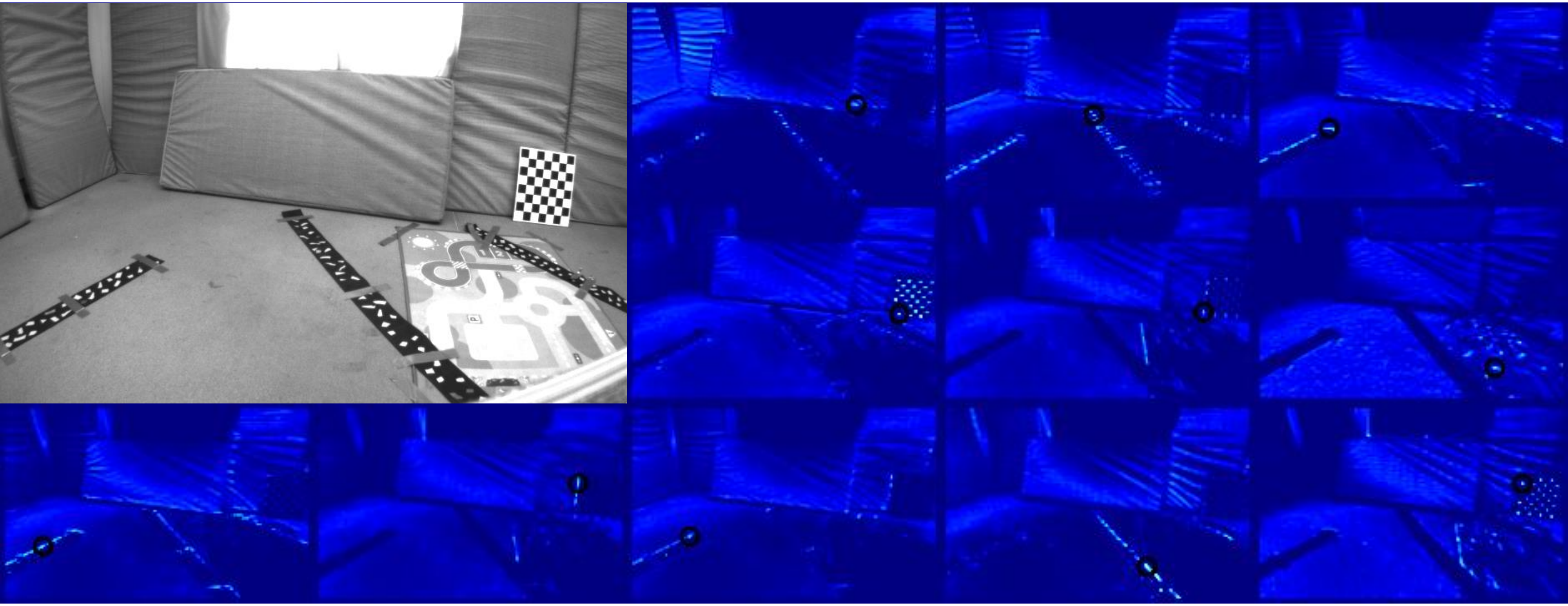


T. Cieslewski, K. Derpanis, D. Scaramuzza: *SIPs: Succinct Interest Points…*, **3DV 2019** PDF Video Code

T. Cieslewski, M. Bloesch, D. Scaramuzza: *Matching Features without Descriptors…*, **BMVC 2019** PDF Code

Matching interest points without descriptors

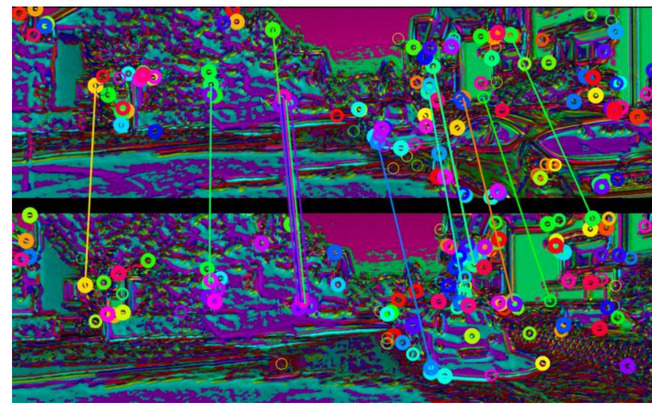
- Idea: Interest point detection network with **n (e.g. 128) output channels** instead of 1 output channel
- Interest points are global maxima of each channel
- Points **extracted by the same channel are implicitly matched**



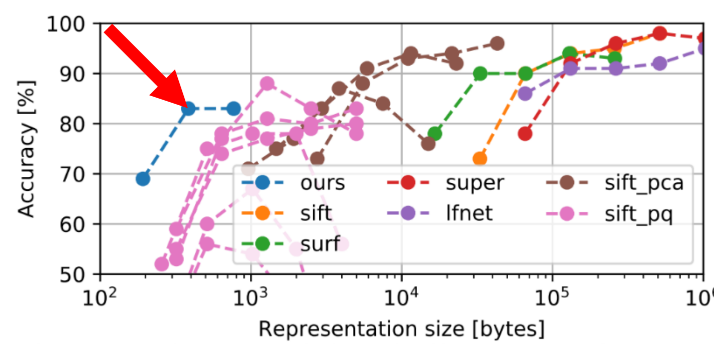
T. Cieslewski, M. Bloesch, D. Scaramuzza: *Matching Features without Descriptors…,* **BMVC 2019** PDF Code

Matching interest points without descriptors

- Self-supervised training:
  - On image pairs
  - Ensure **each channel** picks the **same point** in both images
  - Ensure **different channels** do **not** pick the **same point**



- 128 x (0 + 3) Bytes = **384 Bytes** (VS 17.5 kB using ORB descriptors)



T. Cieslewski, M. Bloesch, D. Scaramuzza: *Matching Features without Descriptors...*, **BMVC 2019** PDF Code

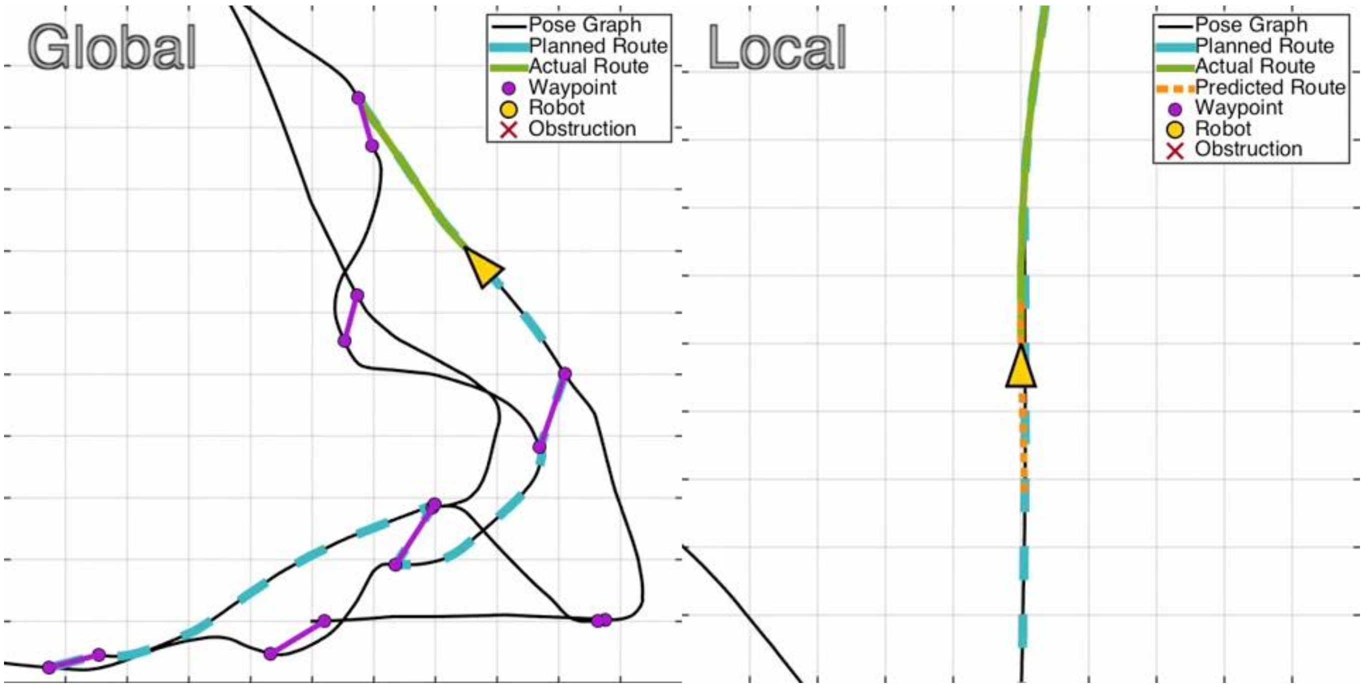
Do we really need to optimize?

[Tabula Rogeriana - Wikimedia]

T. Cieslewski, A. Ziegler, D. Scaramuzza: *Exploration Without Global Consistency…,* **ISRR 2019** PDF Video

For navigation? No!

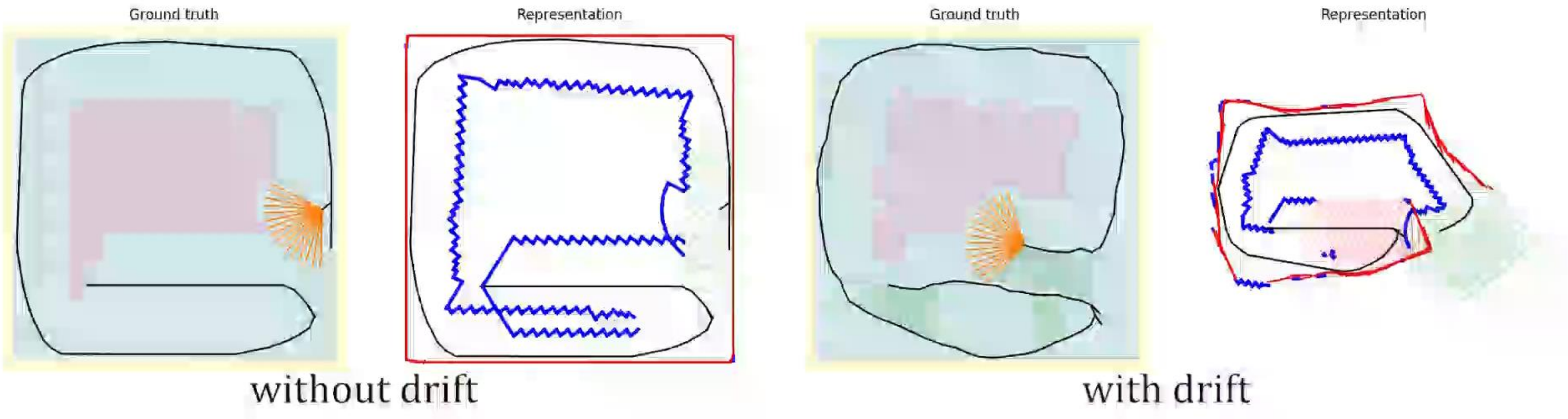
- (Visual) teach and repeat!
- van Es 2015: “Being in two places at once”
- Only local consistency needed



[van Es 2015 Being]

For exploration? Also no!

- Exploration: Want to ensure unknown environment is fully covered
- Keep track of the boundary between known and unknown space
- Traditionally done with globally consistent maps

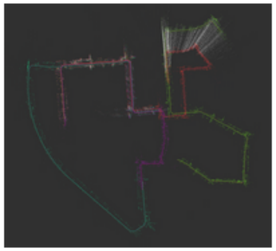


T. Cieslewski, A. Ziegler, D. Scaramuzza: *Exploration Without Global Consistency…,* **ISRR 2019** PDF Video

Interested?

- Further reading
  - http://rpg.ifi.uzh.ch/people_titus.html
- Student projects
  - http://rpg.ifi.uzh.ch/student_projects.php

Decentralized Visual Map Building - Available



Description: In State-of-the-Art decentralized mapping methods, optimization (correcting odometry drift) is typically done using pose graph optimization due to the fact that a pose graph is a very compact representation. Unfortunately, this compression in data results in limitations in precision and robustness. Bundle Adjustment is a map optimization method for visual maps which is much more precise and robust, but also much more data intensive.

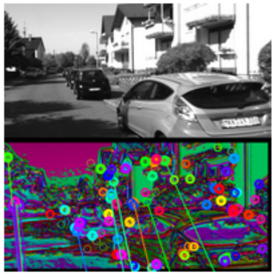
Goal: In this work, you will figure out a way to achieve the superior precision of Bundle Adjustment while minimizing the amount of data that needs to be exchanged between robots in a decentralized setting.

Contact Details: Titus Cieslewski (titus at ifi.uzh.ch), APPLY VIA EMAIL, ATTACH CV AND TRANSCRIPT! Required skills: Matlab or C++, with a preference for the latter. Desirable: Background in optimization (Nonlinear least squares, Gauss-Newton or similar)

Thesis Type: Semester Project / Master Thesis

See project on SiROP

Learning minimal representations of places - Available



Description: Place recognition and 6DoF localization has a wide range of applications, whether in robot autonomy, VR/AR or navigation interfaces. Given sensor readings (we focus on images), the goal is to establish position and orientation of a robot/device with respect to a previously recorded map. Recently, this is generally solved with a mixture of machine learning and geometry (NetVLAD, SuperPoint, LF-NET, PoseNet). Our focus in particular will be to solve this problem with a minimal representation.

Goal: Given query agent A and map agent B, have B establish a pose of A within its map, with minimal data transmission from A to B. We have a couple of ideas on how to solve this (see our most recent publication on this: https://arxiv.org/abs/1811.10681), but you are encouraged to bring your own ideas to the table.

Contact Details: Titus Cieslewski (titus at ifi.uzh.ch), APPLY VIA EMAIL, ATTACH CV AND TRANSCRIPT (also Bachelor)! Preferred skills: Linux, Python, "Vision Algorithms for Mobile Robots" class or equivalent, TensorFlow/PyTorch or equivalent.

Thesis Type: Master Thesis

See project on SiROP