Hierarchical Localization with hloc and SuperGlue

ECCV 2020 Workshop on
Map-based Localization for Autonomous Driving

Paul-Edouard Sarlin
Computer Vision and Geometry Lab
ETH Zurich

psarlin.com

ECCV'20 ONLINE
23-28 AUGUST 2020

ETH zürich

Computer Vision and Geometry Lab

hloc - a toolbox for visual localization

with

Mihai Dusmanu

Viktor Larsson

Marc Pollefeys

ETH zürich

SuperGlue - a graph neural network for feature matching

with

Daniel DeTone

Tomasz Malisiewicz

Andrew Rabinovich

magic leap

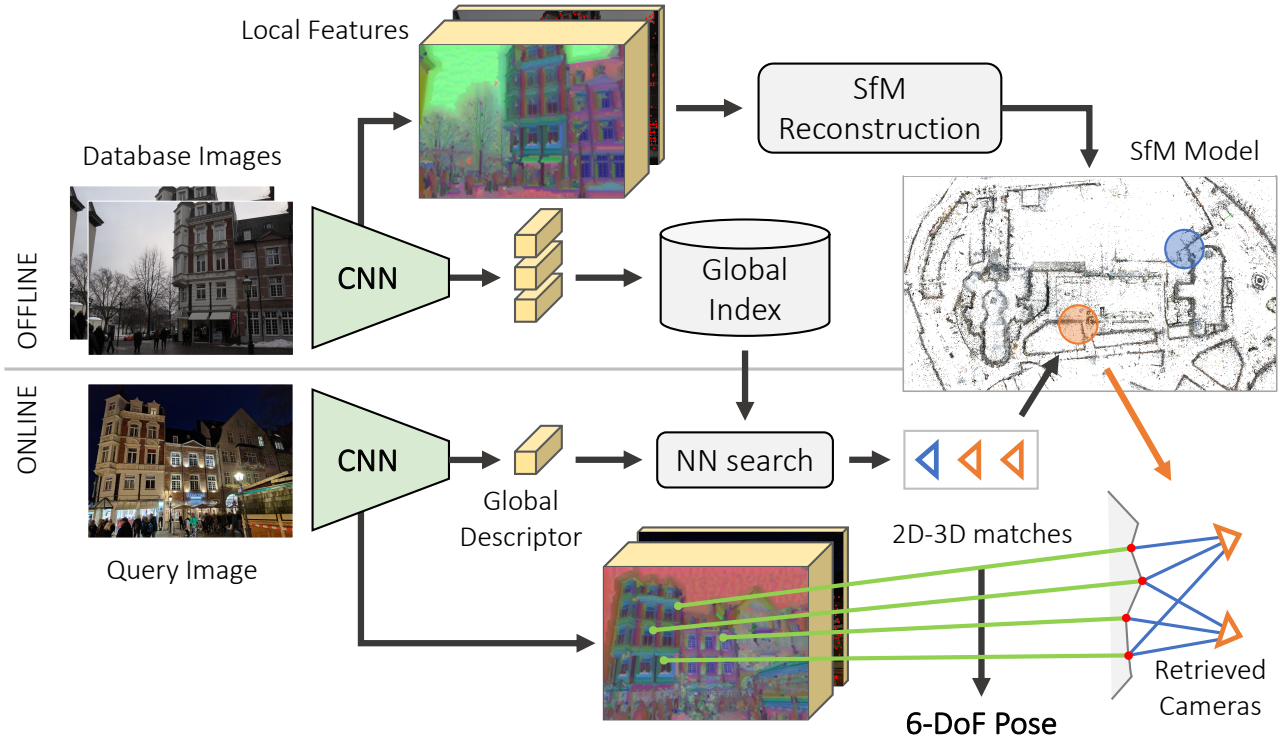
First place in 6 localization challenges!

At CVPR 2020: 2 challenges, local features & handheld devices

At ECCV 2020, workshops:

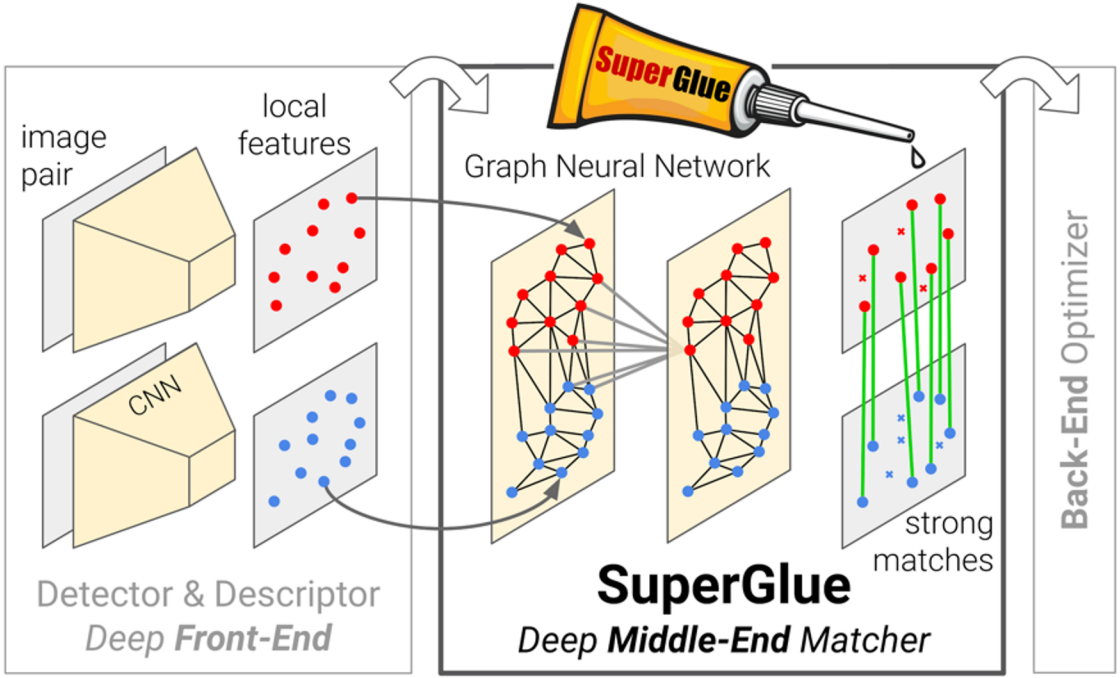
- Map-based Localization for Autonomous Driving — this talk
- Long-Term Visual Localization under Changing Conditions — Friday 28th

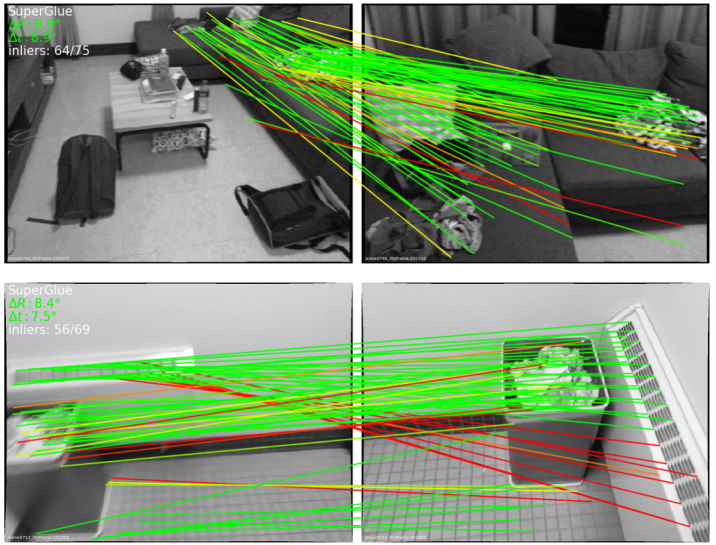
Hierarchical Localization



From Coarse to Fine: Robust Hierarchical Localization at Large Scale, CVPR 2019

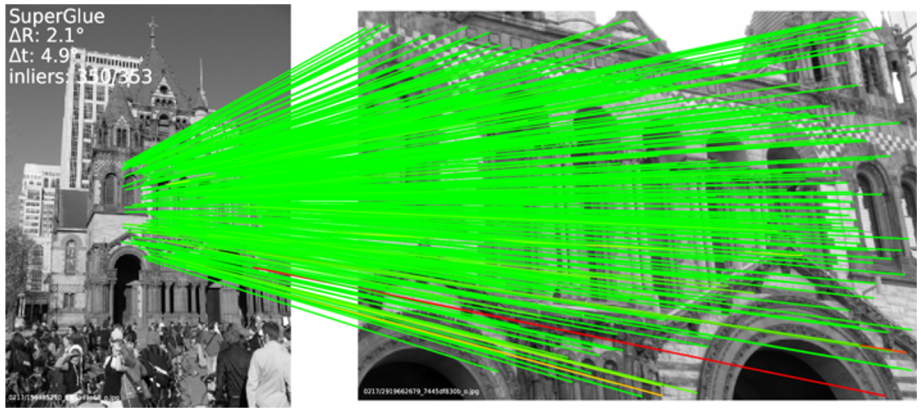
SuperGlue = Graph Neural Nets + Matching

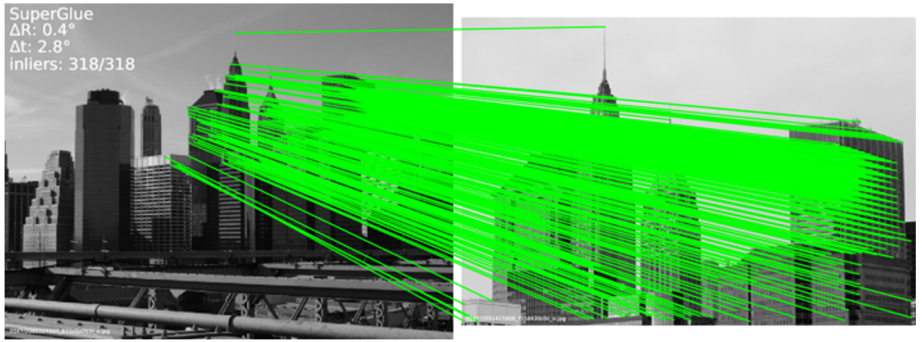




- Extreme **wide-baseline** image pairs in **real-time on GPU**
- State-of-the-art **indoor+outdoor** matching with **SIFT** & **SuperPoint**

SuperGlue = Graph Neural Nets + Matching





CVPR 2020

Paper + code + demo:

psarlin.com/superglue

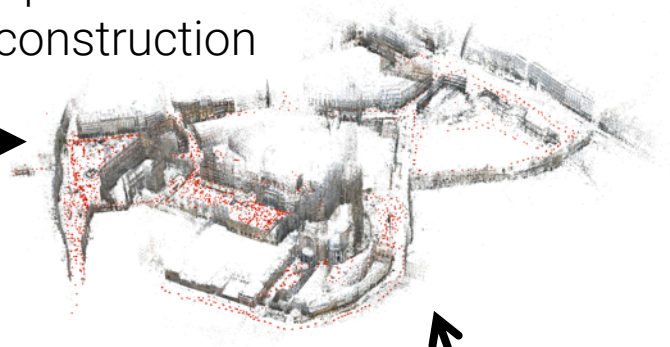
hloc – a toolbox for SfM & localization

images

3D reconstruction

hloc

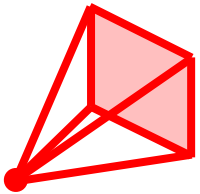
sparse 3D reconstruction



query

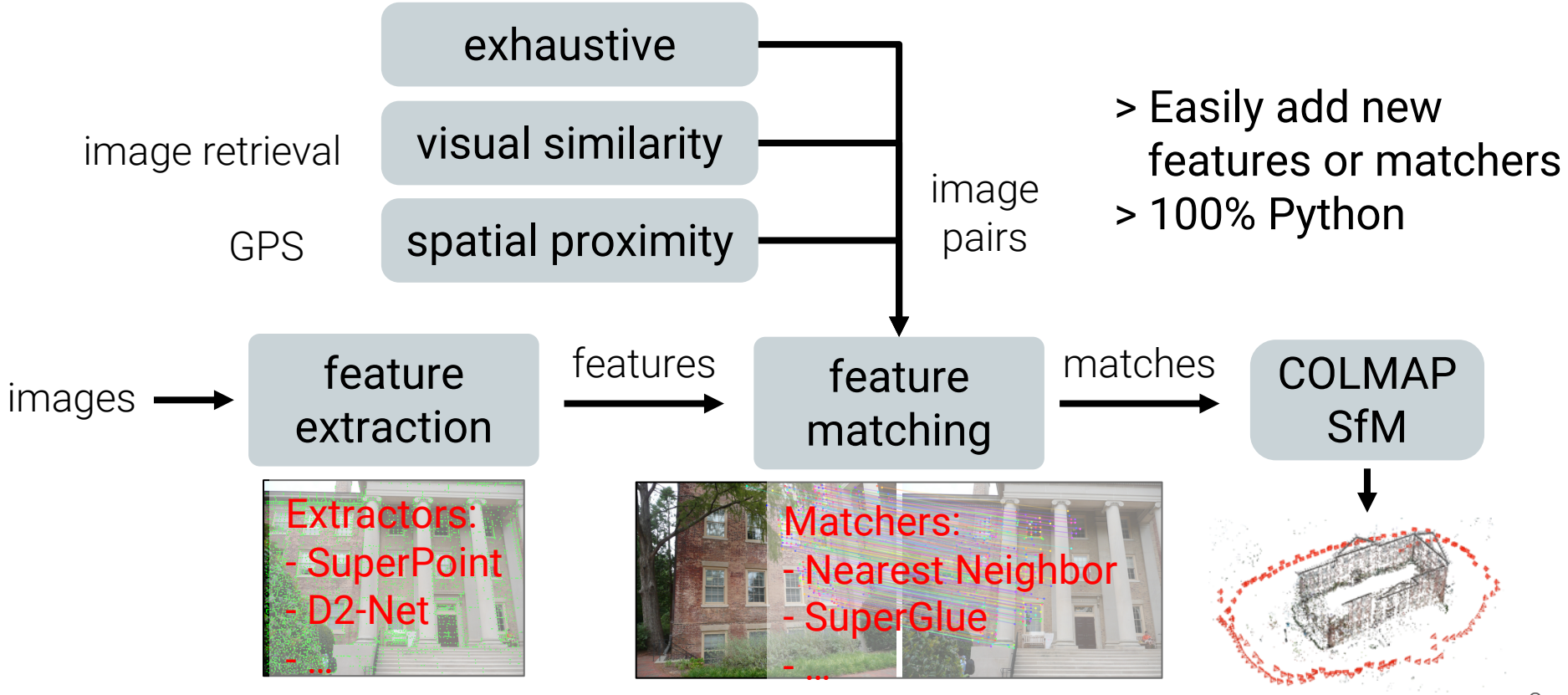
localization

hloc

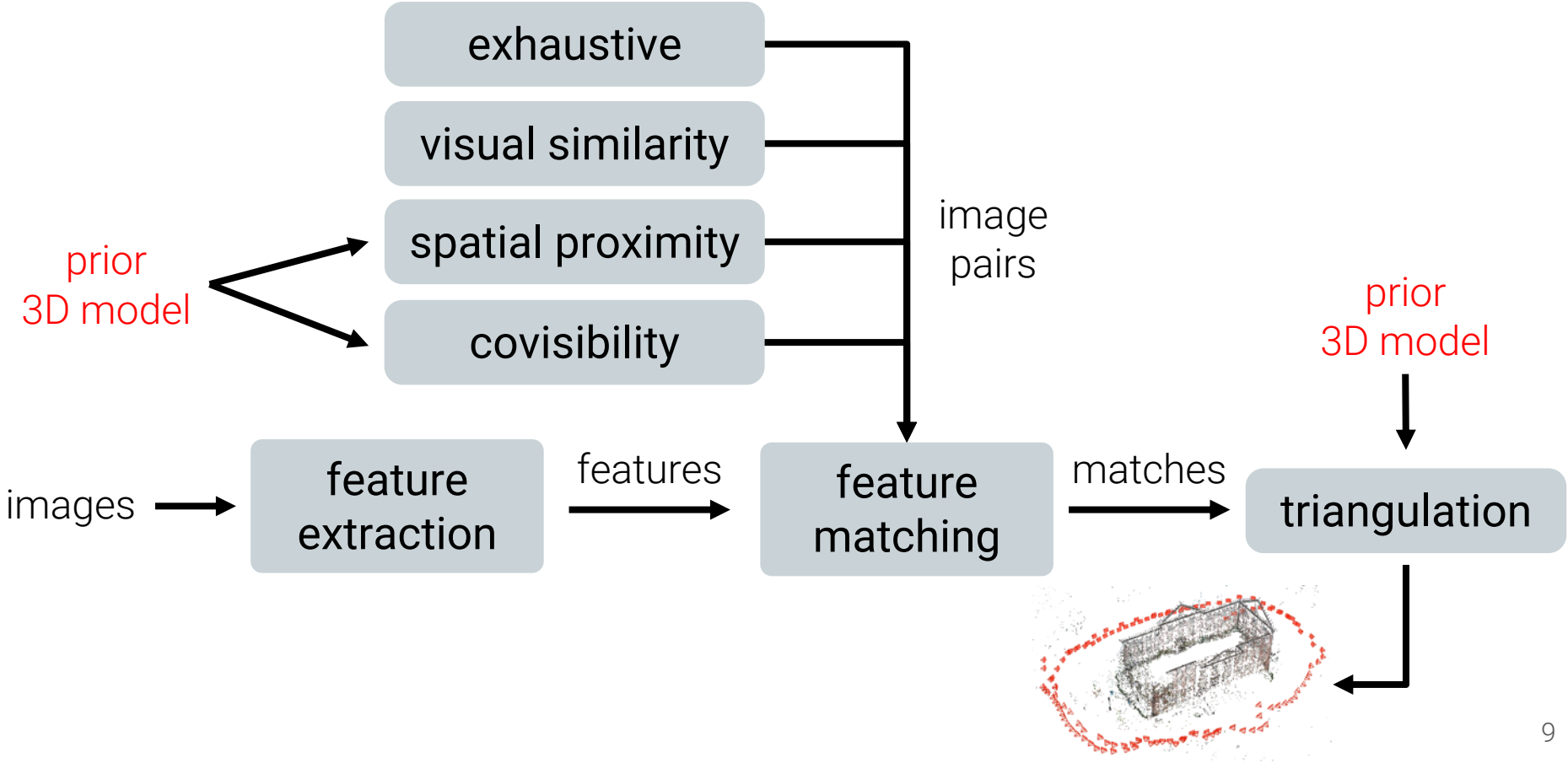


github.com/cvg/Hierarchical-Localization

hloc – reconstruction



hloc – triangulation



hloc – localization

reference

- 3D model
- images
- features

visual similarity

image pairs

query

feature extraction

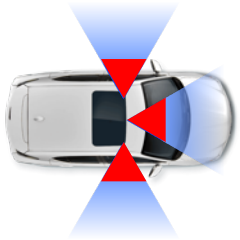
features

feature matching

matches

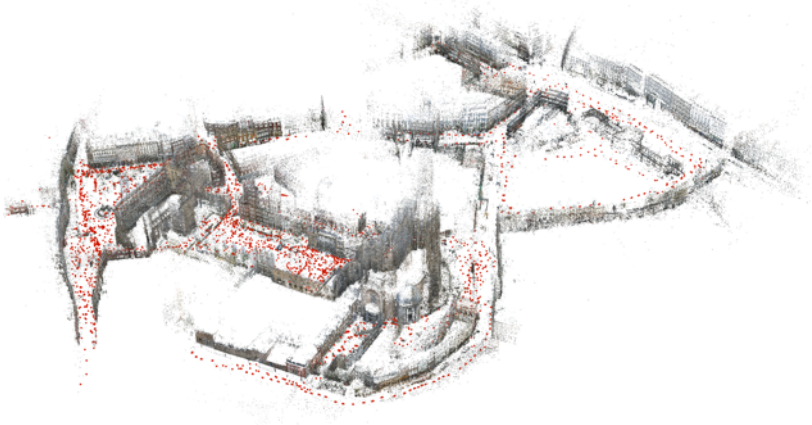
Localization
PnP+RANSAC

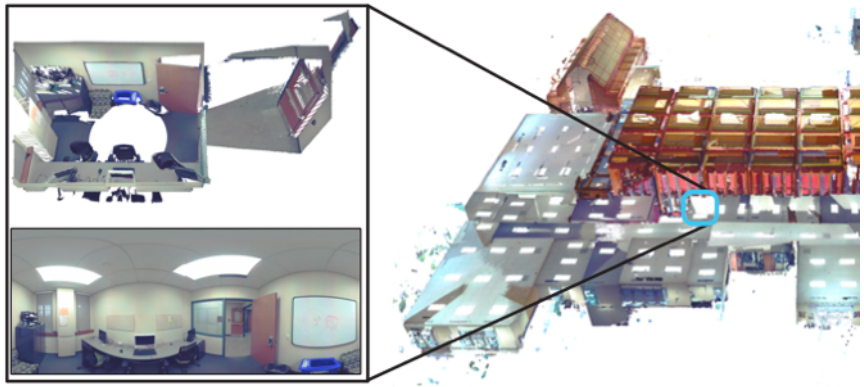
6 DoF pose



Multi-camera rig with generalized PnP

Supported datasets





from https://www.visuallocalization.net/

with 3D SfM model: (outdoor)

- Aachen Day-Night (v1.0 and v1.1)
- RobotCar Seasons
- CMU Seasons
- SILDa (multicamera coming soon)

with 3D Lidar scans: (indoor) InLoc

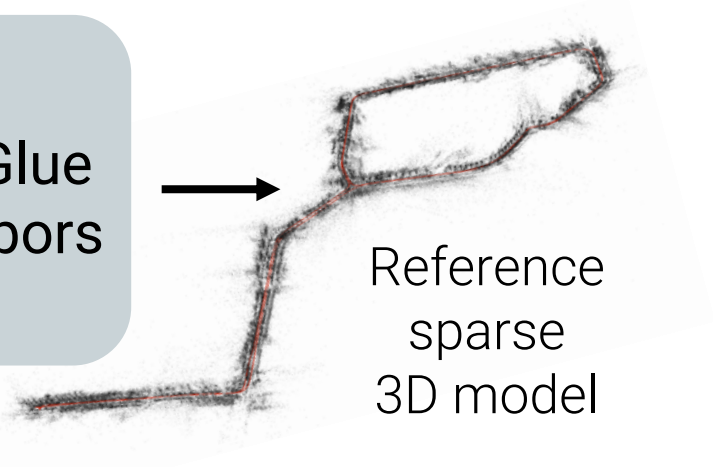
Winning the MLAD Challenge

hloc

reference sequence + camera poses

3D reconstruction

- SuperPoint + SuperGlue
- top 20 spatial neighbors
- triangulation

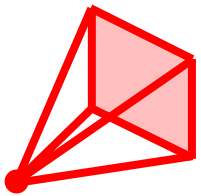


Reference sparse 3D model

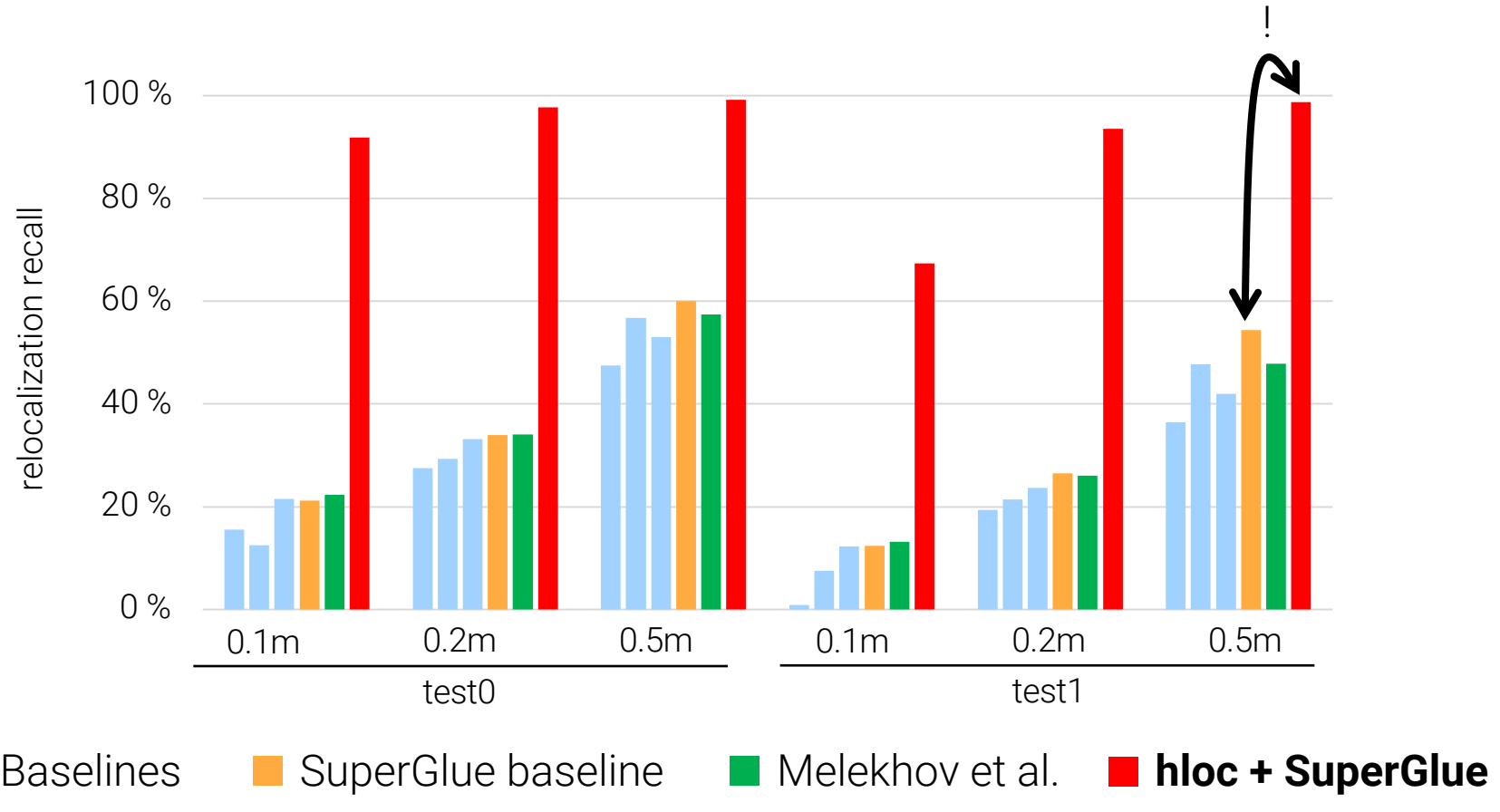
Reference model + Test query image pairs

localization

- top 20 spatial neighbors
- PnP+RANSAC



Winning the MLAD Challenge

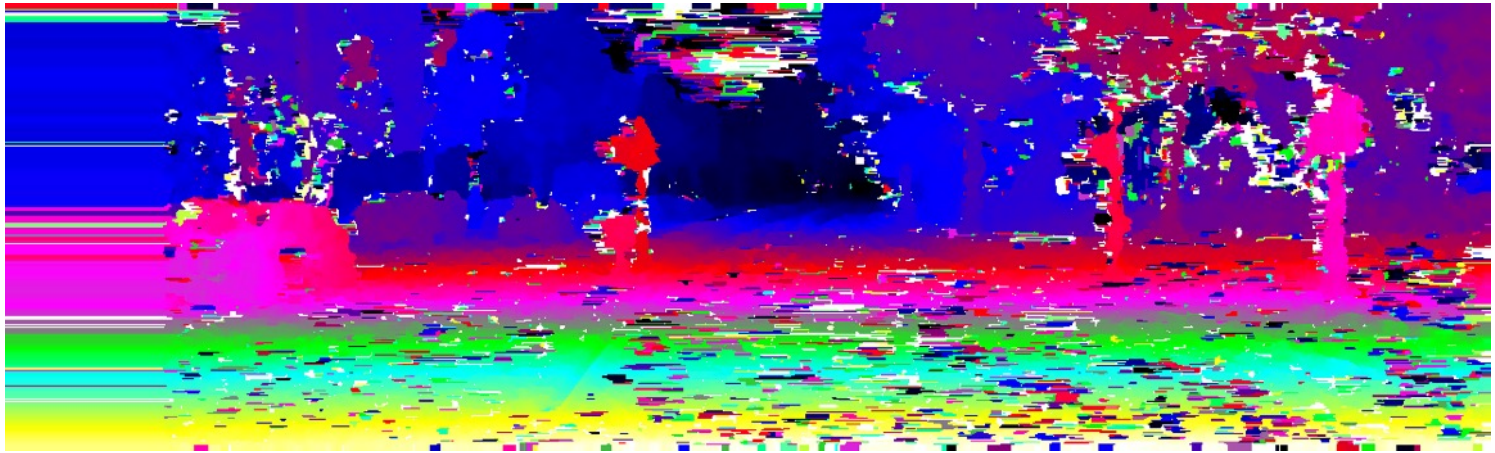


But what are the baselines?

1. Extract features from query and reference images with SuperPoint/D2-Net/R2D2
2. 2D-2D matching with Nearest Neighbor or SuperGlue
3. Compute disparity map with OpenCV stereo matcher
4. Lift ref 2D keypoints to 3D with nearest sampling of disparity
5. Pose estimation with PnP+RANSAC and 2D-3D matches

github.com/pmwenzel/mlad-benchmark-baselines

But what are the baselines?



KITTI Stereo 2015, OpenCV StereoBM [Rank #289]

But what are the baselines?

1. Extract features from query and reference images with SuperPoint/D2-Net/R2D2
2. 2D-2D matching with Nearest Neighbor or SuperGlue
3. Compute disparity map with OpenCV stereo matcher
4. Lift ref 2D keypoints to 3D with nearest sampling of disparity
5. Pose estimation with PnP+RANSAC and 2D-3D matches

vs hloc: triangulation + bundle adjustment

- Is this how GN-Net evaluates SuperPoint and D2-Net ?
- Next time: evaluate local features with hloc :)

What should we work on?

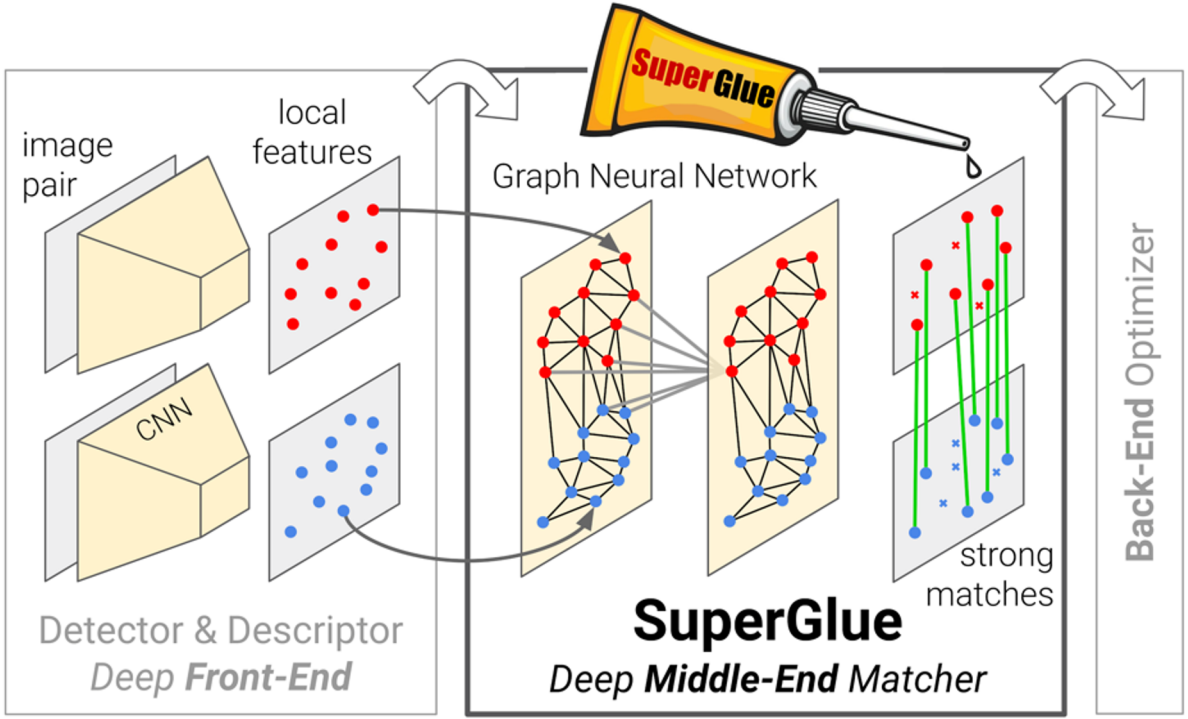
1. Going **beyond keypoints**:
   combining the **robustness** of keypoint-based localization
   with the **accuracy** of direct image alignment
2. More **powerful image retrieval**:
   here we have localization priors, what if we don't?
   image retrieval is **often the bottleneck**
3. Leveraging **sequential information**:
   localize single queries vs sequences

hloc + SuperGlue

hloc = a simple & clean toolbox for visual localization

1. Reproduce our **CVPR & ECCV 2020 winning results**
2. Run SfM with SuperGlue **to localize with your own datasets**
3. **Evaluate your own** local features or image retrieval for localization
4. **Easily implement & debug** new localization pipelines 🔥

github.com/cvg/Hierarchical-Localization



@pesarlin

psarlin@ethz.ch

psarlin.com

Next talk: Friday 28th @ localization workshop

github.com/cvg/Hierarchical-Localization