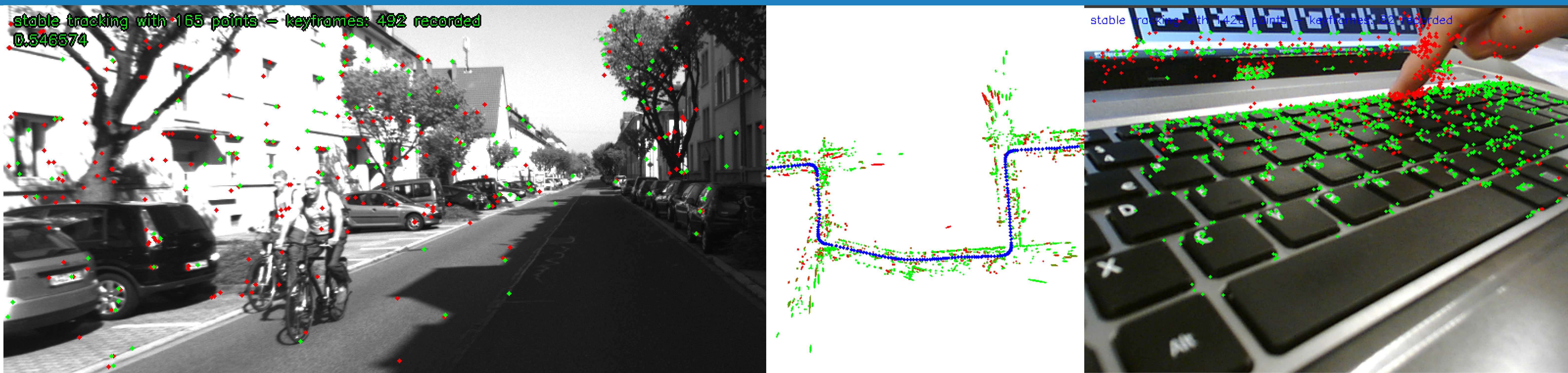




# Fraunhofer

## IOSB

FRAUNHOFER-INSTITUTE OF OPTRONICS, SYSTEM TECHNOLOGIES AND IMAGE EXPLOITATION IOSB



## FLEXIBLE VISUAL SLAM

### Abstract

Visual trajectory estimation and 3D reconstruction is an important task in robotics and has interesting applications like autonomous driving. Current research concentrates on scalability and precision on large environments and real time performance. Here I present you my approach, offering precise real time trajectory estimation while supporting arbitrary camera configurations and efficient loop detection.

### Approach

- Frame to frame motion is estimated via sparse optical flow and RANSAC.
- ORB features serve as landmarks for fast processing and robustness.
- Epipolar guided matching enables support for thousands of keypoints per image.
- Adaptive feature and landmark removal ensures even distribution across the image and removes most outliers.
- PBA offers multithreaded bundle adjustment of independent camera measurements using PCG.
- g2o allows incorporation of additional constraints (e.g. rigid camera setups, IMU and GPS measurements, loop closures).
- Monocular and arbitrary stereo setups are supported.
- Cameras can move independently during reconstruction, while having varying focal lengths, resolution and color modes (RGB or grayscale).
- Robust and fast loop closure detection based on appearance and geometry. A heuristic allows scaling with large maps, no training is required.

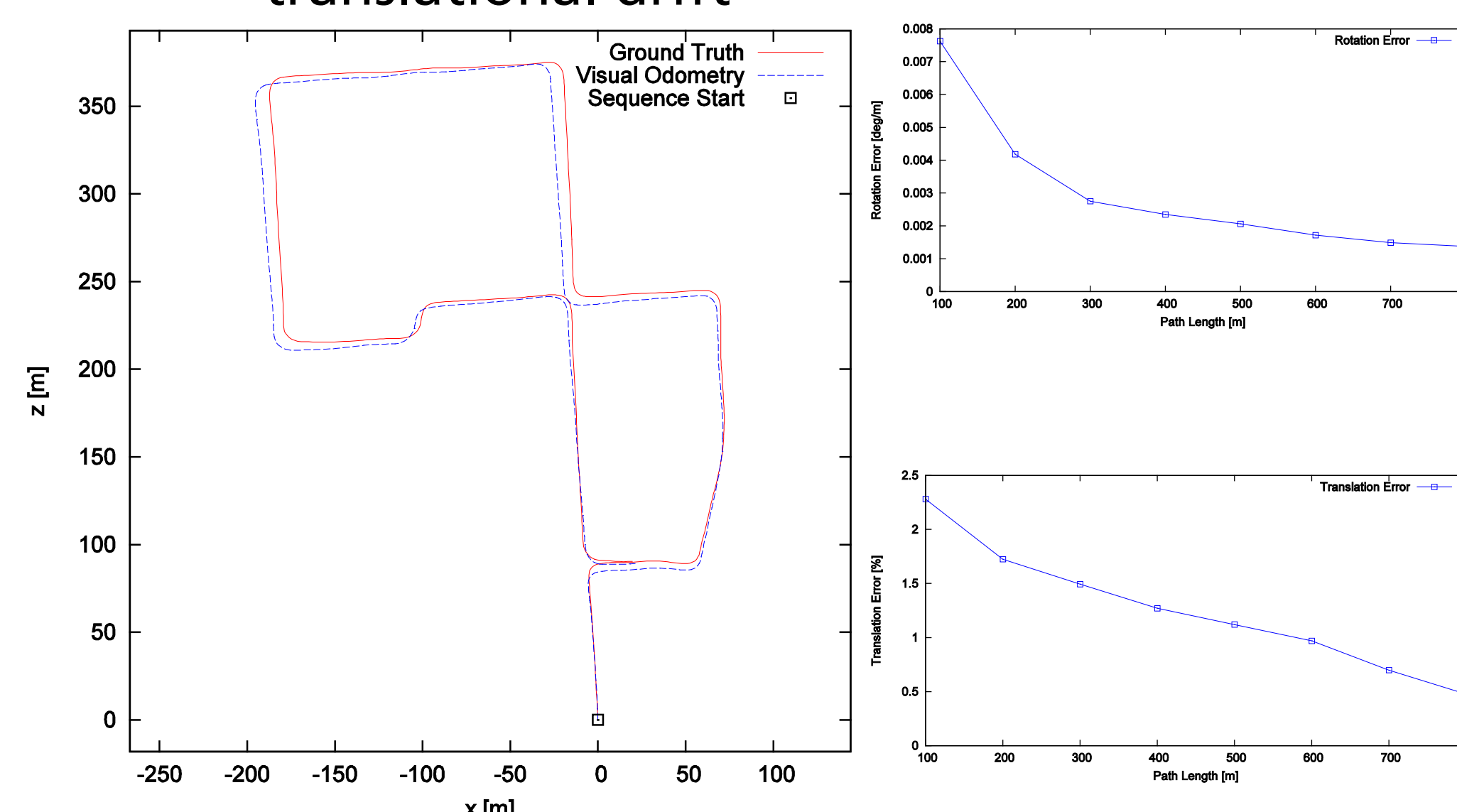
### Test equipment

- High speed cameras have been precisely calibrated for synchronized stereo recording 1280x1024 pixels at 60 FPS.
- An inertial measurement unit and a camera have been autocalibrated using an EKF and will be used for monocular VSLAM.
- The stereo camera was mounted onto a vehicle, recording footage in an urban environment, along with a high resolution panorama camera synchronized with a high accuracy INS.



### Evaluation – Kitti Vision Benchmark

- Specs
  - Car mounted stereo camera
  - Exposure changes and traffic
  - 1241x376 pixels
  - Ground truth from INS and GPS
- Optimization is done with PBA and g2o.
- Stereo constrained bundle adjustment
- Performance
  - ~ 10 FPS (real time)
  - ~ 0.001 °/m rotational and 0.5 % translational drift

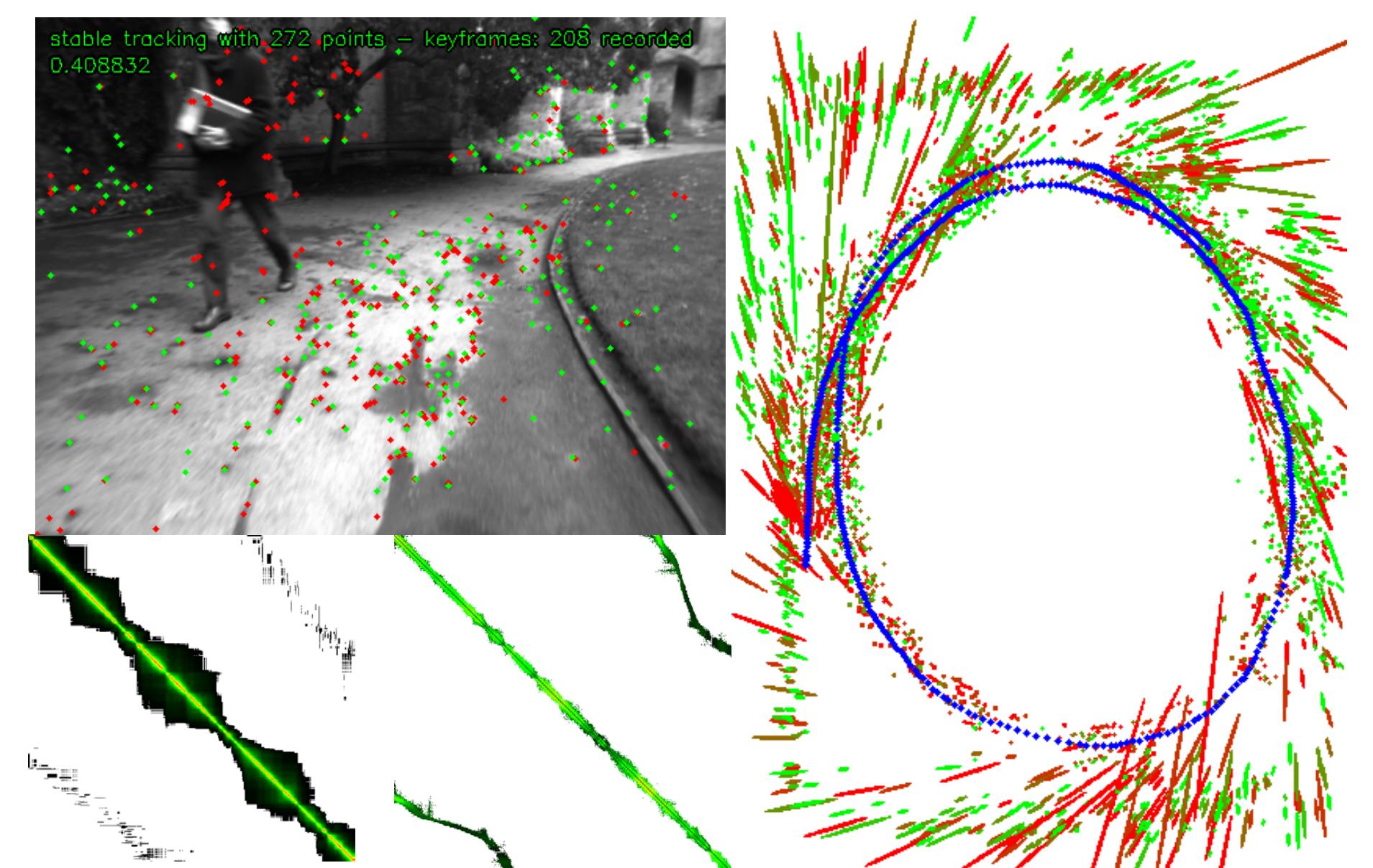


### Evaluation – New Tsukuba

- Specs
  - Synthetic Data
  - Pose and depth ground truth
  - Strong camera panning
  - 640x480 pixels
  - Known stereo baseline used at initialization and keyframe insertion
- Performance
  - above 40 FPS
  - Accuracy comparable to state of the art visual odometry

### Evaluation – New College

- Specs
  - People crossing the path
  - Trajectory forms an exact loop
  - 512x384 pixels
- Performance
  - above 80 FPS (monocular)
  - Loop closure detection in 2 milliseconds per keyframe comparison
  - 575 keyframes with 20k landmarks



Matches between keyframes based on VSLAM and loop closure detection (bottom left). Reconstructed trajectory and landmarks (right).

