# Combining Direct and Indirect Features for Efficient Visual Inertial Odometry

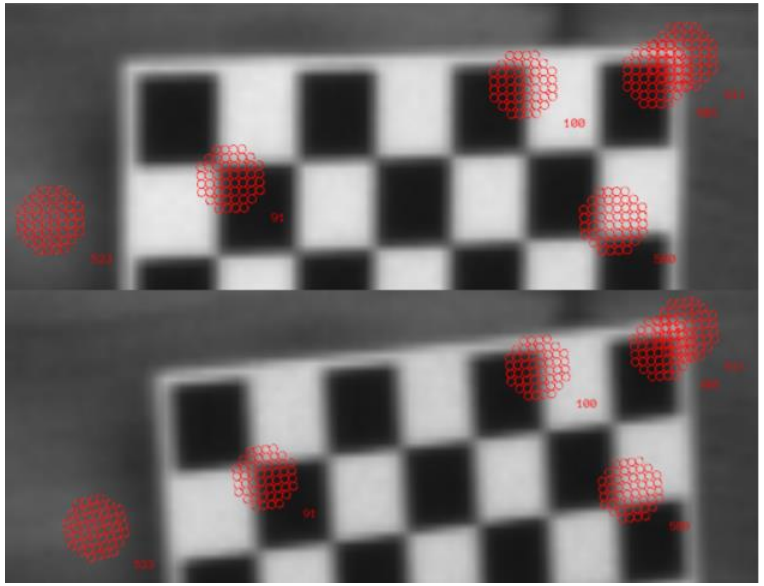
Ahmed Agha

Advisor: Nikolaus Demmel
Supervisor: Prof. Daniel Cremers

# Contents

# Introduction: Indirect Features in Basalt

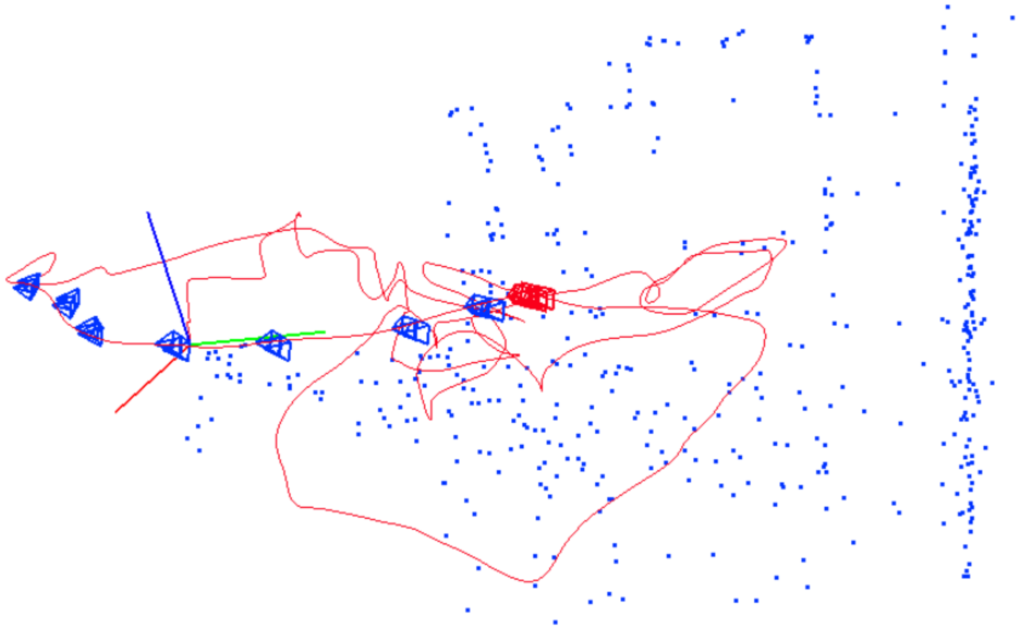
- FAST corner points
- KLT feature tracking
- Inverse distance of the landmark initialized using triangulation



# Introduction: Motion Tracking in Basalt

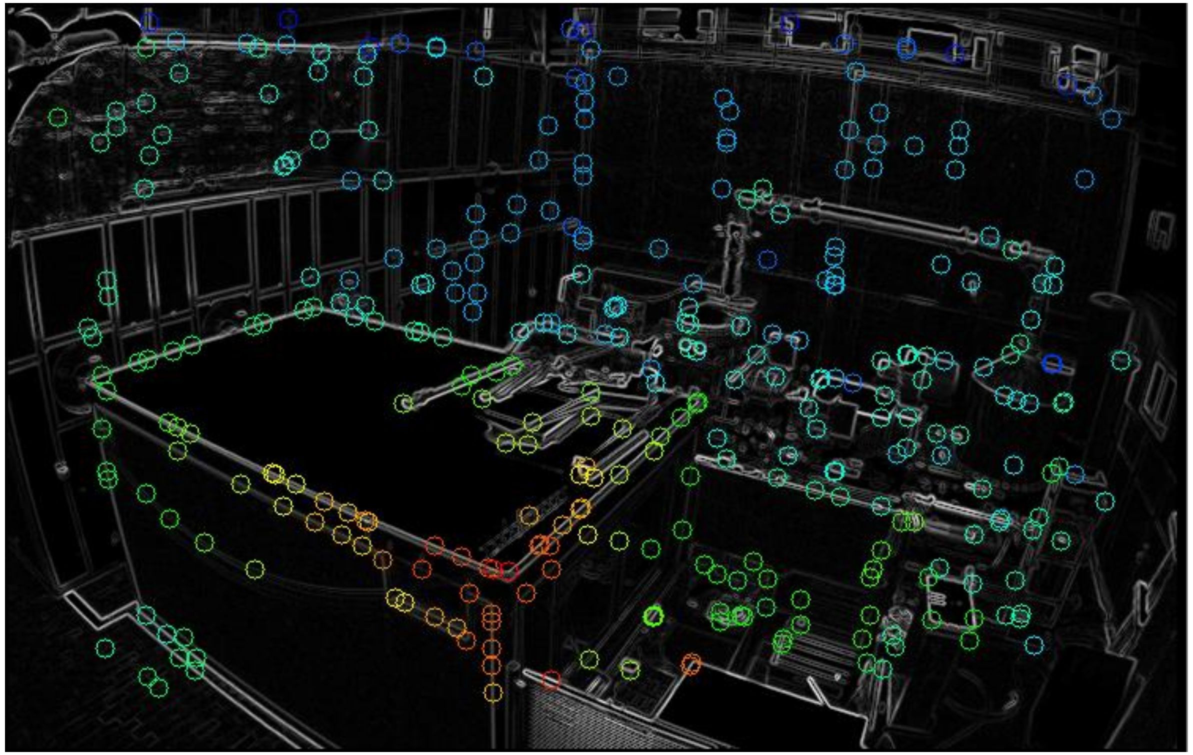
- Bundle adjustment in a sliding window

- The bundle adjustment optimizes poses, 3D positions of landmarks, IMU Biases



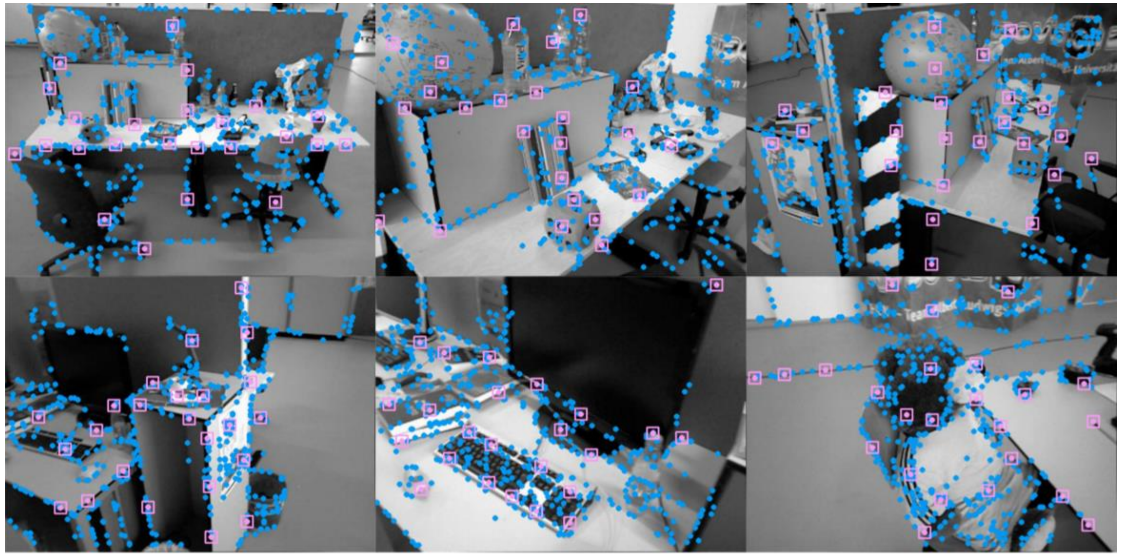
# Introduction: Direct features selection

- Point selection based on their gradients
- Selecting points in grid-based manner
- Inverse distance initialized using static stereo



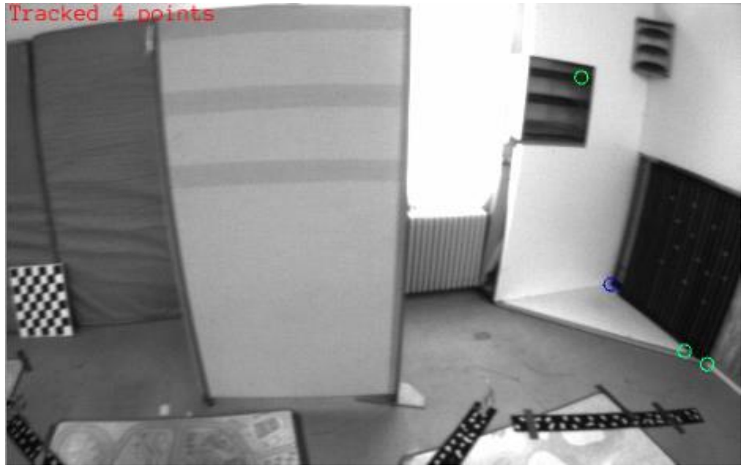
# Introduction: Information Driven Odometry

- Using all direct features is inefficient
- Usually, Heuristics are used for point selection
  - Image Gradients
  - Image Covering
- Select most informative points using an efficient information-based criterion

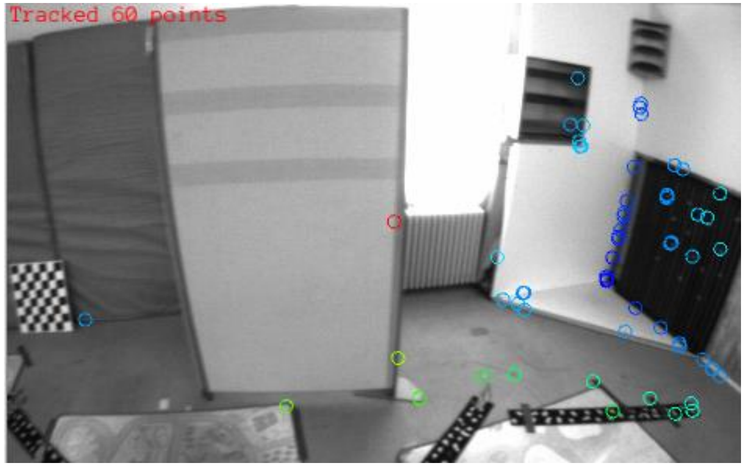


# Motivation

- Complement the indirect features used in Basalt with direct features in an efficient manner



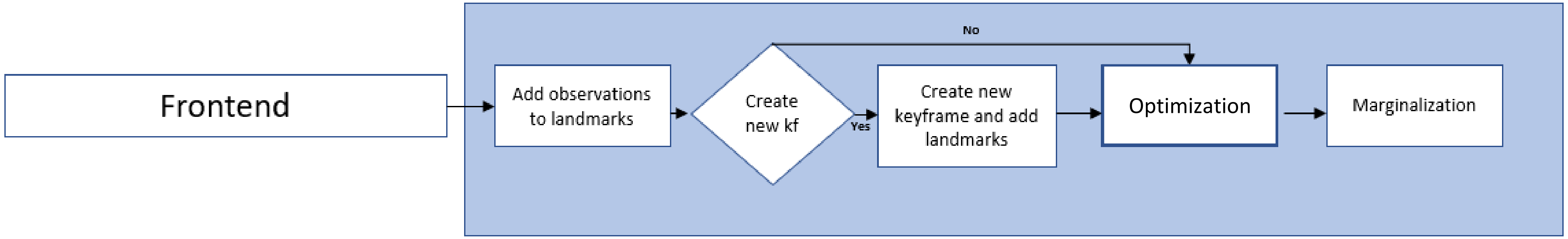

Projection of tracked landmarks when only using indirect features




Projection of tracked landmarks after adding direct features as well

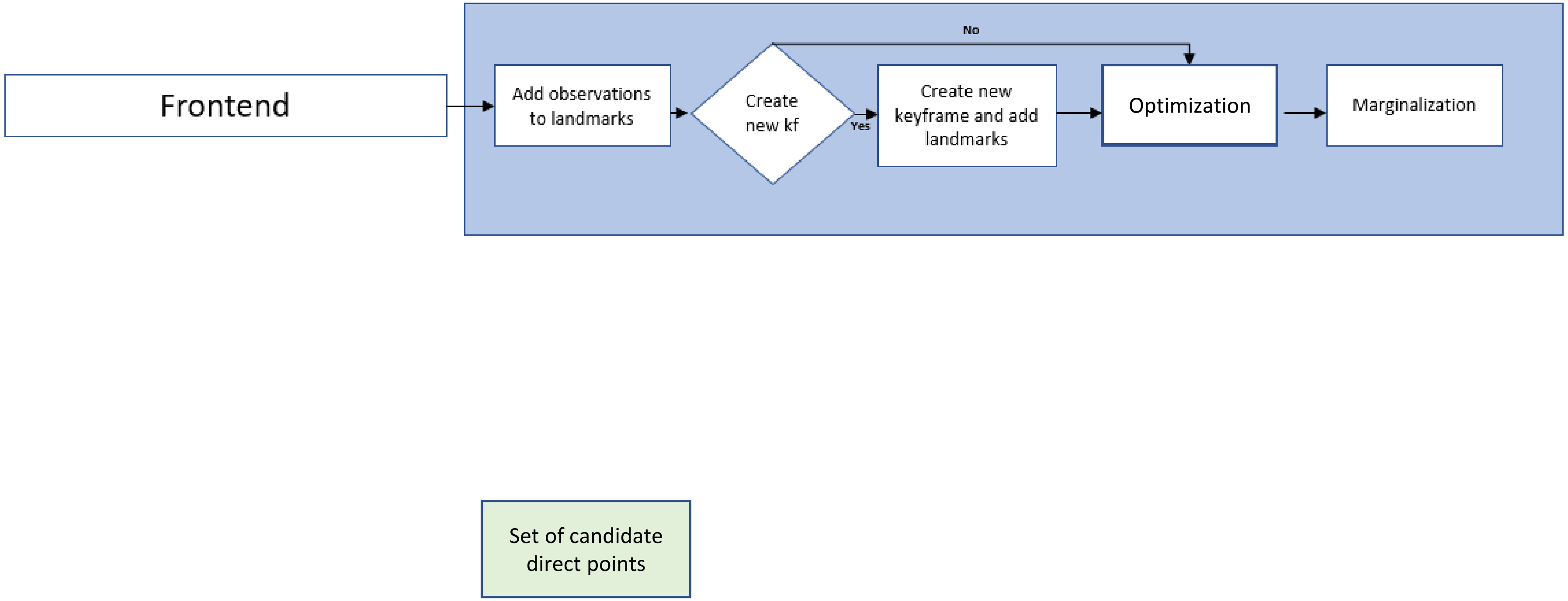
# Overview: Algorithm

- Basalt pipeline



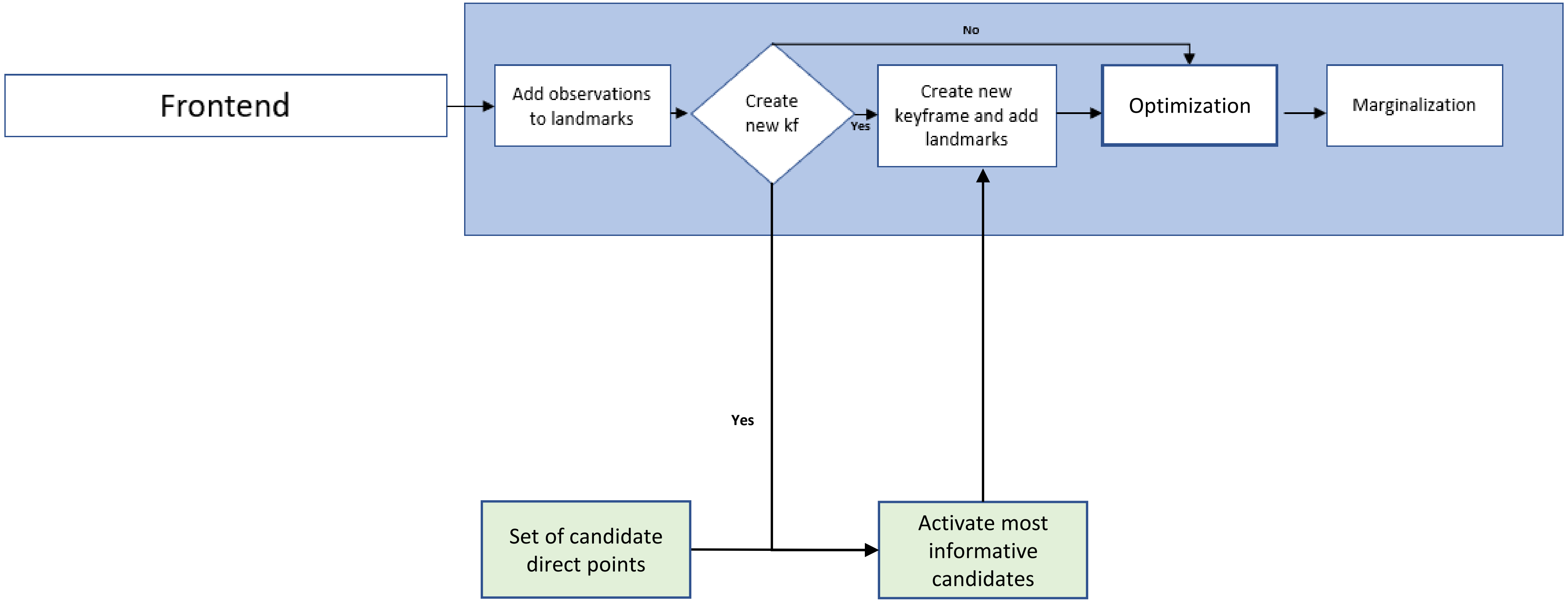
# Overview: Algorithm

- Proposed pipeline



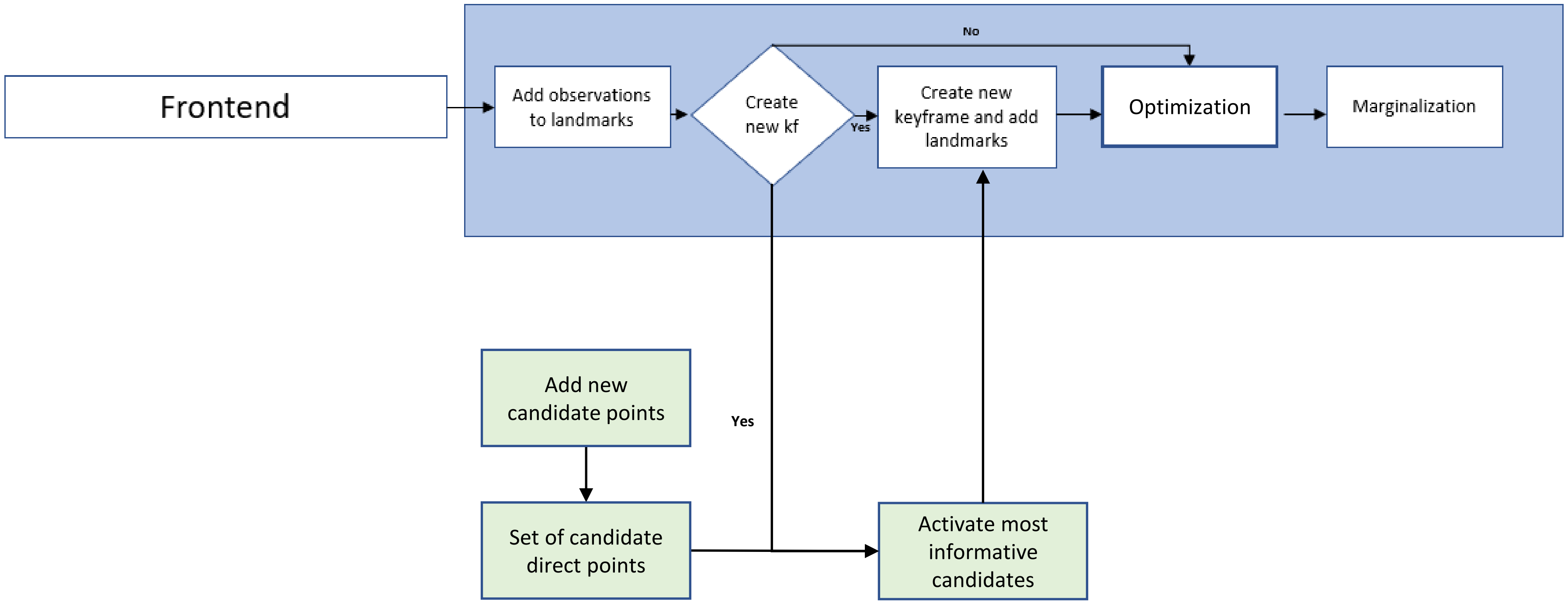
# Overview: Algorithm

- Proposed pipeline



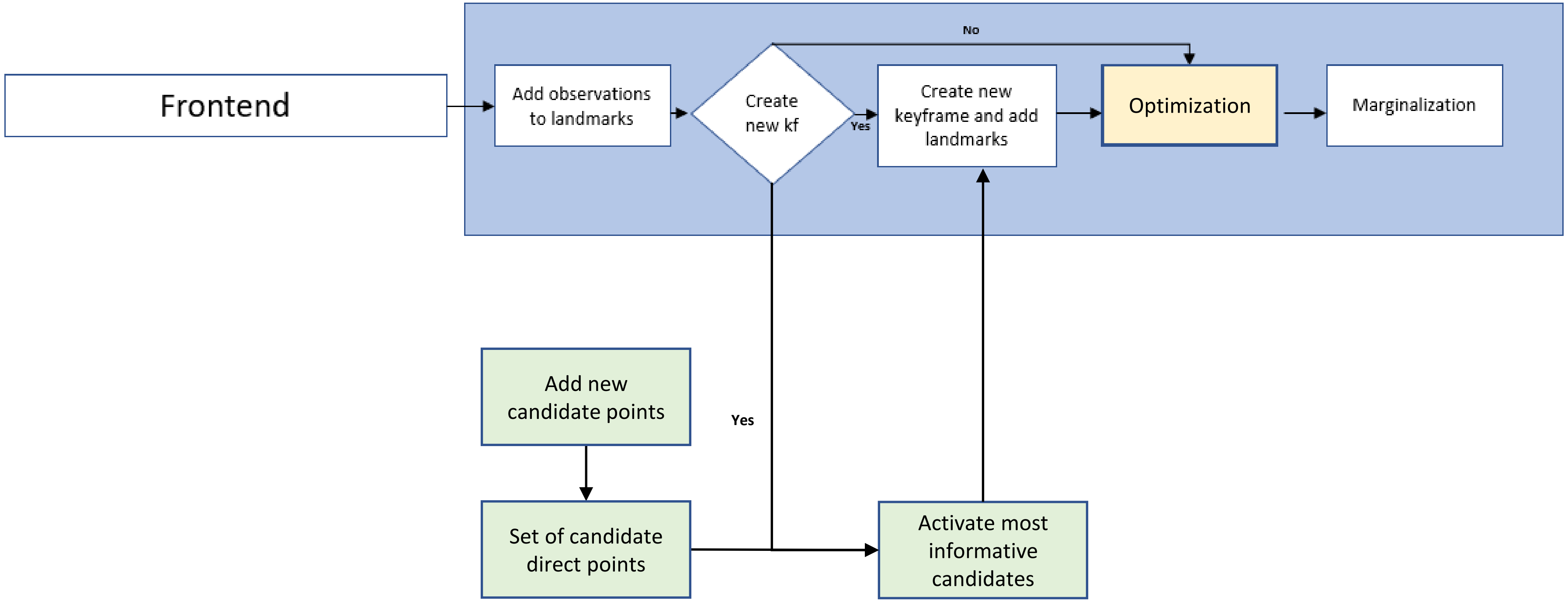
# Overview: Algorithm

- Proposed pipeline



# Overview: Algorithm

- Proposed pipeline



# Optimizing Direct and Indirect Features

- Reprojection error

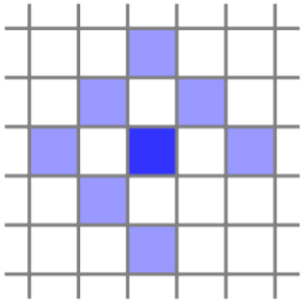
$$E = p_t - \pi(R.P_h + t)$$

- Photometric Error

$$E = I_t(\pi(R\pi^{-1}(p_h, d) + t)) - I_h(p_h)$$

# Photometric Bundle Adjustment

- Photometric residual calculated for an 8-pixel patch
- Each observation of the feature results in an additional residual with 8 dimensions
- Both direct and indirect landmarks are optimized using a photometric error
- Each landmark represented by its inverse distance i.e., each landmark has one degree of freedom



$$E = I_t(\pi(R\pi^{-1}(p_h, d) + t)) - I_h(p_h)$$

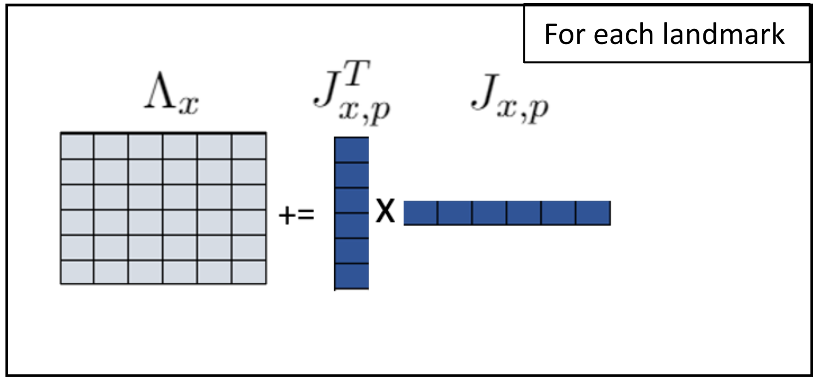
# Selecting Most Informative Points

- Describe the information of the pose using the approximate hessian matrix used in the gauss-newton optimization

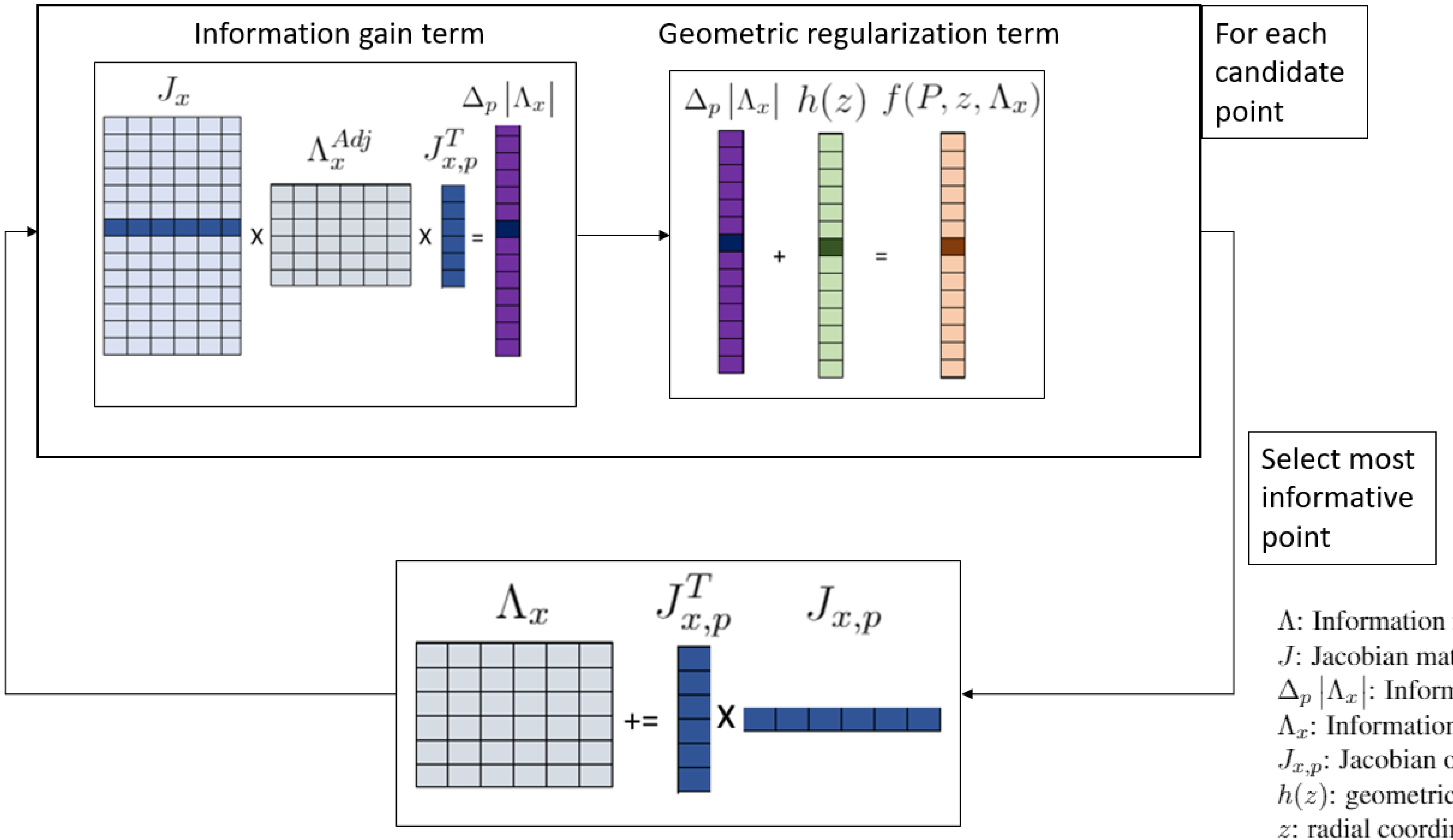
$$\Lambda = J^T J$$

- Select point that would decrease the entropy of the pose

- Initialize the information matrix using existing landmarks



# Selecting Most Informative Points



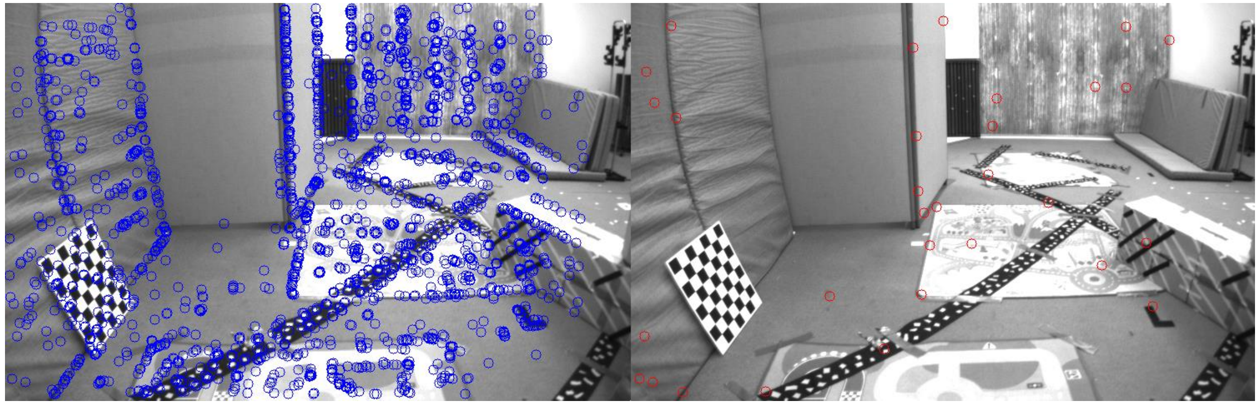

Update information matrix

$\Lambda$: Information matrix
$J$: Jacobian matrix
$\Delta_p |\Lambda_x|$: Information gain of a single point
$\Lambda_x$: Information Matrix
$J_{x,p}$: Jacobian of a single point
$h(z)$: geometric regularization term
$z$: radial coordinates of the point

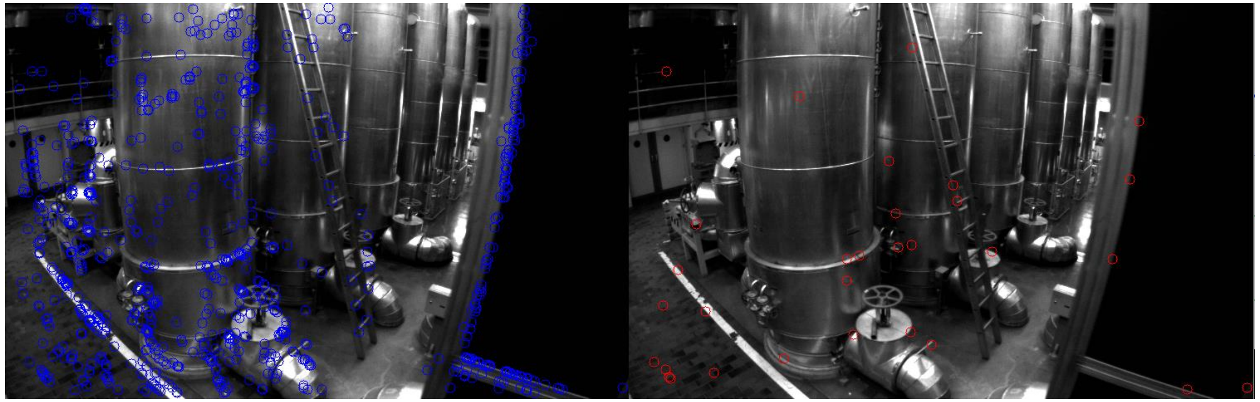
# Selecting Most Informative Points

Candidate points | Selected points

- EuRoC/V1_01



- EuRoC/MH_02



# Robustification

- Outlier observations are rejected based on the median absolute deviation (MAD)
  - Based on photometric residuals
  - Takes place during optimization

$$\begin{aligned}\hat{\sigma}(\mathbf{x}) &= c \cdot \mathrm{MAD} \\ &= c \cdot \mathrm{median}(|\mathbf{x} - \mathrm{median}(\mathbf{x})|)\end{aligned}$$

- Variable Huber Threshold

$$w_h(r_k) = \begin{cases} \frac{1}{\sigma_n^2} & \text{if } |r_k| < \lambda \\ \frac{\lambda}{\sigma_n^2 |r_k|} & \text{otherwise} \end{cases} \qquad \lambda = 1.345\sigma_n$$

# Evaluation

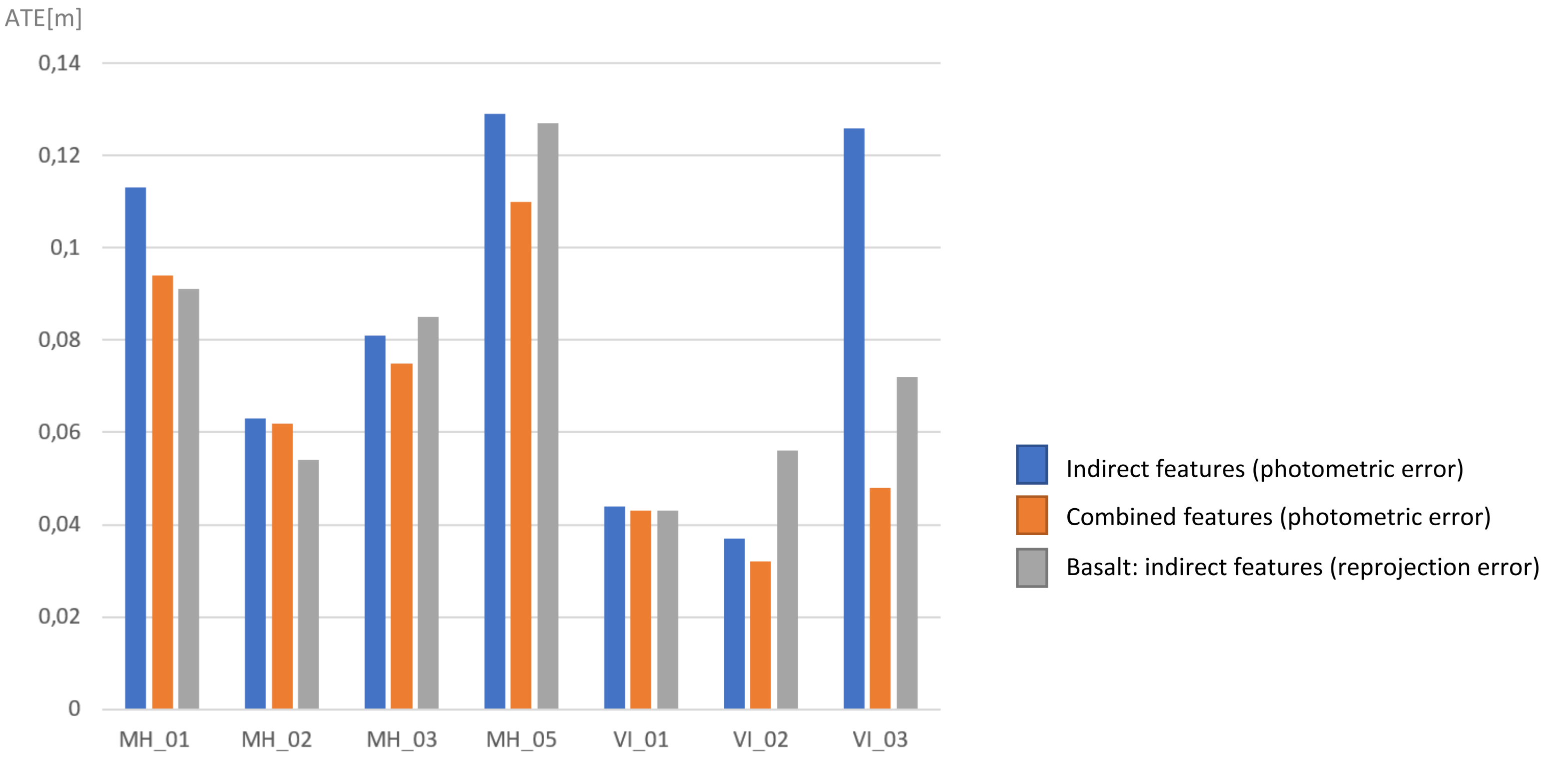
- Metric
  - Absolute Trajectory Error in meters

- Datasets: EuRoC
  - Visual inertial dataset
  - Contains: synchronized stereo images, IMU measurements and ground truth

# Evaluation

- Switching from a reprojection error to a photometric error results in a degradation in performance

- Adding indirect features improves the odometry in 4 out of 7 sequences, but the performance is worse or stays the same in the other 3

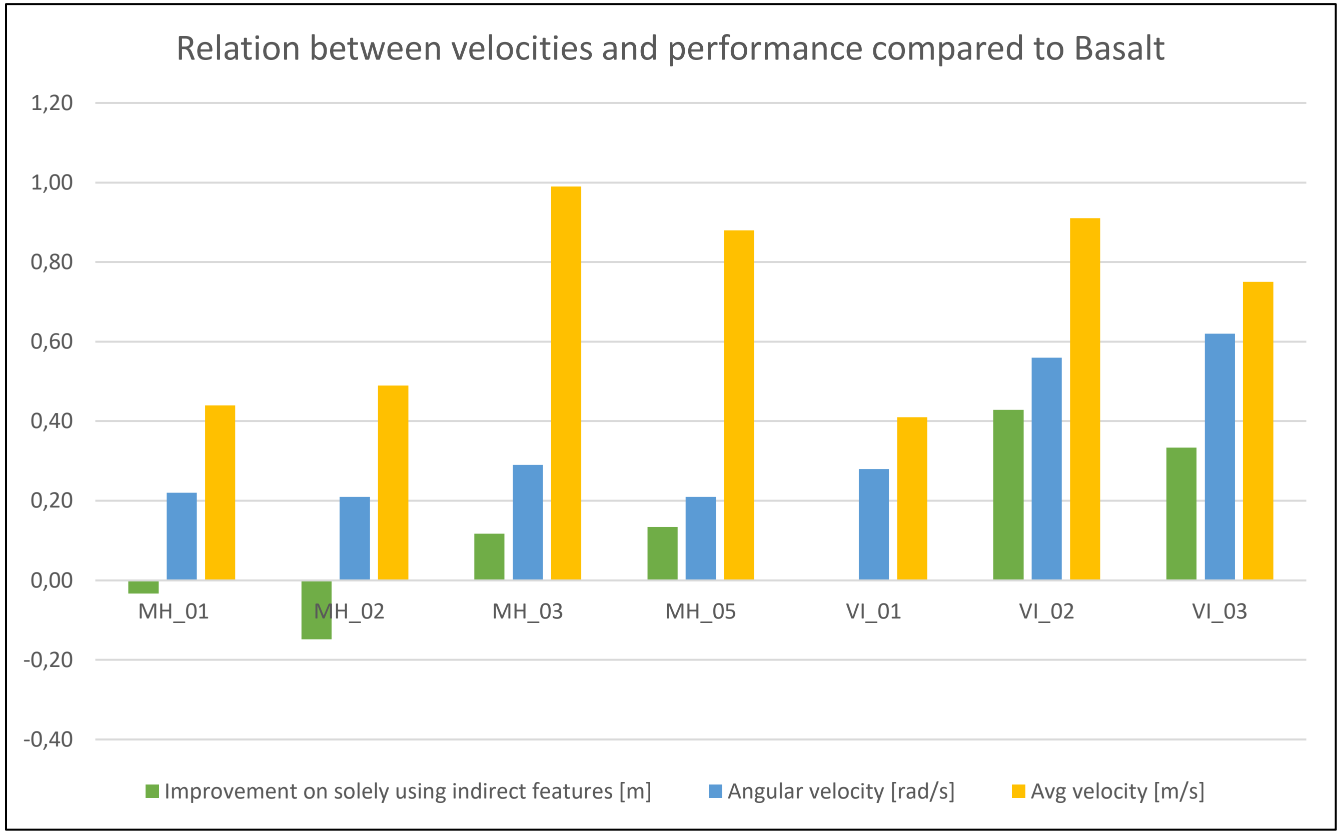


# Evaluation

- Switching from a reprojection error to a photometric error results in a degradation in performance

- Adding indirect features improves the odometry in 4 out of 7 sequences, but the performance is worse or stays the same in the other 3

# Evaluation: When does adding direct features results improve the odometry ?

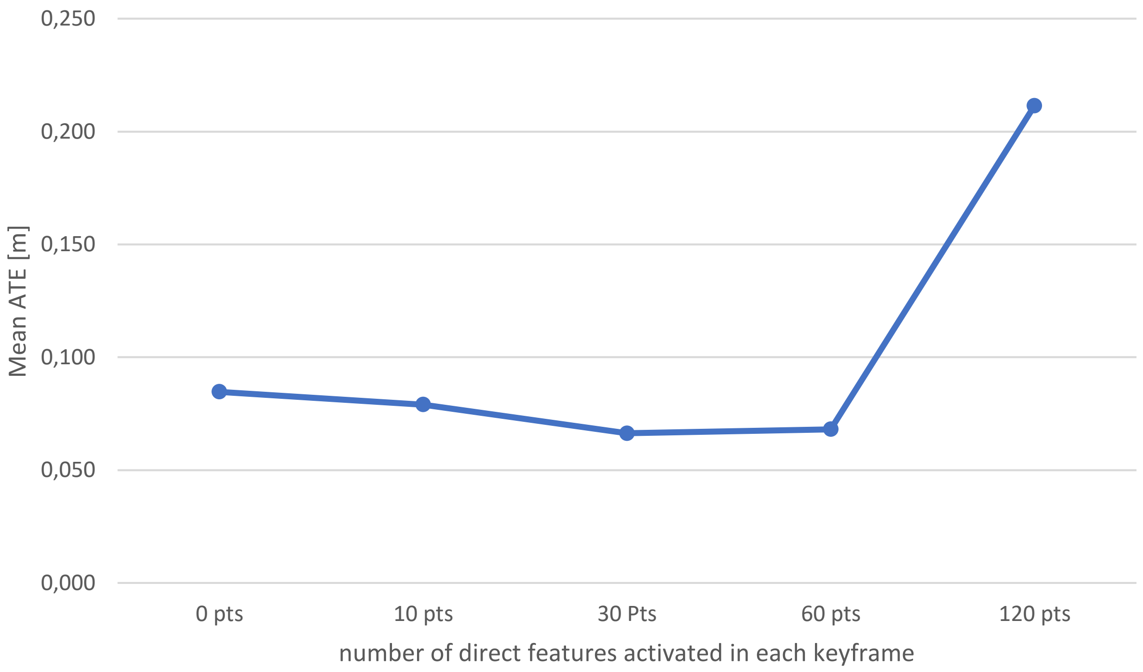



| Name | Length / duration | Avg. vel. / angular vel. | Note |
|---|---|---|---|
| MH_01_easy | 80.6m<br>182s | $0.44 \mathrm{ms}^{-1}$<br>$0.22 \mathrm{rad\,s}^{-1}$ | Good texture, bright scene |
| MH_02_easy | 73.5m<br>150s | $0.49 \mathrm{ms}^{-1}$<br>$0.21 \mathrm{rad\,s}^{-1}$ | Good texture, bright scene |
| MH_03_medium | 130.9m<br>132s | $0.99 \mathrm{ms}^{-1}$<br>$0.29 \mathrm{rad\,s}^{-1}$ | Fast motion, bright scene |
| MH_05_difficult | 97.6m<br>111s | $0.88 \mathrm{ms}^{-1}$<br>$0.21 \mathrm{rad\,s}^{-1}$ | Fast motion, dark scene |
| V1_01_easy | 58.6m<br>144s | $0.41 \mathrm{ms}^{-1}$<br>$0.28 \mathrm{rad\,s}^{-1}$ | Slow motion, bright scene |
| V1_02_medium | 75.9m<br>83.5s | $0.91 \mathrm{ms}^{-1}$<br>$0.56 \mathrm{rad\,s}^{-1}$ | Fast motion, bright scene |
| V1_03_difficult | 79.0m<br>105s | $0.75 \mathrm{ms}^{-1}$<br>$0.62 \mathrm{rad\,s}^{-1}$ | Fast motion, motion blur |

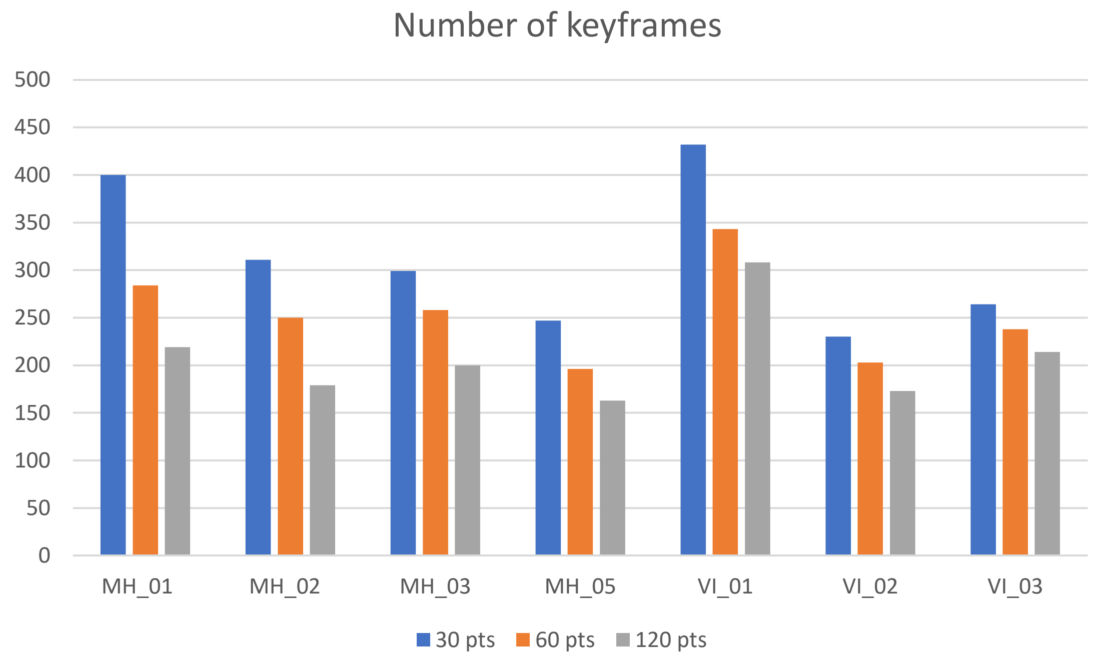
# Evaluation: Comparing different number of activated candidates in each keyframe

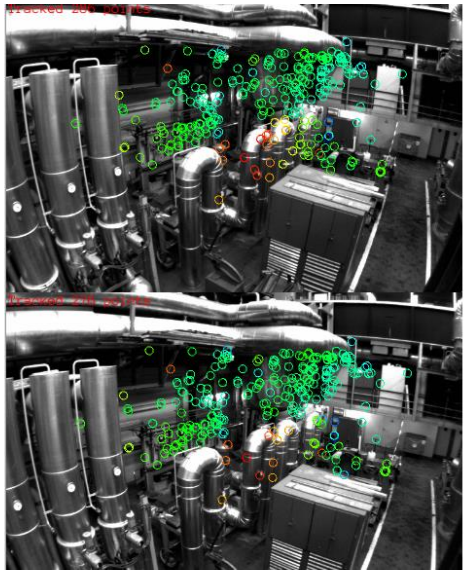
- We expect performance to get better as the number of activated per keyframe increases

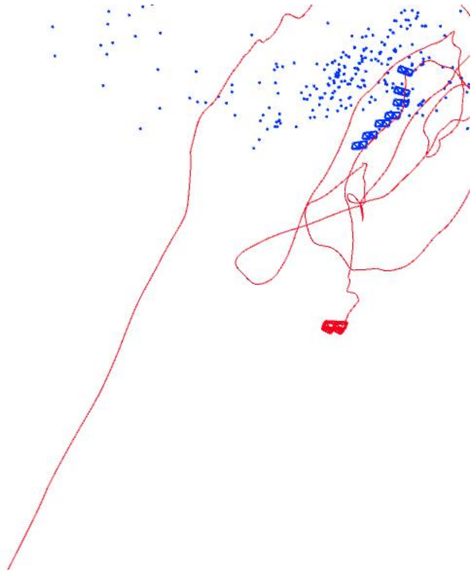
- Possible reasons:
  - Simplistic outlier rejection
  - Keyframing



# Evaluation: Comparing different number of activated candidates in each keyframe





# Conclusion

- Complement indirect features with direct features to improve odometry
- Promising results in sequences with high velocities

# Future Work

- Improve outlier rejection
- New keyframing strategy, possibly using information-based metrics
- New landmark activation strategy to target always having a constant number of landmarks active
- Refining inverse distance of candidate points using temporal stereo

# Thank You For Listening!