Feature-based initialization for Monocular Direct Visual Odometry

Interdisciplinary project (IDP)

Mariia Gladkova

Advisor: Nikolaus Demmel
Supervisor: Prof. Daniel Cremers
Contents

1. Motivation and DSO initialization approach
2. Implemented pipeline
   a. ORB-SLAM initialization
   b. Our optimizations and outlier removal
3. Evaluation
4. Conclusion
DSO Coarse Initializer

1. Select points in a grid-based manner
2. Initialize depth map to 1
3. Iteratively optimize depth map in a coarse-to-fine fashion
DSO Coarse Initializer

1. Select points in a grid-based manner
2. Initialize depth map to 1
3. Iteratively optimize depth map in a coarse-to-fine fashion

Needs good map initialization for convergence!
Project idea

- Address DSO initialization in a feature-based manner

1. Implement robust feature-based initializer
2. Evaluate its performance in comparison with DSO Coarse Initializer

Our overall pipeline

1. Relative pose estimation with homography and fundamental matrix models
   a. Extension to ORB SLAM initialization approach
2. Geometric Bundle Adjustment with outlier removal
3. Additional point extraction with depth initialization using epipolar line search
4. Photometric Bundle Adjustment with outlier removal
Relative pose estimation

1. Find feature correspondences between reference and current frames
2. Fit two models in parallel: a homography $H_{cr}$ and a fundamental matrix $F_{cr}$ (RANSAC scheme)
3. Select based on the score ratio and recover the transformation
4. Perform full Bundle Adjustment

$$x_c = H_{cr}x_r \quad \rightarrow \quad \text{score} \quad S_H$$

$$x_c^T F_{cr} x_r = 0 \quad \rightarrow \quad \text{score} \quad S_F$$

$$R_H = \frac{S_H}{S_H+S_F}$$

Relative pose estimation - changes to ORB SLAM initialization

- Several changes to the ORB SLAM initialization:
  - Extracting corners and tracking them with Lucas Kanade optical flow
  - Attempting to fit homography and fundamental matrix after \( N - 1 \) frames

"Two-way" tracking
Point augmentation

- Grid-based point selection
- 1D search along epipolar line
  - depth prior from triangulated features
  - SSD error over 3 x 3 pixel patch
- Uncertainty propagation through N frames
  - Geometric and photometric components of depth error-variance [1]

Geometric Bundle Adjustment

- Geometric Bundle Adjustment
- Outlier removal based on reprojection error
  - scheme: removal - optimization - removal - optimization

Photometric Bundle Adjustment

- Photometric Bundle Adjustment
  - 8-pixel patch
  - mean intensity normalization

\[
\min_{\{\xi_j\}_{j=1\ldots|c|}, \{x_i\}_{i=1\ldots|p|}} \sum_{i=1}^{|p|} \sum_{j \in \text{obs}(x_i)} \sum_{\Delta \in \mathcal{N}(p_i)} \left\| I_j(\pi(\xi_j(x_i + \pi^{-1}\Delta))) - \psi I(\pi x_i + \Delta) \right\|_\gamma
\]

with \( \psi = \frac{\sum_{\Delta \in \mathcal{N}(p)} I_j(\pi T_j(x + \pi^{-1}\Delta))}{\sum_{\Delta \in \mathcal{N}(x)} I(\pi x + \Delta)} \)

Outlier removal

- T-distribution of residuals
  - mean, variance and degrees of freedom
  - 8 residuals per point

\[
\sigma^2_{k+1} = \frac{1}{n} \sum_{i=1}^{n} \frac{\nu+1}{\nu+r_i^2/\sigma_k^2} r_i^2
\]

Outlier removal

● T-distribution of residuals
  ○ mean, variance and degrees of freedom
  ○ 8 residuals per point

Evaluation setup

1. KITTI dataset (sequences 00 - 10)
2. EuRoC dataset [*]

- initialization accuracy and overall robustness
  - trajectory error over 200 frames
  - different starting points for each run

[*] MH_01_easy, MH_02_easy, MH_03_medium, MH_04_difficult, MH_05_difficult, V1_01_easy, V1_02_medium, V1_03_difficult, V2_01_easy, V2_02_medium, V2_03_difficult
Evaluation setup : metrics used

\[ ATE_{1...n} = \left( \frac{1}{n} \sum_{i=1}^{n} \| \text{transl}(F_i) \|_2^2 \right)^{1/2} \]

where \[ F_i = Q_i^{-1} S P_i \]

\[ RPE_{1...n} = \left( \frac{1}{n-1} \sum_{i=1}^{n-1} \| \text{transl}(E_{i,i+1}) \|_2^2 \right)^{1/2} \]

where \[ E_{i,j} = (Q_i^{-1} Q_j)^{-1}(P_i^{-1} P_j) \]


[2] Choice of the metrics: ATE and number of frames used for initialization are based on the evaluation done by Xingwei Qu in his master thesis Initialization Methods for Visual and Visual-inertial SLAM
Evaluation (initialization)

<table>
<thead>
<tr>
<th></th>
<th>KITTI</th>
<th>EuRoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse Initiator</td>
<td>7.3</td>
<td>7.82</td>
</tr>
<tr>
<td>ORB Initiator</td>
<td>4.78</td>
<td>5.28</td>
</tr>
</tbody>
</table>
Evaluation (ATE)
Evaluation (RPE)
Evaluation with work by Xingwei Qu

- KITTI sequences (00, 01, 02)

<table>
<thead>
<tr>
<th></th>
<th>Coarse Initializer</th>
<th>Xingwei's ORB Initializer</th>
<th>Our ORB Initializer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number frames</td>
<td>7.0</td>
<td>12.367</td>
<td>4.8</td>
</tr>
</tbody>
</table>
Conclusion and future work

- Improvement of runtime
- Further tuning of the outlier removal parameters to boost PBA performance
- Optimization of affine light transform parameters
- Learned features (e.g. SuperPoint)
Thank you for your attention!
Appendix: Pipeline overview

\[
\text{if } \left| \text{frame.frame_id} - \text{ref.frame_id} \right| \geq N \text{ then}
\]
\[
\quad \text{success} \leftarrow \text{map.computeRelativeTransform(frame, ref.frame)};
\]
\[
\quad \text{if success then}
\]
\[
\quad \quad \text{do}
\]
\[
\quad \quad \quad \text{map.performGeometricBA();}
\]
\[
\quad \quad \quad \text{severeOutliers} \leftarrow \text{map.removeSevereOutliers()};
\]
\[
\quad \quad \text{while severeOutliers} > 0;
\]
\[
\quad \quad \text{map.populateStructure3D();}
\]
\[
\quad \quad \text{do}
\]
\[
\quad \quad \quad \text{map.performGeometricBA();}
\]
\[
\quad \quad \quad \text{severeOutliers} \leftarrow \text{map.removeSevereOutliers()};
\]
\[
\quad \quad \text{while severeOutliers} > 0;
\]
\[
\quad \quad \text{if map.performPhotometricBA() then}
\]
\[
\quad \quad \quad \text{map.fitTDistribution();}
\]
\[
\quad \quad \quad \text{map.performPhotometricBA()};
\]
\[
\quad \quad \text{return SUCCESS;}
\]
\[
\text{return FAILURE;}
\]
Appendix: evaluation of ORB Initializer versions (initial pose)
Appendix: evaluation of ORB Initializer versions (ATE)
Appendix: evaluation of ORB Initializer versions (RPE)
Failed cases: wrong depth estimation
Failed cases: incorrect model chosen