Batched Incremental Structure-from-Motion

Hainan Cui, Shuhan Shen, Xiang Gao, and Zhanyi Hu
NLPR, Institute of Automation, Chinese Academy of Sciences
Incremental Structure-from-Motion

1. Image Feature Detection and Matching
2. Epipolar Geometry Graph Construction
3. Camera Poses Initialization
4. Bundle Adjustment: Refine Camera Poses and Scene Structure

Drawbacks:
1. Sensitive to Seed Selection
2. Error Accumulation and Drift
3. Poor Efficiency and Scalability

Additional:
1. Seed Reconstruction
2. Camera Registration
3. BA and Outliers Filtering
Repeat 2 and 3
Batched Incremental SfM

- Image Feature Detection and Matching
- Epipolar Geometry Graph Construction
- Camera Poses Initialization
- Bundle Adjustment: Refine Camera Poses and Scene Structure

**Strengths:**
1. Smaller Error Accumulation
2. Less Bundle Adjustment
3. Faster Scene Reconstruction

**Two Iteration Loops**
- **Inner Loop:** Tracks Triangulation, Selection and Bundle Adjustment
- **Outer Loop:** Batched Camera Registration and Filtering
Main Contributions

- Batched camera registration
- Tracks selection
Quantitative Evaluation

- Three median-scale benchmark image datasets

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (mm — deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FountainP11</td>
</tr>
<tr>
<td></td>
<td>$C_{err}$</td>
</tr>
<tr>
<td>Bundler [35]</td>
<td>7.0</td>
</tr>
<tr>
<td>VSFM [42]</td>
<td>36.0</td>
</tr>
<tr>
<td>Theia [38]</td>
<td>1.9</td>
</tr>
<tr>
<td>COLMAP [34]</td>
<td>4.9</td>
</tr>
<tr>
<td>our BSfM</td>
<td><strong>1.9</strong></td>
</tr>
</tbody>
</table>

Quantitative Evaluation

- One large-scale image dataset Quad
  - Median error: DISCO 1.16m, Bundler 1.01m, VSFM 0.89m, COLMAP 0.85m, and our BSfM 0.69m.

  ![Graphs](image)

  - a: with tracks selection (257K points)
  - b: without tracks selection (3048K points)

## Qualitative Evaluation

- **Sequential image datasets**

<table>
<thead>
<tr>
<th></th>
<th>Bundler</th>
<th>VSFM</th>
<th>Theia</th>
<th>COLMAP</th>
<th>Our BSfM</th>
<th>Our BSfM + Full Tri</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Building</strong></td>
<td>![Building.png]</td>
<td>![Building.png]</td>
<td>![Building.png]</td>
<td>![Building.png]</td>
<td>![Building.png]</td>
<td>![Building.png]</td>
</tr>
<tr>
<td><strong>Campus</strong></td>
<td>![Campus.png]</td>
<td>![Campus.png]</td>
<td>![Campus.png]</td>
<td>![Campus.png]</td>
<td>![Campus.png]</td>
<td>![Campus.png]</td>
</tr>
<tr>
<td><strong>Street View</strong></td>
<td>![Street View.png]</td>
<td>![Street View.png]</td>
<td>![Street View.png]</td>
<td>![Street View.png]</td>
<td>![Street View.png]</td>
<td>![Street View.png]</td>
</tr>
</tbody>
</table>

## Qualitative Evaluation

- **Unordered image datasets**

<table>
<thead>
<tr>
<th></th>
<th>Bundler</th>
<th>VSFM</th>
<th>Theia</th>
<th>COLMAP</th>
<th>Our BSfM</th>
<th>Our BSfM + Full Tri</th>
</tr>
</thead>
</table>

Qualitative Evaluation

• More results

(a) Graham Hall
(b) Gendarmenmarkt
(c) Roman Forum
(d) Trafalgar
Conclusion

• We propose a novel batched incremental SfM algorithm to tackle the **efficiency** and **scalability** challenges in a unified framework.

• **Batched camera registration** registers many cameras at a iteration, which aims to alleviate the **error accumulation**.

• **Tracks selection** finds a compact subset of tracks for the bundle adjustment, which makes the SfM become more **efficient** and **scalable**.
Thank you!