**IMP: Iterative Matching and Pose Estimation with Adaptive Pooling**

Fei Xue, Ignas Budvytis, Roberto Cipolla

---

**Problem**
- Relative pose estimation via keypoints matching

**Limitations of prior methods**
- Matching and pose estimation are independent
- Geometric information is ignored
- Graph-based matchers have good performance but suffer from high computational cost

**Motivation**
- Iterative matching and pose estimation
- Adaptively discarding keypoints without correspondences
- Robust pose-guided pooling

---

**Approach**
- Iterative matching and pose estimation
- Transformer-based recurrent module
- Pose-aware loss in the training process
- Pose-guided matching in the testing process

---

**Results**
- More inliers and accurate poses (matches / inliers / R error / t error)

---

**Adaptive pooling of keypoints (Efficient IMP - EIMP)**
- Pooling with attention scores and matching matrix to remove outliers
- Pose-guide pooling to avoid over pooling when matches are not good

---

**Percentage of pose errors within 5/10/20 deg on YFCC dataset**

<table>
<thead>
<tr>
<th></th>
<th>5 deg</th>
<th>10 deg</th>
<th>20 deg</th>
</tr>
</thead>
<tbody>
<tr>
<td>SuperGlue</td>
<td>32.6%</td>
<td>44.1%</td>
<td>55.6%</td>
</tr>
<tr>
<td>SGMNet</td>
<td>36.8%</td>
<td>49.7%</td>
<td>61.6%</td>
</tr>
<tr>
<td>CLNet</td>
<td>38.9%</td>
<td>51.6%</td>
<td>63.0%</td>
</tr>
<tr>
<td>IMP</td>
<td>42.8%</td>
<td>55.4%</td>
<td>67.1%</td>
</tr>
<tr>
<td>EIMP</td>
<td>45.2%</td>
<td>58.2%</td>
<td>70.4%</td>
</tr>
</tbody>
</table>

---

**Running time**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>SuperGlue</th>
<th>SGMNet</th>
<th>IMP</th>
<th>EIMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (ms)</td>
<td>43</td>
<td>35</td>
<td>37</td>
<td>35</td>
</tr>
</tbody>
</table>

---

**Uncertainty of pose shows the quality of matches**

---

**Attention scores show possibilities of keypoints being inliers**

---

**More robust (inliers) to viewpoint, illumination, and seasonal changes**

---

**More accurate poses and higher efficiency**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>NN</th>
<th>CLNet</th>
<th>SuperGlue</th>
<th>SGMNet</th>
<th>IMP</th>
<th>EIMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inliers</td>
<td>6.5</td>
<td>27.8</td>
<td>37.1</td>
<td>33.0</td>
<td>39.4</td>
<td>37.9</td>
</tr>
<tr>
<td>R error</td>
<td>15.4</td>
<td>46.4</td>
<td>57.2</td>
<td>53.0</td>
<td>59.4</td>
<td>57.9</td>
</tr>
<tr>
<td>t error</td>
<td>28.5</td>
<td>63.8</td>
<td>73.6</td>
<td>70.0</td>
<td>75.2</td>
<td>74.0</td>
</tr>
</tbody>
</table>

---

[2] SGMNet, Chen et al., ICCV 2021
[3] CLNet, Zhao et al., ICCV 2021
[5] Suerpoint, DeTone et al., CVPRW 2018