Place Recognition using Surface Entropy Features

Torsten Fiolka\textsuperscript{1}, Jörg Stückler\textsuperscript{2}, Dominik A. Klein\textsuperscript{3}, Dirk Schulz\textsuperscript{1}, and Sven Behnke\textsuperscript{2}

\textsuperscript{1}Fraunhofer Institute for Communication, Information Processing, and Ergonomics FKIE, Wachtberg, Germany

\textsuperscript{2}Autonomous Intelligent Systems Computer Science Institute VI University of Bonn, Germany

\textsuperscript{3}Intelligent Vision Systems, Computer Science Institute III, University of Bonn, Germany
Place Recognition using Interest Points

- Compress image content into set of prominent points
- Repeatable detection, distinctive description
- Invariance to image transformations
- Place recognition: classify set of point features
Prominent Points on Surfaces

- Why yet another 3D interest point?
  - Feature quality determines recognition rate
  - Entropy as a principled measure of prominence
  - 3D: Characterize local distribution of surface orientation changes
Measuring Interest on Surfaces using Entropy

- Distribution of surface curvature directions in local point neighborhood (scale)
- Entropy attains spatial maxima at points with prominent surface variation
Approximation of Curvature Distribution

- Curvature directions from pairs of surface normals
  \[ \frac{(\hat{n}(\mathcal{E}) \times \hat{n}_r(\vec{q}_k))}{\|\hat{n}(\mathcal{E}) \times \hat{n}_r(\vec{q}_k)\|_2} \]

- Orientation histogram with center bin for parallel normals
- Interpolate between center and directional bins
  \[ w_\times = (1 - \langle \hat{n}(\mathcal{E}), \hat{n}_r(\vec{q}_k) \rangle) \]
Efficient Implementation

• Octree
  – Multi-resolution
  – Regular sampling
  – Efficient aggregation of statistics within node volumes

• Efficient interest point detection
  – Parameters: Histogram and normal scale / sampling rate
  – Set octree maximum depth to match highest normal sampling rate
  – Set ratios between parameters to powers of 2 to hit octree resolutions
Occlusion Handling in Depth Images

- Fill-in points along view direction
- Virtual background detections
- Foreground detections
  - Fill-in points along view direction
Shape Descriptor

- Rotation-invariant description of surface curvature in local point neighborhood
- Histogram of surfel-pair relations between normal at interest point and neighboring normals
- Euclidean distance for descriptor comparison
Colorful Texture Descriptor

- Rotation-invariant description of
  - hue and saturation (HSL color space),
  - luminance contrasts

- Fast Earth Mover’s Distance with truncated ground distance to handle illumination changes
Place Recognition – Training

extract visual vocabulary

training images with features

living room

kitchen
Place Recognition – Training

- Extract visual vocabulary
- Generate word histograms

`training images with features`

`living room`

`kitchen`

`training images with word histograms`
Place Recognition – Recall

- **Category?**
  - Query Image

- **Query Word Histogram**

- **Generate Word Histogram**

- **Trained Visual Vocabulary**
Place Recognition – Recall

category?

query image

query word histogram

intersect histograms

living room

kitchen

... living room

kitchen

bathroom

20 best matches
Place Recognition – Recall

- Category?
- Query image
- Query word histogram

- Kitchen
- Match features

- 20 best matches
Evaluation – Setup

• Interest points
  – comparison with NARF
  – 4 datasets (3 objects, clutter) with ground truth poses
  – 3 scales (12, 24, 48cm)

• Place Recognition
  – comparison with NARF
  – 2 location datasets with 500 images from 3 rooms and 200 images from 5 rooms (kitchen, bathroom, living room, bedroom, corridor)
Evaluation – Repeatability

- simple: bidirectional consistent nearest neighbor matching
- unique: interest points also unique in scale range
- type of measure has little effect on SURE, few redundant interest points
• SURE without colorful texture descriptor (shape only) achieves similar or better scores than NARF
• Use of SURE’s colorful texture descriptor has a clear advantage
Evaluation – Run-Time

<table>
<thead>
<tr>
<th>dataset</th>
<th>SURE</th>
<th></th>
<th>NARF 160x120</th>
<th></th>
<th>NARF 320x240</th>
<th></th>
<th>NARF 640x480</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#features</td>
<td>run-time (sec)</td>
<td>#features</td>
<td>run-time (sec)</td>
<td>#features</td>
<td>run-time (sec)</td>
<td>#features</td>
<td>run-time (sec)</td>
</tr>
<tr>
<td>box</td>
<td>8.8</td>
<td>0.62</td>
<td>14.8</td>
<td>0.27</td>
<td>18.2</td>
<td>1.95</td>
<td>32.5</td>
<td>160.18</td>
</tr>
<tr>
<td>rocking horses</td>
<td>19.8</td>
<td>0.72</td>
<td>44.6</td>
<td>0.36</td>
<td>72.4</td>
<td>3.25</td>
<td>121.6</td>
<td>133.36</td>
</tr>
<tr>
<td>teddy</td>
<td>3.9</td>
<td>0.72</td>
<td>15.3</td>
<td>0.26</td>
<td>26.9</td>
<td>2.09</td>
<td>43.0</td>
<td>164.43</td>
</tr>
<tr>
<td>clutter</td>
<td>26.4</td>
<td>0.84</td>
<td>26.5</td>
<td>0.27</td>
<td>48.4</td>
<td>3.24</td>
<td>93.3</td>
<td>179.20</td>
</tr>
</tbody>
</table>

- SURE 640x480 requires computation time between NARF 160x120 and NARF 320x240
### Evaluation – Place Recognition

<table>
<thead>
<tr>
<th>dataset</th>
<th>method</th>
<th>correct</th>
<th>avg. run-time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>NARF 160x120</strong></td>
<td>53.3%</td>
<td>0.3 sec</td>
</tr>
<tr>
<td>3 rooms</td>
<td><strong>NARF 320x240</strong></td>
<td>55.1%</td>
<td>4 sec</td>
</tr>
<tr>
<td></td>
<td><strong>NARF 640x480</strong></td>
<td>28.1%</td>
<td>4 min</td>
</tr>
<tr>
<td></td>
<td><strong>SURE shape only</strong></td>
<td>56.2%</td>
<td>1.4 sec</td>
</tr>
<tr>
<td></td>
<td><strong>SURE color+shape</strong></td>
<td>91.5%</td>
<td>1.6 sec</td>
</tr>
<tr>
<td>5 rooms</td>
<td><strong>NARF 160x120</strong></td>
<td>37.5%</td>
<td>0.3 sec</td>
</tr>
<tr>
<td></td>
<td><strong>NARF 320x240</strong></td>
<td>39.1%</td>
<td>4 sec</td>
</tr>
<tr>
<td></td>
<td><strong>NARF 640x480</strong></td>
<td>24.0%</td>
<td>4 min</td>
</tr>
<tr>
<td></td>
<td><strong>SURE shape only</strong></td>
<td>43.4%</td>
<td>1.4 sec</td>
</tr>
<tr>
<td></td>
<td><strong>SURE color+shape</strong></td>
<td>88.3%</td>
<td>1.9 sec</td>
</tr>
<tr>
<td></td>
<td><strong>SURE, direct comparison</strong></td>
<td>91.1%</td>
<td>4.8 sec</td>
</tr>
</tbody>
</table>

- Shape alone not sufficient
- Inclusion of colorful texture descriptors achieves high recognition rates
- Direct feature comparison better but computationally more demanding
Summary and Future Work

- SURE is a novel pairing of interest point detector and descriptor for 3D point clouds and depth images based on surface entropy
- Promising results w.r.t. detector repeatability and matching by descriptor
- High recognition rates with our place recognition approach

Future Directions:

- Complement detector with texture
- Automatic scale selection
- Further applications, e.g., loop-closing
Thank you!
Detected Interest Points