A Performance Evaluation of 3D Keypoint Detectors

Samuele Salti, Federico Tombari, Luigi Di Stefano

CVLab – University of Bologna, Italy
Lots of 3D applications rely on **3D keypoints** (or features)

- **Distinctive**, i.e. suitable for effective description and matching

- **Repeatable** with respect to point-of-view variations

3DIMPVT - A Performance Evaluation of 3D Keypoint Detectors

18/05/2011
3D keypoints for object recognition

- To be useful in an object recognition scenario 3D keypoints must be also repeatable in presence of occlusions, clutter, noise
- Several techniques proposed recently but we don’t know which one performs better

3DIMPVT - A Performance Evaluation of 3D Keypoint Detectors 18/05/2011
Several recent proposals, two main categories

- **Fixed scale detectors**
  - Local Surface Patches (LSP) [Chen & Bhanu, PRL07]
  - Intrinsic Shape Signatures (ISS) [Zhong, 3DRR09]
  - KeyPoint Quality (KPQ) [Mian et al., IJCV10]

- **Scale-invariant detectors**
  - Several proposal working only on range maps (e.g. [Novatnack & Nishino, ECCV08], [Akagunduz & Ulusoy, ICCV07])
  - MeshDoG [Zaharescu et al., CVPR09]
  - KPQ Scale Invariant (KPQ-SI) [Mian et al., IJCV10]
  - Laplace-Beltrami Scale-Space (LBSS) [Unnikrishnan & Hebert, S3D08]
  - Heat Kernel Signature (HKS) [Sun et al., SGP09]
  - 3D-SURF [Knopp et al., ECCV10]
Methodology

- We focused on **repeatability**
  - **Absolute repeatability**: the number of repeatable keypoints of a model in a scene
  - **Relative repeatability**: the percentage of repeatable keypoints out of the number of model keypoints.

- A keypoint is said to be repeatable if \[ R_{ms} k^i_m + t_{ms} - k^j_s \] < \( \varepsilon \)

Model keypoint rotated and translated in the scene
Scene keypoint
Repeatability threshold
For scale invariant detectors we also consider scale repeatability

\[ R_{ms} k^i_m + t_{ms} \]

\[ k^j_s \]

\[ R_{ms} k^i_m + t_{ms} \]
Datasets

- 2 “Synthetic” datasets created with models from the Stanford Repository
  - 1 for Retrieval
  - 1 for Object Recognition
  - Synthetic Gaussian noise

- the laser scanner dataset [Mian et al., IJCV10]
  - Detailed shapes
  - Full 3D models, 2.5D scenes
  - Point density variations

- the SpaceTime stereo dataset [Tombari et al., ECCV10]
  - Smoother, noisier data
  - 2.5D models and scenes
  - Same point density
- LSP performs poorly in presence of noise (based on curvature)
- EVD based fixed-scale detectors (i.e. ISS and KPQ) suffer 2.5D views and occlusions
- Presence of clutter favors smaller supports
- KPQ is largely unaffected by noise (smooth surface fitting step)
- Best performer
  - Retrieval: ISS
  - ObjRec: KPQ
LBSS exhibits unsatisfactory repeatability

KPQ-SI is greatly more robust to noise than MeshDoG (smooth surface fitting step vs use of curvatures)

KPQ and KPQ-SI have similar performance
- LBSS shows superior scale repeatability but poor spatial repeatability.
- MeshDoG suffers point density variations (from Laser Scanner dataset results)
## Timing (in seconds)

<table>
<thead>
<tr>
<th></th>
<th>Retrieval</th>
<th>Random Views</th>
<th>Laser Scanner</th>
<th>Space Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSP</td>
<td>56 ~ 65</td>
<td>31 ~ 100</td>
<td>65 ~ 76</td>
<td>74 ~ 92</td>
</tr>
<tr>
<td>ISS</td>
<td>2 ~ 10</td>
<td>2 ~ 7</td>
<td>5 ~ 13</td>
<td>6 ~ 18</td>
</tr>
<tr>
<td>KPQ*</td>
<td>266 ~ 493</td>
<td>413 ~ 662</td>
<td>799 ~ 1109</td>
<td>544 ~ 1222</td>
</tr>
<tr>
<td>LBSS</td>
<td>1585</td>
<td>461</td>
<td>1148</td>
<td>1397</td>
</tr>
<tr>
<td>MeshDoG</td>
<td>198</td>
<td>185</td>
<td>425</td>
<td>469</td>
</tr>
<tr>
<td>KPQ-SI*</td>
<td>303</td>
<td>364</td>
<td>634</td>
<td>767</td>
</tr>
</tbody>
</table>

- Average detection time / scene (in seconds)
- ISS is one or more orders of magnitude faster than the other detectors