A Benchmark Dataset for 6DOF Object Pose Tracking

Po-Chen Wu1, Yueh-Ying Lee1, Hung-Yu Tseng1, Hsu-an Ho1, Ming-Hsuan Yang2, Shao-Yi Chien1
1Media IC & System Lab, National Taiwan University, 2Vision and Learning Lab, UC Merced

Abstract

Accurately tracking the six degree-of-freedom pose of an object in real scenes is an important task in computer vision and augmented reality with numerous applications. Although a variety of algorithms for this task have been proposed, it remains difficult to evaluate existing methods in the literature as oftentimes different sequences are used and no large benchmark datasets close to real-world scenarios are available. In this paper, we present a large object pose tracking (OPT) benchmark dataset consisting of RGB-D video sequences of 2D and 3D targets with ground-truth information. The videos are recorded under various lighting conditions, different motion patterns and speeds with the help of a programmable robotic arm. We present extensive quantitative evaluation results of the state-of-the-art methods on this benchmark dataset and discuss the potential research directions in this field.

Introduction

Using a programmable robotic arm (as shown in the figure above), we can record images under different motion patterns and different speed.

- The proposed object pose dataset is also the only one where color and depth information.

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>Device</th>
<th>Mechanism</th>
<th>Point Establishment Time (s)</th>
<th>2D Target</th>
<th>3D Target</th>
<th>Motion Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lietherknecht</td>
<td>3D-2D</td>
<td>Headled</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>George et al.</td>
<td>2D-2D</td>
<td>Headled</td>
<td>6</td>
<td>16</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Hinterstoisser</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Kouskouridas</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Krull et al.</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Rennie et al.</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Bekaert et al.</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Druon et al.</td>
<td>2D-2D</td>
<td>Headled</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>OPT</td>
</tr>
<tr>
<td>Proposed</td>
<td>2D-2D</td>
<td>Programmable</td>
<td>6</td>
<td>6</td>
<td>23</td>
<td>OPT</td>
</tr>
</tbody>
</table>

- 2D objects with low (wing, duck), normal (city, beach), and rich (maple, firework) texture.

- 3D objects with simple (soda, chest), normal (ironman, house), and complex (bike, jet) geometry.

OPT Dataset

- Motion patterns
  - Translation. An object moves along a circle parallel to the camera sensor plane with motion blur in all directions.
  - Zoom. An object moves forward first and then backward.
  - In-plane Rotation. An object rotates along an axis perpendicular to the camera sensor plane.
  - Out-of-plane Rotation. An object rotates along an axis parallel to the camera sensor plane.
  - Flashing Light. The light source is turned on and off repeatedly, and the object moves slightly.
  - Moving Light. The light source moves and results in illumination variations while the object moves slightly.
  - Free Motion. An object moves in arbitrary directions.

- Ground-truth object pose annotation

- Provide two different resolutions: 1920 x 1080 and :512 x 424

Evaluation Results

- Overall performance evaluation. The AUC score for each approach is shown in the legend.

References


Contact

Po-Chen Wu
Taiwan National University
Email: pcwu@media.ee.ntu.edu.tw
Website: http://media.ee.ntu.edu.tw/research/OPT/