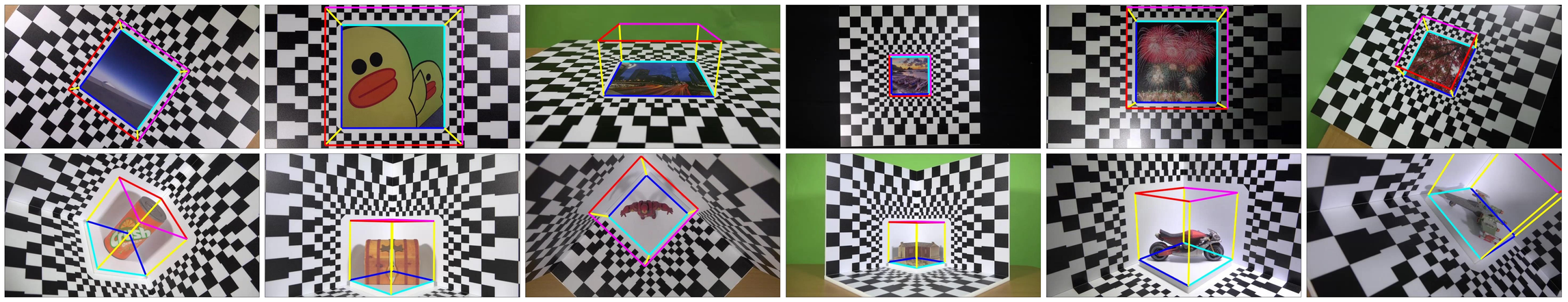
A Benchmark Dataset for 6DoF Object Pose Tracking

Po-Chen Wu[1], Yueh-Ying Lee[1], Hung-Yu Tseng[1], Hsuan-I Ho[1], Ming-Hsuan Yang[2], Shao-Yi Chien[1]

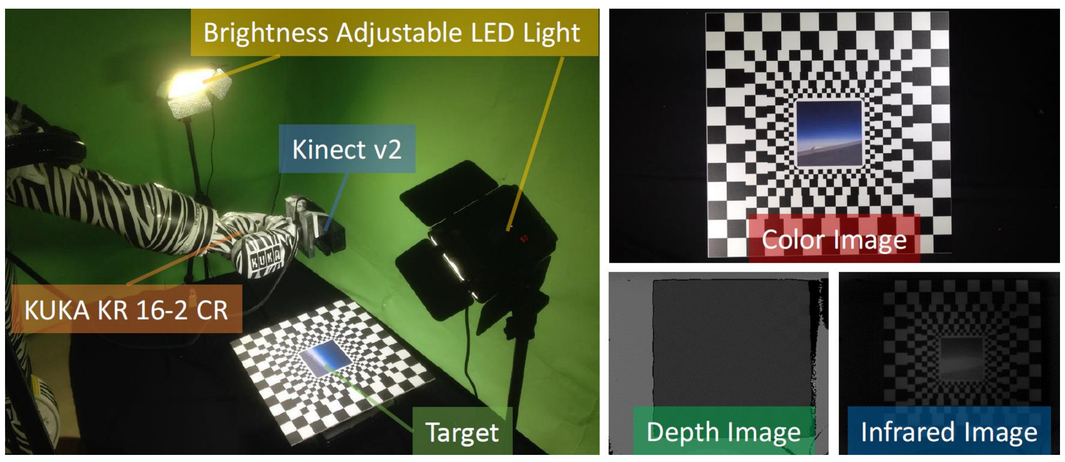
[1]Media IC & System Lab, National Taiwan University, [2]Vision and Learning Lab, UC Merced



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 image sequences are recorded by a **Microsoft Kinect v2 sensor**.

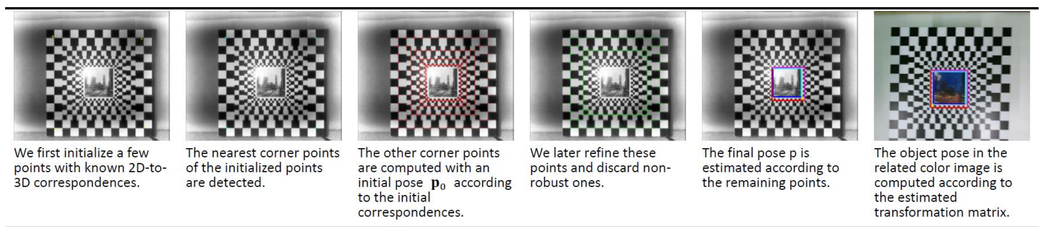
Benchmark	Device	Mechanism	Pose Establishment	Video Clips	# 2D Targets	# 3D Targets	# Motion Patterns	# Frames
Lieberknecht [1]	Marlin F-080C	Handheld	Marker-based	Yes	8	-	5	48,000
Gauglitz [2]	Fire-i	Manually Operated Contraption	Direct Alignment	Yes	6	-	16	6,889
Hinterstoisser [3]	Kinect v1	Handheld	Marker-based	No	-	15	-	18,000
Tejani [4]	Kinect v1	Handheld	Marker-based	No	-	3	-	5,229
Brachmann [5]	Kinect v1	Handheld	Marker-based	No	-	20	3	10,000
Rennie [6]	Kinect v1	Robotic Arm	Manual	No	-	24	-	10,368
Krull [7]]	Kinect v1	Handheld	ICP	Yes	-	3	-	1,100
Choi [8]	Synthetic	-	Synthetic	Yes	-	4	-	4,000
Proposed	**Kinect v2**	**Programmable Robotic Arm**	**Checkerboard-based**	Yes	6	6	23	**100,956**

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

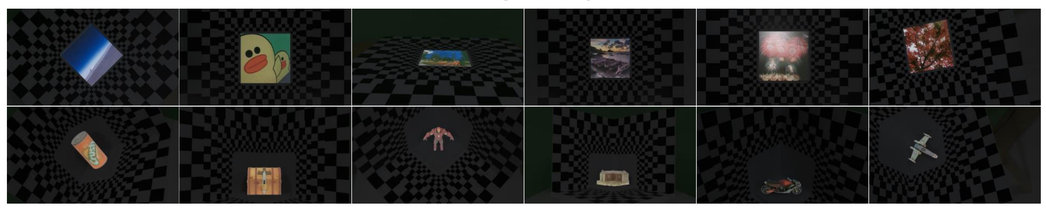
- 3D objects with simple (**soda**, **chest**), normal (**ironman**, **house**), and complex (**bike**, **jet**) geometry.

OPT Dataset

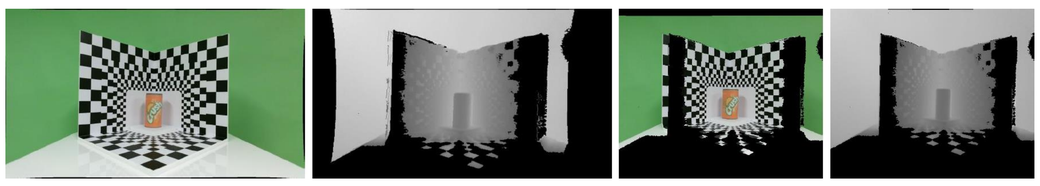
- Motion patterns
  - ① **Translation**. An object moves along a circle parallel to the camera sensor plane with motion blur in all directions.
  - ② **Zoom**. An object moves forward first and then backward.
  - ③ **In-plane Rotation**. An object rotates along an axis perpendicular to the camera sensor plane.
  - ④ **Out-of-plane Rotation**. An object rotates along an axis parallel to the camera sensor plane.
  - ⑤ **Flashing Light**. The light source is turned on and off repeatedly, and the object moves slightly.
  - ⑥ **Moving Light**. The light source moves and results in illumination variations while the object moves slightly.
  - ⑦ **Free Motion**. An object moves in arbitrary directions.
- Ground-truth object pose annotation



- Camera frames for 2D and 3D target objects blended with **masks**.

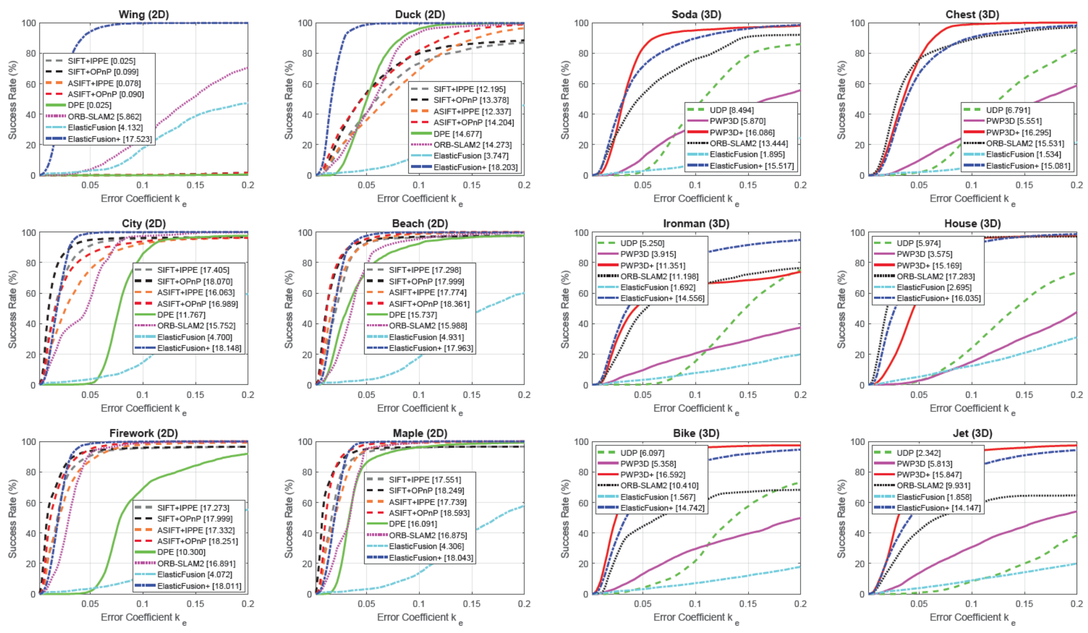


- Provide two different resolutions: $\mathbf{1920 \times 1080}$ and : $\mathbf{512 \times 424}$



Evaluation Results

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



References

1. S. Lieberknecht, S. Benhimane, P. Meier, and N. Navab. A Dataset and Evaluation Methodology for Template-based Tracking. In ISMAR, 2009.
2. S. Gauglitz, T. H¨ollerer, and M. Turk. Evaluation of Interest Point Detectors and Feature Descriptors for Visual Tracking. IJCV, 94(3):335– 360, 2011.
3. S. Hinterstoisser, V. Lepetit, S. Ilic, S. Holzer, G. Bradski, K. Konolige, and N. Navab. Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes. In ACCV, 2012.
4. A. Tejani, D. Tang, R. Kouskouridas, and T.-K. Kim. Latent-Class Hough Forests for Object Detection and Pose Estimation. In ECCV, 2014.
5. E. Brachmann, A. Krull, F. Michel, S. Gumhold, J. Shotton, and C. Rother. Learning 6D Object Pose Estimation Using 3D Object Coordinates. In ECCV, 2014.
6. C. Rennie, R. Shome, K. E. Bekris, and A. F. De Souza. A Dataset for Improved RGBD-based Object Detection and Pose Estimation for Warehouse Pick-and-Place. RAL, 1(2):1179–1185, 2016.
7. A. Krull, F. Michel, E. Brachmann, S. Gumhold, S. Ihrke, and C. Rother. 6-DOF Model Based Tracking via Object Coordinate Regression. In ACCV, 2014.
8. C. Choi and H. I. Christensen. RGB-D Object Tracking: A Particle Filter Approach on GPU. In IROS, 2013.

Contact

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