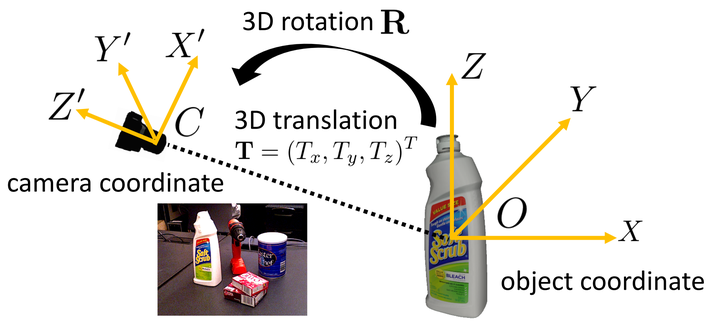
DeepIM: Deep Iterative Matching for 6D Pose Estimation

Yi Li[1], Gu Wang[1], Xiangyang Ji[1], Yu Xiang[2], Dieter Fox[2,3]

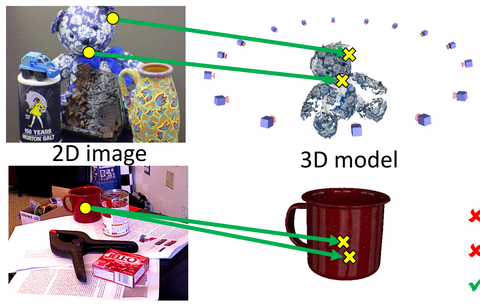
[1]Tsinghua University, [2]NVIDIA Research, [3]University of Washington

❑6D Object Pose Estimation



❑Related Work

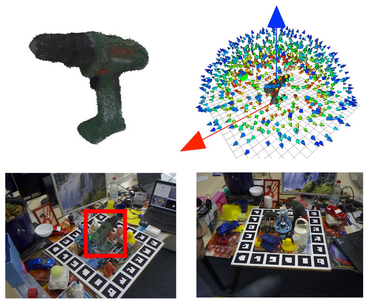
- Feature Matching



- Lowe, ICCV, 1999
- Rothganger et al., IJCV, 2006
- Savarese & Fei-Fei, ICCV, 2007
- Collet et al., IJRR, 2011
- Brachmann et al., ECCV, 2014
- Krull et al., ICCV, 2015
- Kehl et al., ECCV, 2016
- Michel et al., CVPR, 2017
- Pavlakos et al., ICRA, 2017
- Rad & Lepetit, ICCV, 2017

✘ Texture-less objects
✘ Symmetric objects
✓ Occlusions

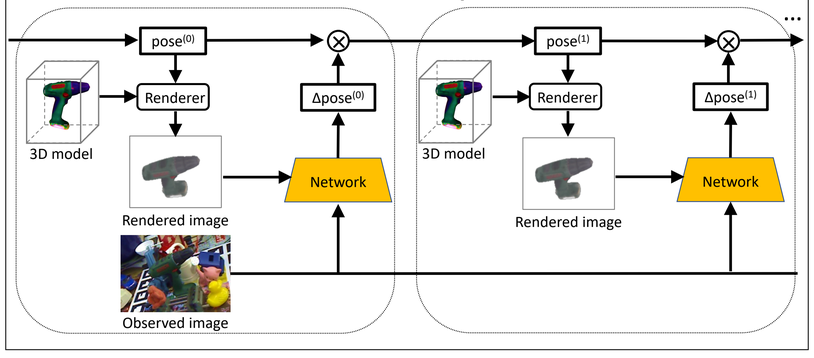
- Template Matching



- Thomas et al., CVPR, 2006
- Ozuysal et al., CVPR, 2009
- Gu & Ren, ECCV, 2010
- Hinterstoisser et al., ACCV, 2012
- Xiang & Savarese, CVPR, 2012
- Pepik et al., CVPR, 2012
- Su et al., ICCV, 2015
- Cao et al., ICRA, 2016
- Tekin et al, CVPR, 2018

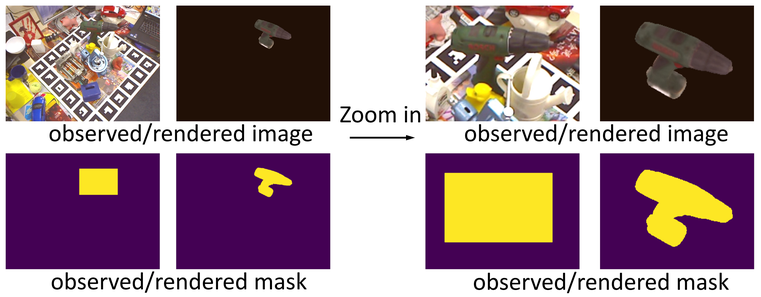
✓ Texture-less objects
✓ Symmetric objects
✘ Occlusions

❑Our Contribution: DeepIM

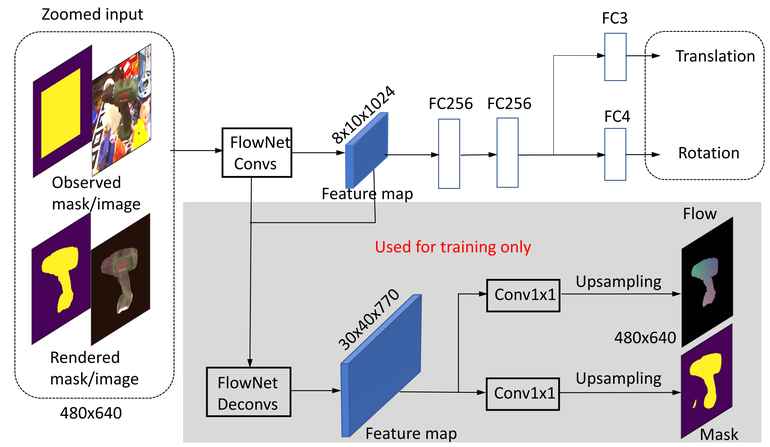


❑Deep Iterative Matching

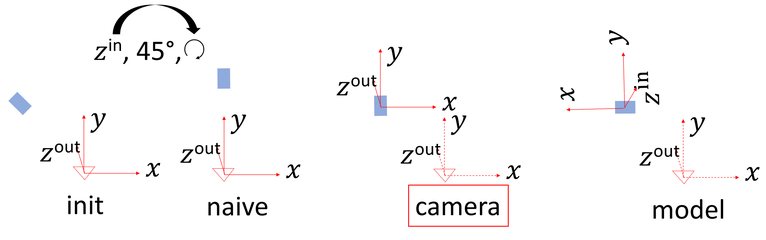
- High-resolution Zoom in



- Network Architecture



- Untangled Transformation Representation



❑Acknowledgement

This work was funded in part by a Siemens grant. We would also like to thank NVIDIA for generously providing DGX used for this research via the NVIDIA Robotics Lab and the UW NVIDIA AI Lab (NVAIL).

❑Experiments on the LINEMOD Dataset

- Ablation study on iterative training and testing

Train iter	init	1			2			4		
Test iter		1	2	4	1	2	4	1	2	4
5cm 5°	19.4	57.4	58.8	54.6	76.3	86.2	86.7	70.2	83.7	85.2
6D Pose	62.7	77.9	79.0	76.1	83.1	88.7	89.1	80.9	87.6	88.6
Proj. 2D	70.2	92.4	92.6	89.7	96.1	97.8	97.6	94.6	97.4	97.5

- Ablation study on two different initial poses

Method	PoseCNN	PoseCNN+Ours	Faster R-CNN	Faster R-CNN+Ours
5cm 5°	19.4	85.2	11.9	83.4
6D Pose	62.7	88.6	33.1	86.9
Proj. 2D	70.2	97.5	20.9	95.7

- Comparison with state-of-the-art methods

Method	[1]	[2]	[3]	[4]	[5]	Ours
5cm 5°	40.6	69.0	-	-	19.4	85.2
6D Pose	50.2	62.7	79.0	55.95	62.7	88.6
Proj. 2D	73.7	89.3	-	90.37	70.2	97.5

[1] Brachmann et al., Uncertainty-driven 6D pose estimation of objects and scenes from a single RGB image. In CVPR, 2016.
[2] Rad et al., BB8: A scalable, accurate, robust to partial occlusion method for predicting the 3D poses of challenging objects without using depth. In ICCV, 2017.
[3] Kehl et al., SSD-6D: Making rgbbased 3D detection and 6D pose estimation great again. In CVPR, 2017.
[4] B. Tekin, S. N Sinha, and P. Fua. Real-time seamless single shot 6d object pose prediction. In CVPR, 2018.
[5] Xiang et al., PoseCNN: A convolutional neural network for 6D object pose estimation in cluttered scenes. In RSS, 2018.