RoboCoop

V-A SK-AT project RoboCoop with the number V212

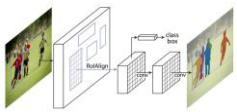
Object detection and Pose estimation for grasping

Markus Vinzce

Overview

Object Detection - Mask-RCNN

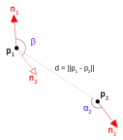
- Region Proposal Network (RPN) using a CNN backbone
- Classification and mask prediction for most likely Region of Interest (RoI)
- RoIAlign using bilinear interpolation



K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 2980-2988

Object Pose Estimation - PPF

- Point Pair Features:
  $\mathbf{F}(\mathbf{m}_1, \mathbf{m}_2, \mathbf{n}_1, \mathbf{n}_2) = [\|\mathbf{d}\|_2, \angle(\mathbf{n}_1, \mathbf{d}), \angle(\mathbf{n}_2, \mathbf{d}), \angle(\mathbf{n}_1, \mathbf{n}_2)]$
- Lookup table from discretized PPF to rotation around normal of the first point to a canonical reference frame
- Hough voting strategy for final pose
- Various subsampling approaches



Alexandrov S.V., T. Patten, M. Vincze "Leveraging Symmetries to Improve Object Detection and Pose Estimation from Range Data", in Computer Vision Systems. ICVS 2019. Lecture Notes in Computer Science, vol 11754. Springer, Cham.

Object Pose Estimation - PPF

- Originally for detection and pose estimation
- Does not scale well with more models and bigger scenes

  → Used in combination with Mask-RCNN

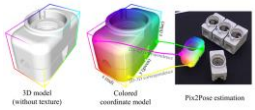
	Date (UTC)	Method	Test image	AR_{Core}	AR_{LM-O}	AR_{T-LESS}	AR_{TUD-L}	AR_{IC-MV}	AR_{ITODD}	AR_{HB}	AR_{YCB-V}	Time (s)
1	2020-08-19	CosyPose-ECCV20-SYNT+REAL-TVIEW-ICP	RGB-D	0.698	0.714	0.701	0.959	0.647	0.313	0.712	0.861	13.743
2	2020-08-19	Koenig-Hybrid-DL-PointPairs	RGB-D	0.659	0.651	0.655	0.920	0.430	0.483	0.651	0.701	0.635
3	2020-08-18	CosyPose-ECCV20-SYNT+REAL-TVIEW	RGB	0.637	0.633	0.728	0.823	0.583	0.216	0.656	0.821	0.449
4	2020-08-17	Pix2Pose-BOP20_wICP-ICCV19	RGB-D	0.591	0.588	0.512	0.820	0.390	0.351	0.695	0.780	4.844

BOP Challenge 2019/2020 Leaderboard (https://bop.felk.cvut.cz/leaderboards/) 02/2021

Alexandrov S.V., T. Patten, M. Vincze "Leveraging Symmetries to Improve Object Detection and Pose Estimation from Range Data", in Computer Vision Systems. ICVS 2019. Lecture Notes in Computer Science, vol 11754. Springer, Cham.

Object Pose Estimation - Pix2Pose

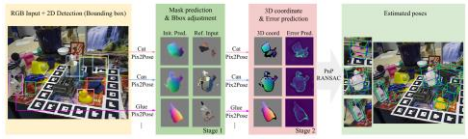
- Learn from untextured models
- Predicts coordinates image
  - Each pixel encodes a position in the object reference frame
  - Final pose is estimated using PnP with RANSAC



K. Park, T. Patten, M. Vincze, "Pix2Pose: Pixel-Wise Coordinate Regression of Objects for 6D Pose Estimation"; Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 7668-7677

Object Pose Estimation - Pix2Pose

- 2-stage estimation
  - Mask prediction and Bbox adjustment
  - 3D Coordinate & Error prediction



K. Park, T. Patten, M. Vincze, "Pix2Pose: Pixel-Wise Coordinate Regression of Objects for 6D Pose Estimation"; Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 7668-7677

Pose Estimation under Domain Shift - PyraPose

Training on synthetic RGB images:

- To generalize to novel domains
- To fully automate training

Testing on real-world images:

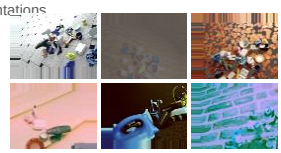
- Detection, pose and coarse mask estimation
- Scalable (multi-instance per forward pass)
- Good occlusion handling through hypotheses sampling scheme
- Fast! ~27 fps on Nvidia Titan V

S. Thalhammer, M. Leitner, . Patten and M. Vincze, "PyraPose: Feature Pyramids for Fast and Accurate Object Pose Estimation under Domain Shift", Under submission at IEEE Robotics and Automation Letters.

Pose Estimation under Domain Shift - PyraPose

Domain adaptation via:

- Coalesced multi-scale features
- Image augmentations



S. Thalhammer, M. Leitner, . Patten and M. Vincze, "PyraPose: Feature Pyramids for Fast and Accurate Object Pose Estimation under Domain Shift", Under submission at IEEE Robotics and Automation Letters.

Pose Verification and Refinement

Alignment

Rendering-based verification score for varying object pose (warmer color for higher fit).

Plausibility

Runs a simulation (showing rotation only) from initial pose (blue) and a registration step (green).

D. Bauer, T. Patten and M. Vincze, "VeREFINE: Integrating Object Pose Verification With Physics-Guided Iterative Refinement," in IEEE Robotics and Automation Letters, vol. 5, no. 3, pp. 4289-4296, July 2020

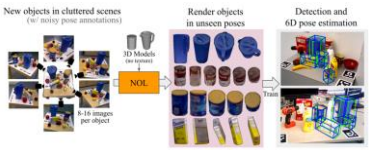
How to obtain Object Models ?

- Recognition Tracking & Modeling toolbox
  - RGB-D camera and turntable setup
  - Textured support surface for tracking
  - Generates point cloud models

K. Park, T. Patten and M. Vincze, "Neural Object Learning for 6D Pose Estimation Using a Few Cluttered Images", in European Conference on Computer Vision (ECCV), August 2020

How to obtain Object Models ?

- Neural Object Learning to render under unseen views



K. Park, T. Patten and M. Vincze, "Neural Object Learning for 6D Pose Estimation Using a Few Cluttered Images", in European Conference on Computer Vision (ECCV), August 2020