

Progress Report

Bardia Mojra

August 28, 2020

Robotic Vision Lab
University of Texas at Arlington

1 Progress

Following items are listed in order of priority:

- OCRTOC: Chris and I are working on implementing instance segmentation and inverse kinematics on UR5 simulation. PoseCNN, [1], uses a convolutional neural network architecture to estimate 6D pose of every object using instance segmentation with RGB-depth camera. First it uses semantic labeling to detect and classify each object within the RGB image, then it estimates the 3D translation by localizing 2D object center in the image and then estimates its distance to the camera. That is done at pixel level and then they use Hough voting method to estimate the center of the object, in 2D. This is a novel approach, but perhaps it could be improved if we first estimate the 3D contour of the object then estimate the center point of the object directly in 3D. Moreover, they use a regression model to estimate each object's 3D rotation. To train their 3D rotation regression model, authors propose two L2 loss functions, PoseLoss (PLoss) and ShapeMatch-Loss (SLoss) where SLoss does not require the specification of object symmetries. At last, they use depth image/video for improving 6D pose annotation and training their model in a global optimization step. Similar learning models can be developed using Deep Reinforcement Learning since ground truth is available via depth camera.
- DenseFusion: I will read and dissect DenseFusion paper, [2], some time this weekends.
- UR5: I began looking at installing Robotiq parallel gripper, it should be fairly straightforward and simple. We still need to secure UR5e wires to the floor and table.
- TensorFlow: I went through chapter 1 TensorFlow implementation provided in [3]. I need to begin using Anaconda to manage my Python workspaces. Nolan and Zongyao will assist me with getting started on that.
- Fellowship: I still need to work on my applications, I will write a new draft by next week. I have been avoiding this for far too long.
- Machine Learning: I did not work on ML course this week.

- (On pause) Nolan and I have continued to work on TISR paper, I started a new document where I summarize background information. A copy of this document has been uploaded to GitHub.

2 Plans

Following items are listed in order of priority:

- Continue to work on OCRTOC with Chris and begin working with instance segmentation ROS node.
- Go through ROS Industrial tutorials and documentation.
- Resume Robotic Perception course as soon as possible.
- (On pause) Begin working on quaternions tutorial.
- (On pause) Need to read [4], [5], [6], and [7]; these papers seem fundamental to understanding the overall picture.
- (On pause) Read Digital Image Processing by Gonzalez and Woods.
- (On pause) I still need to dissect [8], [9], [10], [11], [12], and [13].
- (On pause) I still need to dissect [14].

References

- [1] V. N. D. F. Yu Xiang, Tanner Schmidt, “Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes,” *CoRR*, 2017.
- [2] Y. Z. R. M.-M. C. L. Chen Wang, Danfei Xu and S. S. Li Fei-Fei, “Densefusion: 6d object pose estimation by iterative dense fusion,” *CoRR*, 2019.
- [3] E. A. Benjamin Planche, *Hands-On Computer Vision with TensorFlow 2*. Packt Publishing, 5 2019.
- [4] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *CoRR*, vol. abs/1501.00092, 2015.
- [5] D. Liu, Z. Wang, N. M. Nasrabadi, and T. S. Huang, “Learning a mixture of deep networks for single image super-resolution,” *CoRR*, vol. abs/1701.00823, 2017.
- [6] K. Doherty, D. Fourie, and J. Leonard, “Multimodal semantic slam with probabilistic data association,” in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 2419–2425, May 2019.
- [7] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, “Residual attention network for image classification,” *CoRR*, vol. abs/1704.06904, 2017.
- [8] A. Kirillov, K. He, R. B. Girshick, C. Rother, and P. Dollár, “Panoptic segmentation,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9396–9405, 2019.
- [9] C. Forster, M. Pizzoli, and D. Scaramuzza, “Svo: Fast semi-direct monocular visual odometry,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 15–22, May 2014.
- [10] B. Horn, “Closed-form solution of absolute orientation using unit quaternions,” *Journal of the Optical Society A*, vol. 4, pp. 629–642, 04 1987.
- [11] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, “Indoor segmentation and support inference from rgb-d images,” pp. 746–760, 10 2012.

- [12] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, “Dynamic graph CNN for learning on point clouds,” *CoRR*, vol. abs/1801.07829, 2018.
- [13] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–1, 2018.
- [14] W. K. Zinsser, *On writing well : the classic guide to writing nonfiction*. New York: HarperCollins, 30th anniversary ed., 7th ed., rev. and updated ed., 2006.