Progress Report

Bardia Mojra

September 5, 2020

Robotic Vision Lab
University of Texas at Arlington

1 Progress

Following items are listed in order of priority:

- OCRTOC: This week Jerry and I worked on implementing a pose estimation solution for OCRTOC competition, we looked at PoseCNN [1], DenseFusion [2], DOPE [3]. As discussed, DOPE is the most promising solution and Jerry is leading that effort. I started back on ROS industrial tutorials, it introduces a lot of features that I still need to get comfortable with. I also believe we might be able to draw inspiration from the latest kitting implementations found on GitHub.

- DenseFusion [2]: DenseFusion is a novel approach for estimating 6D position for a given RGB-Depth input. First, it generates object segmentation masks and bounding boxes from RGB images. Moreover, it generate point cloud based on processed masked depth images. Then, RGB image crops are used to generate color embeddings using a CNN. In parallel, masked point clouds are used to generate geometry embeddings. The two embeddings are then pixel-wise dense fused together as one tensor for each training sample. Then a MLP [4] is trained PointNet with Average Pooling, instead of Max Pooling, to encode the stochastic features or "information about the vicinity of each point and of the point cloud as a whole". Then with global features it is stacked onto the tensor so that pixel-wise features are preserved. Next, a pose predictor proposed to train a model to estimate rotation R, translation t and confidence c. The paper points out its short comings for symmetric objects and introduces another equation and cites ambiguous learning objectives as the root cause. Perhaps, use of quaternions can solve this learning ambiguity, use of mathematical tools and concepts that can describe projecting geometry accurately is quite useful when processing raw data from a RGB-D sensor.

- AI Class Project Idea: An auto trader for detecting general market corrections while it is happening and make trades to make maximum gain of sell-off and rebound. Each correction on average last around five to ten days and could happen once a month for periods of time. It is a sell-off over too much speculations in an over-heated market, and it is a low hanging fruit! The signs are remarkably obvious in my opinion. You could wake up one day and see the market and stocks all red for no apparent reason. There is nothing on the news to indicate a new economic down turn on the horizon but few hours

later, financial news begins to be filed with words such as "sell-off", "correction", "market-down" and etc. There is so much data available online both raw and derivative, all while the pattern is clear, it drops from the peak or all-time-high and corrects itself with a range of 10 to 20 percent.

- UR5e: I began going through ROS industrial tutorials and UR5e ROS package.
- Quaternion: Will try to send out short reading materials and summary on quaternions on weekly basis.
- TensorFlow: Need to keep referencing to [5] and do few projects.
- Fellowship:
- Machine Learning:

2 Plans

Following items are listed in order of priority:

- Continue to work on OCRTOC.
- Continue going through ROS Industrial tutorials and documentation.
- (On pause) Resume Robotic Perception course as soon as possible.
- (On pause) Read Digital Image Processing by Gonzalez and Woods.

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] NVlabs, "Deep object pose," Sep 2020.

[4] H. Bourlard and C. J. Wellekens, "Links between markov models and multilayer perceptrons," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 12, pp. 1167–1178, 1990.

[5] E. A. Benjamin Planche, *Hands-on with Computer Vision with Tensorflow 2*. 2017.