Progress Report

Bardia Mojra

November 28, 2020

Robotic Vision Lab
University of Texas at Arlington

1 To Do

- Implement and play with PoseCNN and DOPE.
- Generate new data set using UE4.
- Reconstruct a pose estimation model to familiarize myself and then start modifying it.
- Read more papers on pose estimation.
- Look into transfer learning. Read [1].
- Look into domain randomization and adaptation.
- Read [2].
- Learn to use UE4.

2 Progress

Following items are listed in order of priority:

- Pose Estimation: This week, I pulled and tested Docker container Jerry created based on OCRTOC container. Moreover, I finished a tutorial on UE4, [5], this is needed for using Nvidia's Dataset Synthasizer. Next, I will generate some data and train Dope and PoseCNN networks with it With some basic understanding of what Lie Algebra is about, I still need to finish dissecting [1] for Pose Estimation work.
- Lie Algebra (no new development): It is a vector space V over a base field F along with bracket operation that satisfies bilinearity, anti-symmetry, and the Jacobian Identity conditions. Considering the fact that robotic vision applications are process heavy, I find it immensely important to be familiar with mathematical tools (such as Dynamic Primitive of Motor Control [3]) that enables us to encode important information into our models, whether it is actuator manipulation or dynamic scene understanding. [4] provides a good starting point on Lie Algebra. I am putting this new theory lead on pause till after Pose Estimation paper for this semester.

- Bayesian Scanning: The goal is to scan a scene or an object using a monocular RGBD sensor with its pose changing. The novelty behind it would be to use Bayesian principles to build a system that can produce increased certainty using motion rather than losing accuracy due to compounding noise. Dealing with accumulated or compounded noise is an important topic of discussion and I will expand on that later. In short, I believe some kind of recalibration-on-the-fly is going to be needed. This scanning system work by update beliefs for entire domain and by tracking each voxel cube in the scene. Bayes' theory relies on the fact that the whole truth equates to one and with it we examine degrees of certainty over equal fractions of the whole domain. Now, if we decide to examine a scene we can assume there is definitely something to be scanned in the domain or the lack thereof and at each step we update the belief for each voxel based on observations and emissions.

  I need to learn ORB-SLAM implementation, develop a new dynamic data structure for dynamic-resolution voxel space, develop 3D ORB feature detector and descriptor for 3D point-cloud KNN topology or for the contour. I need to familiarize myself much more with the mentioned theories and code based before I can fully mature the idea. I am sure there are technical difficulties that are still unknown-unknown to me.

- OCRTOC: I successfully setup Unreal Engine 4, [5]. Next, we need to look into Domain Randomization and other techniques for developing a data with diverse features and patterns.

- TensorFlow [6]: I am still working through chapter 2.

- MoreFusion [7]: Still need to write a literature review on this.

- Reading list: [8] and [9].

- Project Alpe with Nolan: On pause for right now.

- Quaternions:

- UR5e: I can work on putting together something presentable with UR5e but that might take some time.

- Fellowship:

- System Identification Presentation:

3 Plans

Following items are listed in order of priority:

- (On pause) Continue with 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] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, "Domain randomization for transferring deep neural networks from simulation to the real world," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 23–30, IEEE, 2017.

[2] C. Choi and H. I. Christensen, "Rgb-d object tracking: A particle filter approach on gpu," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1084–1091, IEEE, 2013.

[3] "Dynamic primitives of motor behavior." `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3735361/#R88`. (Accessed on 11/20/2020).

[4] "L1_defsandexamples.pdf." `https://www.math.upenn.edu/~brweber/Courses/2012/Math650/Notes/L1_DefsandExamples.pdf`. (Accessed on 11/20/2020).

[5] M. S. Unreal, "Unreal engine [en línea]."

[6] B. Planche and E. Andres, "Hands-on computer vision with tensorflow 2," 2019.

[7] K. Wada, E. Sucar, S. James, D. Lenton, and A. J. Davison, "Morefusion: Multi-object reasoning for 6d pose estimation from volumetric fusion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14540–14549, 2020.

[8] J. Lampinen and A. Vehtari, "Bayesian approach for neural networks—review and case studies," *Neural networks*, vol. 14, no. 3, pp. 257–274, 2001.

[9] R. Li, Z. Liu, and J. Tan, "A survey on 3d hand pose estimation: Cameras, methods, and datasets," *Pattern Recognition*, vol. 93, pp. 251–272, 2019.