Progress Report

Bardia Mojra

December 12, 2020

Robotic Vision Lab
University of Texas at Arlington

1 To Do

- Write a literature review for [1] and add it Pose Estimation survey paper.
- Implement and play with PoseCNN and DOPE.
- Generate new data set using UE4.
- Read more papers on pose estimation.
- Look into transfer learning. Read [2].
- Look into domain randomization and adaptation.
- Read [3].
- Learn to use UE4.
- Reconstruct a pose estimation model to familiarize myself and then start modifying it.

2 Progress

Following items are listed in order of priority:

- Pose Estimation: I worked on the draft some more, read more papers and read [1] again. I was able to understand [1] more but it is a heavy paper. It is based on [4], VoxelNet [5], MV3D [6], [7], [8], and [9]. They use deep neural nets; they have developed object detection for deep navigation stacks and use methods introduced in [9] to estimation uncertainty. BayesOD was tested on four major 2D object detection datasets, COCO, Pascal VOC, Berkeley Deep Drive and Kitti. They reduced the minimum Gaussian uncertainty error metric by about 11 percent, the minimum Categorical uncertainty error metric by about 3.5 percent, and increased the probabilistic detection quality over the next best method from state of the art by about 1.5 percent. Moreover, they provide Bayesian procedure for deep object detector inference which allows for anchor-level and object-level prioris computation in closed form. They also replace the standard non-maximum suppression method with Bayesian inference which allows for priori information to retain for both the bounding box and the category of a detected object instance.

- OCRTOC:
- Chaotic Systems: Chaotic systems are a class of nonlinear dynamical systems. Classic control systems, introduced by James Clerk Maxwell was based on Newtonian physics. Newtonian physics is a point-mass and force-field system which means it is incapable of solving problems where there are more than two mass bodies exert force. Newton was famously puzzled by the three-body problem which is still difficult unless using linear algebra and modern computers.

  Chaotic systems are defined by being made of completely deterministic entities but over time, as a whole, they behave in a probabilistic manner. This means no matter how much information we collect about initial conditions, no chaotic system, even in simulation, can be repeated except for certain situations, such as Van der Pol oscillator and the Lorenz Attractor. This is due to the fact that no measurement is exact, even slightest errors in initial conditions eventually result in different convergent results. This is known as the "butterfly effect" and was introduced along with Chaos Theory by Ed Lorenz in [10].

  Chaos theory allows for calculation of a range of possible outcomes for a given chaotic system and initial conditions. This is why weather forecasts now provide a range possible hurricane paths and news outlets provide multiple paths to the presidency for each candidate or for each party to control of the house or the senate. I can guess they can make predictions about the weights of underlying (dynamic) factors. Weight estimation would be similar to Adaptive Control, which is also known to be used in Chaotic system control.
- 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 [11]) that enables us to encode important information into our models, whether it is actuator manipulation or dynamic scene understanding. [12] 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.
- TensorFlow [13]: I am still working through chapter 2.
- MoreFusion [14]: Still need to write a literature review on this.

- Reading list: [15] and [16].
- 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] A. Harakeh, M. Smart, and S. L. Waslander, "Bayesod: A bayesian approach for uncertainty estimation in deep object detectors," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 87–93, IEEE, 2020.

[2] 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.

[3] 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.

[4] J. Ku, M. Mozifian, J. Lee, A. Harakeh, and S. L. Waslander, "Joint 3d proposal generation and object detection from view aggregation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1–8, IEEE, 2018.

[5] Y. Zhou and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3d object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4490–4499, 2018.

[6] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3d object detection network for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1907–1915, 2017.

[7] N. Sünderhauf, O. Brock, W. Scheirer, R. Hadsell, D. Fox, J. Leitner, B. Upcroft, P. Abbeel, W. Burgard, M. Milford, *et al.*, "The limits and potentials of deep learning for robotics," *The International Journal of Robotics Research*, vol. 37, no. 4-5, pp. 405–420, 2018.

[8] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*, pp. 1050–1059, 2016.

[9] A. Kendall and Y. Gal, "What uncertainties do we need in bayesian deep learning for computer vision?," in *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach,

R. Fergus, S. Vishwanathan, and R. Garnett, eds.), vol. 30, pp. 5574–5584, Curran Associates, Inc., 2017.

[10] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.

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

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

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

[14] 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.

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

[16] 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.