Progress Report

Bardia Mojra

April 25, 2020

Robotic Vision Lab
The University of Texas at Arlington

1 Progress

Following items are listed in order of priority:

- Finally with Chris' help, I was able to resolve the issue by reinstalling the OS and ROS. I was able to launch TurtleBot from my PC and continued on the tutorials. Unfortunately, I didn't get much time to continue ROS tutorials but I will continue at my own pace until semester ends, and then-on I will work on it full time.
- Regarding SLAM tutorial, I am getting familiar with ROS and continue with SLAM tutorials.
- Continue reading on RL, DRL and SLAM. I began reading on Turtlebot3 and ROS.
- Read , began to dissect it. It's a heavy paper, I have already read it twice, I need to go back and read it again and learn the material.
- In the fall semester I plan on taking Optimization, AI and Intelligent Systems. If it becomes too much like this semester I will drop one because this semester I didn't really get time to do research as much as I wanted to.
- My AI and Robotic goals for the summer comprise of learning ROS, Gazebo, and SLAM and begin working towards a model that can compete in Niko Suenderhauf's Scene Understanding Challenge. Other than that, I will work on finishing everything on this list.
- I still need to work on fellowship applications.
- Still working on [1], [2], [3], [4], [5], and [6].

2 Plans

Following items are listed in order of priority:

- (On pause) Resume Machine Learning course with Andrew Ng as soon as possible.
- (On pause) I am working through Jason Brownlee's ML Mastery book, [7].

- (On pause) Resume Robotic Perception course as soon as possible.
- (On pause) Need to read [8], [9], [10], and [11]; these papers seem fundamental to understanding the overall picture.
- (On pause) Get intimate with Python, Numpy, Pandas, Scipy, and Matplotlib, TensorFlow and PyTorch.
- (On pause) Read Digital Image Processing by Gonzalez and Woods.
- (On pause) Learn ROS.
- (On pause) (Supremely important) Read on scene understanding, semantic SLAM, graph SLAM, visual odometer, place recognition, and Kalman Filtering. Read Niko Sunderhauf's research publications.

3 General Notes

This section summarizes general research leads. The following item are to be investigated, understood and briefly summarized.

- Open3D: An open source toolbox used for truth occupancy grid application and probably other things. Should review.
- Horn's paper [3]: It introduces Unit Quaternions which allow for complex domain representation of kinematics. Very important paper for robotic motion. Should review.
- Bayesian Learning: This probabilistic ML approach treats model parameters as random variables. Read [12] for more details.
- Convex Optimization:
- Q-Learning: A learning model of reinforcement (RL), learning from delayed reward.
- Deep Reinforcement Learning:

4 Literature Review

4.1 Dynamic Graph CNN for Learning on Point Clouds [5]

This paper introduces a new model for training CNNs to learn similar features of point cloud objects.

4.1.1 Keywords

- PointNet
- Extrinsic and intrinsic descriptors:
- Permutation variance:

4.2 Single Image Super-Resolution Using Multi-Scale Convolutional Neural Network - MSSR [13]

Paper proposes an architecture with two parallel path with different depths (which correspond to scales) for residual learning; where one path (module L) is used for large factor up-scaling (x4, x8) and the other (module S) for small factor up-scaling (x2). At the end, it combines the outputs by summation (a form of ensembles). In contrast to previous work where the focus is on small factor up-scaling (x2) and repeat if needed, this model takes higher factor up-scaling into consideration while training the network which helps with reducing blurriness of output image for higher factor up-scaling. The model uses multi-scale residual learning to train on general model for multiple up-scaling factors; hence, saving memory and processing time. This paper provides experiment results that show higher output image integrity where peak signal to noise ratio (PSNR) and structural similarity index (SSIM) are higher or comparatively close to state-of-the-art methods.

4.2.1 Keywords

- Lanczos re-sampling:
- Statistical priors:
- Stochastic Neighbor Embedding (paper by G. Hinton [14]): Read paper.

- Parse coding:
- GoogLeNet:
- YCbCr color space:
- Caffe package (paper by Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.): Convolutional architecture for fast feature embedding. Read and write literature review.
- Adam method (paper by Kingma, D., Ba, J.): A method for stochastic optimization.
- PSNR (performance metric): Peak Signal to Noise Ratio represents the ratio between max (peak) possible value (power) of a signal (image) and the power of distorting noise.
- SSIM (performance metric): Structural Similarity Index, is the ratio of structural features of a processed image to the original image. The value represent percentage of structural features/information retained throughout image processing.
- A+ (SR method, paper):
- SelfEx (SR method, paper):
- SRCNN (SR method, paper): Image Super Resolution Using Deep Convolutional Networks. Read again and write literature review.
- FSRCNN (SR method, paper):
- VDSR (SR method, paper):

4.3 Value Iteration Networks - VIN [15]

4.3.1 Keywords

- Imitation Learning: **need to finish reading [16].**
- CNNs applied to reinforcement learning:
- MDP (Markov Decision Process):
- VI Algorithm (Value Iteration):

- SGD (stochastic gradient decent):
- Theano Code: "Theano is a Python library that allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently. It can use GPUs and perform efficient symbolic differentiation."
- Grid-World Domain:

References

[1] 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.

[2] 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.

[3] B. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society A*, vol. 4, pp. 629–642, 04 1987.

[4] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from rgbd images," pp. 746–760, 10 2012.

[5] 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.

[6] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–1, 2018.

[7] J. Brownlee, *Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End.* Machine Learning Mastery, 2016.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *CoRR*, vol. abs/1501.00092, 2015.

[9] 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.

[10] 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.

[11] 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.

[12] "Bayesian learning," 2012.

[13] X. Jia, X. Xu, B. Cai, and K. Guo, "Single image super-resolution using multi-scale convolutional neural network," *Lecture Notes in Computer Science*, p. 149–157, 2018.

[14] G. Hinton and S. Roweis, "Stochastic neighbor embedding," *Advances in neural information processing systems*, vol. 15, pp. 833–840, 2003.

[15] A. Tamar, Y. WU, G. Thomas, S. Levine, and P. Abbeel, "Value iteration networks," in *Advances in Neural Information Processing Systems 29* (D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, eds.), pp. 2154–2162, Curran Associates, Inc., 2016.

[16] V. Kurin, "Introduction to imitation learning," Aug 2017.