Suggested projects for ECE-GY 6123 Image and Video Processing (Spring 2019)

Updated 2/17/2019

Please email the contact person to set up an appointment to learn more about the project. If you want to choose a project outside this list, you should discuss your idea with the instructor before deciding.

- Tracking the movements of plants, as part of our PlantTracer Project (http://planttracer.com/) (Contact: Yixiang Mao, yixiang.mao@nyu.edu)
    - Tracking plant apex using classical tracking approaches (e.g. KLT tracker, multi-view block matching)
    - Using deep learning approach (you will help to annotate ground truth data as well as training and testing the network)

- Moving Foreground detection in video using RPCA or other sparse-representation-based optimization methods. If you have excellent results, you can enter the competition http://www.changedetection.net/
(Contact: Amirhossein Khalilian akg404@nyu.edu).

Ref:
Candès, Emmanuel J., et al. "Robust principal component analysis?." *Journal of the ACM (JACM)* 58.3 (2011): 11.
X. Cao, L. Yang and X. Guo, "Total Variation Regularized RPCA for Irregularly Moving Object Detection Under Dynamic Background," in *IEEE Transactions on Cybernetics*, vol. 46, no. 4, pp. 1014-1027, April 2016.

Background needed: Linear Algebra. Convex optimization background is desired but not necessary.

- Image/video deblurring or super-resolution with or without knowing the blurring kernels (Contact: Amirhossein Khalilian akg404@nyu.edu)

Ref: S. Farsiu, M. D. Robinson, M. Elad and P. Milanfar, "Fast and robust multiframe super resolution," in *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327-1344, Oct. 2004.
Background needed: Linear Algebra background is required. Additionally, you need to know some basic topics of convex optimization.

- 360 degree or panoramic video or image stitching (contact: Yixiang Mao ym1496@nyu.edu)

Ref:

Brown, Matthew, and David G. Lowe. "Automatic panoramic image stitching using invariant features." *International journal of computer vision* 74.1 (2007): 59-73.
Szeliski, Richard. "Image alignment and stitching: A tutorial." *Foundations and Trends in Computer Graphics and Vision* 2.1 (2006): 1-104.
Shum, Heung-Yeung, and Richard Szeliski. "Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment." *International Journal of Computer Vision* 36.2 (2000): 101-130.

Background needed: Image registration and warping

- 360 Degree Video Saliency Estimation and/or View Prediction (contact: Weixi Zhang, wz1219@nyu.edu)

References:

Pan, Junting, et al. "Shallow and deep convolutional networks for saliency prediction." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
Wu, Chenglei, et al. "A Dataset for Exploring User Behaviors in VR Spherical Video Streaming." *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 2017.
Lo, Wen-Chih, et al. "360 Video Viewing Dataset in Head-Mounted Virtual Reality." *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 2017.
Bao, Yanan, et al. "Viewing 360 degree videos: Motion prediction and bandwidth optimization." *Network Protocols (ICNP), 2016 IEEE 24th International Conference on*. IEEE, 2016.
Bao, Yanan, et al. "Shooting a moving target: Motion-prediction-based transmission for 360-degree videos." *Big Data (Big Data), 2016 IEEE International Conference on*. IEEE, 2016.
https://code.facebook.com/posts/118926451990297/enhancing-high-resolution-360-streaming-with-view-prediction/

Background needed: deep learning

- Segmentation of ultrasound images of mouse embryos. (contact: Ziming Qiu zq415@nyu.edu and Nitin Nair nn1174@nyu.edu)

Ziming Qiu, Jack Langerman, Nitin Nair, Orlando Aristizabal, Jonathan Mamou, Daniel H. Turnbull, Jeffrey Ketterling, Yao Wang, “Deep BV: A Fully Automated System for Brain Ventricle Localization and Segmentation in 3D Ultrasound Images of Embryonic Mice“, IEEE Signal Processing in Medicine and Biology Symposium (SPMB), 2018, arXiv preview link.

Background: deep learning

- Medical Image registration (cross-modality, or within the same modality) (Contact: Amirhossein Khalilian akg404@nyu.edu)

- Image or video compression using deep learning (Contact: Changyue Ma, cm5154@nyu.edu)

Refs.:
Video Compression through Image Interpolation,
https://arxiv.org/abs/1804.06919

Learned Video Compression, **https://arxiv.org/abs/1811.06981**

Background: image compression and deep learning

- Multi-mode intra-prediction for image compression using diversity encouraged ensemble learning (Contact: Yao Wang, yaowang@nyu.edu)

References

· Guzman-Rivera, D. Batra, and P. Kohli. Multiple choice learning: Learning to produce multiple structured outputs. In Advances in Neural Information Processing Systems, pages 1799–1807, 2012.
· S. Lee, S. Purushwalkam, M. Cogswell, D. Crandall, and D. Batra. Why m heads are better than one: Training a diverse ensemble of deep networks. arXiv preprint arXiv:1511.06314, 2015.
· Yilin Song, Jonathan Viventi, and Yao Wang, Diversity encouraged learning of unsupervised LSTM ensemble for neural activity video prediction, July 2018. https://arxiv.org/abs/1611.04899.

Background: image compression and deep learning.

- Multiple objective deep learning: feature learning for image compression and visual analytics (Contact: Yao Wang, yaowang@nyu.edu)

Background: image compression and deep learning.

- Other possible projects in medical image analysis: you can pick one of the challenges in the link below and work on it. Please discuss with the instructor once you have several choices after you have read the background material. I will gauge whether you have sufficient background and time to do it or suggest a subset to work on.
https://grand-challenge.org/challenges/

- Other possible projects in computer vision: you can pick one of the challenges in the links below and work on it. Please discuss with the instructor once you have several choices after you have read the background material. I will gauge whether you have sufficient background and time to do it or suggest a subset to work on.
http://www.robustvision.net/

http://www.icme2019.org/conf_challenges

Added topics (2/17/2019)

· Optical flow estimation using deep learning:

[PWCnet] Deqin Sun, Xiaodong Yang, Ming-Yu Liu, Jan Kautz, PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. CVPR 2018. https://arxiv.org/abs/1709.02371

[Flownet2] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "Flownet 2.0: Evolution of optical flow estimation with deep networks.," in CVPR . 2017, pp. 1647–1655, IEEE Computer Society
https://lmb.informatik.uni-freiburg.de/Publications/2017/IMKDB17/

[FlowFields] C. Bailer, B. Taetz, and D. Stricker. Flow fields: Dense correspondence fields for highly accurate large displacement optical flow estimation. In IEEE Int. Conference on Computer Vision (ICCV), 2015.

Older papers:
[DeepFlow] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid. Deepflow: Large displacement optical flow with deep matching. In IEEE Int. Conference on Computer Vision (ICCV), 2013.

[EpicFlow] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid. Epicflow: Edge-preserving interpolation of correspondences for optical flow. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

· Depth estimation from single image:

Ref: Zhengqi Li and Noah Snavely, "Megadepth: Learning single-view depth prediction from internet photos," in Computer Vision and Pattern Recognition (CVPR), 2018.

· Video prediction: predicting future frames from prior frames using deep learning method

[PredRNN] Y. B.Wang, Z. F. Gao, M. S. Long, J. M.Wang, and P. S. Yu, "Predrnn++: Towards a resolution of the deep-intime dilemma in spatiotemporal predictive learning.," in ICML. 2018, vol. 80 of JMLRWorkshop and Conference Proceedings, pp. 5110–5119, JMLR.org.

[TBnet] J. W. Xu, B. B. Ni, Z. F. Li, S. Cheng, and X. K. Yang, "Structure preserving video prediction," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

Older papers:

[DVF] Z. W. Liu, R. A. Yeh, X. O. Tang, Y. M. Liu, and A. Agarwala, “Video frame synthesis using deep voxel flow.,” in ICCV. 2017, pp. 4473–4481, IEEE Computer Society.

[DualGAN] X. D. Liang, L. Lee, W. Dai, and E. P. Xing, “Dual motion gan for future-flow embedded video prediction.,” in ICCV . 2017, pp. 1762–1770, IEEE Computer Society.

[EpicFlow] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, “Epicflow: Edge-preserving interpolation
of correspondences for optical flow.,” in CVPR . 2015, pp. 1164–1172, IEEE Computer Society.

[NextFlow] N. Sedaghat, “Next-flow: Hybrid multi-tasking with next-frame prediction to boost optical-flow estimation in the wild.,” CoRR, vol. abs/1612.03777, 2016.

[CNDA] C. Finn, I. J. Goodfellow, and S. Levine, “Unsupervised learning for physical interaction through video prediction.,” in NIPS , Daniel D. Lee, Masashi Sugiyama, Ulrike V. Luxburg, Isabelle Guyon, and Roman Garnett, Eds., 2016, pp. 64–72.

[MCnet] R. Villegas, J. M. Yang, S. Hong, X. Y. Lin, and H. Lee, “Decomposing motion and content for natural video sequence prediction.,” CoRR, vol. abs/1706.08033, 2017.

M. Mathieu, C. Couprie, and Y. LeCun, “Deep multiscale video prediction beyond mean square error.,” CoRR, vol. abs/1511.05440, 2015.

· Video interpolation: predicting intermediate frames from surrounding frames using deep learning.

· Image inpainting: filling holes in an image

G. L. Liu, F. A. Reda, K. J. Shih, T. C. Wang, A. Tao, and B. Catanzaro, “Image inpainting for irregular holes using partial convolutions.,” in ECCV (11) . 2018, vol. 11215 of Lecture Notes in Computer Science , pp. 89–105, Springer.