جـــامـــعـــة مـحـمـد بـــن زايـــد
لـلـــذكـــاء الاصـــطـــنـــاعـــي
MOHAMED BIN ZAYED UNIVERSITY
OF ARTIFICIAL INTELLIGENCE

Core Courses Syllabi

CV705 - Advanced 3D Computer Vision

Title	Advanced 3D Computer Vision
Code	CV705
Loading	4 Credit-hours
Prerequisites	• A thorough understanding of the material covered in the following Masters course: 1. CV 702 Geometry for Computer Vision • Hands-on experience with Python and Pytorch or equivalent language/library
Catalog Description	The course exercises an in-depth coverage of special topics in 3D computer vision. The students will be able to critique the state-of-the-art methods on 3D reconstruction, 3D visual scene understanding and multi-view stereo. In addition, students will have to implement papers to accomplish the following goals: (1) reproduce results reported in the papers, and (2) improve the performance of published peer-reviewed works. This course builds upon concepts from Human and Computer Vision (CV701), Geometry for Computer Vision (CV702) and assumes that the students are familiar with the basic concepts of machine learning and optimization.
Goal	This PhD course aims to inculcate a deeper understanding of the algorithms for 3D reconstruction, 3D scene understanding, and multi-view stereo so the students are capable of researching, developing, and implementing these methods for solving computer vision problems of real-world scale. Additionally, a significant goal of this course is to enhance students' teamwork skills by requiring them to participate in group projects and develop their communication and analytical skills by engaging them in reading group activities.
Contents	The course covers three modules: **(I)** 3D reconstruction from RGB-D images, **(II)** 3D Visual Scene Understanding, and **(III)** Multi-view Stereo
Recommended Textbooks	1. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, second edition, Cambridge, 2003. 2. Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3D Vision: from Images to Geometric Models*, Springer Science & Business Media, 2012.
Recommended References & Supplemental Material	Relevant research papers, tech reports, and surveys for each topic, where needed, are identified in the teaching plan ahead. In addition, the following textbooks will be useful: 1. E. Trucco, and A. Verri, *Introductory Techniques for 3D Computer Vision*, Prentice Hall, 1998. 2. C. Wöhler, *3D Computer Vision: Efficient Methods and Applications*, Springer Science & Business Media, 2012.

Teaching Week	Topics
1	**3D Reconstruction from RGB-D images (Module I)** This module covers the following topics: 3D reconstruction from RGB-D images, 3D reconstruction of static scenes, 3D reconstruction of dynamic scenes and appearance reconstruction approaches. **Lecture** • Introduction to 3D reconstruction from RGB-D images • Overview of RGB-D cameras and their core algorithms • Relevant papers: - S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera", *ACM symposium on User interface software and technology*, 2011. - J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review", *IEEE transactions on cybernetics*, 2013. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-1 work
2	**3D Reconstruction from RGB-D images (contd.)** **Lecture** • 3D reconstruction of static scenes • Reading group activity on selected papers related to "3D reconstruction of static scenes": - D. Scharstein, and C. Pal, "Learning conditional random fields for stereo", *IEEE Conference on computer vision and pattern recognition*, 2017. - A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang, "Matterport3D: Learning from RGB-D data in indoor environments", *arXiv preprint arXiv:1709.06158*, 2017. **Lab** • Continue Project-1 work
3	**3D Reconstruction from RGB-D Images (contd.)** **Lecture** • 3D reconstruction of dynamic scenes • Reading group activity on selected papers related to "3D reconstruction of dynamic scenes": - M. Dou, and H. Fuchs, "Temporally enhanced 3d capture of room-sized dynamic scenes with commodity depth cameras", *IEEE Virtual Reality*, 2014. - V. Golyanik, T. Fetzer, and D. Stricker, "Accurate 3d reconstruction of dynamic scenes from monocular image sequences with severe occlusions", *IEEE winter conference on applications of computer vision*, 2017. **Lab** • Continue Project-1 work

Teaching Week	Topics
4	**3D Reconstruction from RGB-D Images (contd.)** **Lecture** • Appearance reconstruction approaches • Group discussion on the relevant papers - M. Hachama, B. Ghanem, and P. Wonka, "Intrinsic scene decomposition from RGB-D images", *IEEE International conference on computer vision*, 2017. - X. Zuo, S. Wang, J. Zheng, and R. Yang, "Detailed surface geometry and albedo recovery from RGB-D video under natural illumination", *IEEE International conference on computer vision*, 2017. **Lab** • Preparation of presentation on Project-1 work • Continue Project-1 work
5	**3D Reconstruction from RGB-D images** **Assessment 1.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 1.2** • In-class exam covering Module I – 3D Reconstruction from RGB-D images.
6	**3D Visual Scene Understanding (Module II)** This module covers the following: 1) 3D object detection, and 2) 3D semantic segmentation. **Lecture** • Introduction to 3D scene understanding • Relevant papers and assigned reading: - B. S. Kim, P. Kohli, and S. Savarese, "3D scene understanding by voxel-CRF", *IEEE International conference on computer vision*, 2013. - M.Z. Zia, M. Stark, and K. Schindler, "Towards scene understanding with detailed 3D object representations", *International journal of computer vision*, 2015. - Y. Shi, P. Long, K. Xu, H. Huang, and Y. Xiong, "Data-driven contextual modeling for 3d scene understanding", *Computers & graphics*, 2015. - A. Geiger, M. Lauer, and R. Urtasun, "A generative model for 3d urban scene understanding from movable platforms", *IEEE Conference on computer vision and pattern recognition*, 2011. - M. Naseer, S. Khan, and F. Porikli, "Indoor Scene Understanding in 2.5/3D for Autonomous Agents: A Survey", *IEEE Access*, 2019. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-2 work

Teaching Week	Topics
7	**3D Visual Scene Understanding (contd.)** **Lecture** • 3D object detection - I • Reading group activity on relevant papers: - B. Pepik, M. Stark, P. Gehler, and B. Schiele, "Multi-view and 3d deformable part models", *IEEE transactions on pattern analysis and machine intelligence*, 2015. - X. Zhao, Z. Liu, R. Hu, and K. Huang, "3D object detection using scale invariant and feature reweighting networks", *arXiv preprint arXiv:1901.02237*, 2019. **Lab** • Continue Project-2 work
8	**3D Visual Scene Understanding (contd.)** **Lecture** • 3D object detection - II • Reading group activity on relevant papers: - B. Yang, W. Luo, and R. Urtasun, "PIXOR: real-time 3D Object detection from point clouds", *IEEE Conference on computer vision and pattern recognition*, 2018. - S. Shi, X. Wang, and H. Li, "Pointrcnn: 3d object proposal generation and detection from point cloud", *arXiv preprint arXiv:1812.04244*, 2018. **Lab** • Continue Project-2 work
9	**3D Visual Scene Understanding (contd.)** **Lecture** • Introduction to 3D semantic segmentation • Discussion of relevant papers: - Q. H. Pham, B. S. Hua, T. Nguyen, and S. K. Yeung, "Real-time progressive 3D semantic segmentation for indoor scenes", *IEEE winter conference on applications of computer vision*, 2019. - F. Engelmann, T. Kontogianni, J. Schult, and B. Leibe, "Know what your neighbors do: 3D semantic segmentation of point clouds", *European conference on computer vision*, 2018. **Lab** • Preparation of presentation on project-2 work • Continue Project-2 work
10	**3D Visual Scene Understanding (contd.)** **Assessment 2.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 2.2** • In-class exam covering Module II – 3D Visual Scene Understanding

Teaching Week	Topics
11	**Multi-View Stereo (Module III)** This module covers an introduction to multi-view stereo, motion estimation from stereo and depth estimation from stereo pairs. **Lecture** • Introduction to multi-view stereo • Relevant papers and assigned reading: - A. Kar, C. Häne, C. and J. Malik, "Learning a multi-view stereo machine", *Advances in neural information processing systems*, 2017. - J. L. Schonberger, and J. M. Frahm, "Structure-from-motion revisited", *IEEE conference on computer vision and pattern recognition*, 2016. - W. Luo, A. G. Schwing, and R. Urtasun, "Efficient deep learning for stereo matching", *IEEE conference on computer vision and pattern recognition*, 2016. - A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. Van Der Smagt, D. Cremers, and T. Brox, "Flownet: Learning optical flow with convolutional networks", *IEEE international conference on computer vision*, 2015. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-3 work
12	**Multi-View Stereo (contd.)** **Lecture** • Motion estimation from stereo pairs • Discussion of relevant papers: - Weinzaepfel, P., Revaud, J., Harchaoui, Z. and Schmid, C., 2013. DeepFlow: Large displacement optical flow with deep matching", *IEEE international conference on computer vision*, 2013. - C. Bailer, B. Taetz, and D. Stricker, "Flow fields: Dense correspondence fields for highly accurate large displacement optical flow estimation", *IEEE international conference on computer vision*, 2015. - J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, "Epicflow: Edge-preserving interpolation of correspondences for optical flow", *IEEE conference on computer vision and pattern recognition*, 2015. - E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, "Flownet 2.0: Evolution of optical flow estimation with deep networks", *IEEE conference on computer vision and pattern recognition*, 2017. **Lab** • Continue Project-3 work

Teaching Week	Topics
13	**Multi-View Stereo (contd.)** **Lecture** • Depth estimation from stereo pairs • Discussion of relevant papers: - J. Zbontar, and Y. LeCun, “Stereo matching by training a convolutional neural network to compare image patches”, Journal of machine learning research, 2016. - W. Luo, A. G. Schwing, and R. Urtasun, “Efficient deep learning for stereo matching”, *IEEE conference on computer vision and pattern recognition*, 2016. - A. Kendall, H. Martirosyan, S. Dasgupta, P. Henry, R. Kennedy, A. Bachrach, and A. Bry, “End-to-end learning of geometry and context for deep stereo regression”, *IEEE conference on computer vision and pattern recognition*, 2017. **Lab** • Continue Project-3 work
14	**Multi-View Stereo (contd.)** **Lecture** • Joint depth and motion estimation from stereo images • Discussion of relevant papers: - Vijayanarasimhan, S., Ricco, S., Schmid, C., Sukthankar, R. and Fragkiadaki, K., 2017. SFM-Net: Learning of structure and motion from video. arXiv preprint arXiv:1704.07804. - B. Ummenhofer, H. Zhou, J. Uhrig, N. Mayer, E. Ilg, A. Dosovitskiy, and T. Brox, “Demon: Depth and motion network for learning monocular stereo”, *IEEE conference on computer vision and pattern recognition*, 2017. **Lab** • Preparation of presentation on project-3 work • Continue Project-3 work
15	**Multi-View Stereo** **Assessment 3.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 3.2** • In-class exam covering Module III – Multi-View Stereo