3DV 2016 Tutorials

Tutorial 1
Title: 3D Object Geometry from Single Image
Organizers:
Xiaowei Zhou, Postdoctoral Researcher, University of Pennsylvania
Yu Xiang, Postdoctoral Researcher, University of Washington
Kostas Daniilidis, Professor, University of Pennsylvania
Description: In the past decade, we have seen a revolution in 3D object recognition techniques from geometric approaches with handcrafted features to learning-based approaches with data-driven representations. With the fast development of deep learning approaches and increasing availability of datasets, recovering 3D geometry of objects from a single image become a possible task. In this course, we will talk about the state-of-the-art techniques for 3D object category detection, pose estimation, keypoint localization, shape reconstruction as well as 3D human pose estimation. We will have hands-on opportunities using datasets and open source code.

Tutorial 2
Title: Large-scale 3D Modeling from Crowdsourced Data
Organizers:
Jan Michael Frahm, Associate Professor, University of North Carolina at Chapel Hill
Enrique Dunn, Associate Professor, Stevens Institute of Technology
Jared Heinly, Senior Research Engineer, URC Ventures
Johannes L. Schönberger, Ph.D. Student, ETH Zürich
Description: Large scale image based 3D modeling has been a major goal of computer vision, enabling a wide range of applications including virtual reality, image based localization, and autonomous navigation. One of the most diverse data sources for modeling is Internet photo collections. In the last decade the computer vision community has made tremendous progress in large scale structure from motion and multi-view stereo from Internet datasets. However, utilizing this wealth of information for 3D modeling remains a challenging problem due to the ever increasing amount of image data. In a short period of time, research in large scale modeling has progressed from modeling using several thousand images, to modeling from city-scale datasets of several million, and recently to reconstructing an Internet scale dataset comprising of 100 million images. This tutorial will present the main underlying technologies enabling these innovations.

Tutorial 3
Title: Semantic and Structured 3D Modeling
Organizers:
Christian Häne, UC Berkeley
Srikumar Ramalingam, MERL
Sudipta Sinha, Microsoft Research
Description: 3D scene reconstruction from multiple images is a fundamental topic in computer vision which has witnessed rapid progress in the last two decades. Since 3D reconstruction techniques involve solving parameter estimation problems that are inherently ill-posed, they require appropriate regularization to deal with noise and ambiguities in the input. While traditional methods mostly relied on low-level geometric priors, nowadays, mid and high-level scene information in the form of structured and semantic scene priors are increasingly being used. Deep learning methods are also being used to learn end-to-end models that can be trained on labeled images. In this tutorial, we will review all these new developments.

Throughout the tutorial, we will informally use the term -- semantic and structured modeling to refer to methods that recover the semantic classes (implicitly or explicitly) or use representations or constraints that encode geometrically meaningful 3D scene priors. We will not cover topics such as structure-from-motion in detail but will instead focus on recent dense reconstruction techniques, optimization methods and learning-based approaches.

Tutorial 4
Title: Understanding 3D and Visuo-Motor Learning
Organizers:
John Schulman, OpenAI
Sergey Levine, UC Berkeley
Amir Zamir, Stanford
Saurabh Gupta, UC Berkeley
Erwin Coumans, Google
Joseph Lim, Stanford
Description: Recent advances in deep reinforcement learning for closing the visuomotor loop have led to interesting questions about the representation of state spaces and the level of detail required from visual recognition.

A number of recent results demonstrated the performance of dynamic tasks, such as grasping and poking, without explicit 3D models of the world, using rewards in visual feature spaces. However, to scale the approach up to more complex tasks involving affordance-based manipulation the use of 3D understanding has potential to benefit policy learning through the use of more benign state space representations.

This tutorial aims to bring together recent results in both Visuo-Motor learning along with advances in the detailed understanding of 3D environment. The hope is for the speakers to present their work as a tutorial to include both fundamentals and new results.