NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

C S   S E M I N A R

Title: **"Advanced Robotics Center Colloquium" in conjunction with "Computer Vision for Robotics" Workshop**

Speaker: ARC Colloquium:
                 Speaker: Prof. Dr. Marc Pollefeys
                 Department of Computer Science, ETH Zurich

                 Computer Vision for Robotics Workshop:
                 Speaker 1: Prof. Dr. Andreas Geiger
                 Max Planck Research Group Leader,
                 MPI for Intelligent Systems, Tubingen

                 Speaker 2: Dr. Torsten Sattler
                 Post-doctoral researcher,
                 Computer Vision and Geometry Group, ETH Zurich

                 Speaker 3: Dr. Lionel Heng
                 Deputy Lab Head & Senior Member of Technical Staff,
                 DSO National Laboratories

                 Speaker 4: Prof. Dr. Gim Hee Lee
                 Assistant Professor, Department of Computer Science,
                 National University of Singapore

Date/Time: 25 November 2016, Friday, 01:00 PM to 06:00 PM

Venue: LT19, COM2 1evel 1

Chaired by: Assistant Professor Lee Gim Hee
                 (mpelgh@nus.edu.sg)

There will be time for interaction with the speaker at the end of the seminar. Light refreshments will be served. Please register at https://goo.gl/forms/d38sRSY3ORMXbZud2 by 18 November 2016.

ARC Colloquium Talk: Semantic 3D Reconstruction

Abstract:

While purely geometric models of the world can be sufficient for some applications, there are also many applications that need additional semantic information. In this talk I will focus on 3D reconstruction approaches which combine geometric and appearance cues to obtain semantic 3D reconstructions. Specifically, the approaches I will discuss are formulated as multi-label volumetric segmentation, i.e. each voxel gets assigned a label corresponding to one of the semantic classes considered, including free-space. We propose a formulation representing raw geometric and appearance data as unary or high-order (pixel-ray) energy terms on voxels, with class-pair-specific learned anisotropic smoothness terms to regularize the results. We will see how by solving both reconstruction and segmentation/recognition jointly the quality of the results improves significantly and we can make progress towards 3D scene understanding.

Biodata:

Marc Pollefeys is director of science at Microsoft HoloLens and a full professor of the Dept. of Computer Science at ETH Zurich. Previously he was on the faculty at the University of North Carolina at Chapel Hill. He obtained his MS and PhD degrees from the KU Leuven in Belgium. His main area of research is computer vision. Dr. Pollefeys has received several prizes for his research, including a Marr prize, an NSF CAREER award, a Packard Fellowship and a ERC Starting Grant. He is the author or co-author of more than 280 peer-reviewed papers. He will be general chair of ICCV 2019, was a general chair for ECCV 2014 in Zurich and one of the program chairs for the IEEE Conf. on Computer Vision and Pattern Recognition 2009. He was also general chair of the predecessor conferences to 3DV in 2006 and 2012. Prof. Pollefeys was on the Editorial Board of the IEEE Transactions on Pattern Analysis and Machine Intelligence, the International Journal of Computer Vision, Foundations and Trends in Computer Graphics and Computer Vision and several other journals. He is an IEEE Fellow.

Computer Vision for Robotics Workshop:
Talk 1: Robust Visual Perception for Intelligent Systems

Abstract:

Perception is a key component of every intelligent system as it enables actions within a changing environment. While humans perceive their environment with seemingly little efforts, computers first need to be trained for these tasks. One of the biggest challenges in computer vision are ambiguities which arise due to the complex nature of our environment and the information loss caused by observing two-dimensional projections of our three-dimensional world. In this talk, I will present several recent results in stereo estimation, 3D reconstruction and motion estimation which integrate high-level non-local prior knowledge for resolving ambiguities that can't be resolved using local assumptions alone. Furthermore, I will discuss the "curse of dataset annotation" and present a method for augmenting video sequences efficiently with semantic information.

Biodata:

Andreas Geiger is a Max Planck Research Group Leader at the MPI for Intelligent Systems

in Tubingen and a Visiting Professor at ETH Zurich. Before this, he was a research scientist in the Perceiving Systems department at MPI Tubingen. He studied at KIT, EPFL and MIT and received his PhD degree in 2013 from the Karlsruhe Institute of Technology. His research interests are at the intersection of 3D reconstruction and visual scene understanding with a particular focus on rich semantic and geometric priors for bridging the gap between low-level and high-level vision. His work has received several prices, including the Ernst-Schoemperlen Award 2014, as well as best paper awards at CVPR 2013, GCPR 2015 and 3DV 2015. He is an associate member of the Max Planck ETH Center for Learning Systems and serves as area chair and associate editor in computer vision (CVPR, ECCV, PAMI).

Talk 2: Image-Based Localization at Large and on Small Devices

Abstract:

Image-based localization is the problem of determining the exact position and orientation of a camera based on an image. It is encountered in numerous applications such as self-driving cars and other autonomous vehicles, robotics, Structure-from-Motion, Augmented / Mixed / Virtual Reality, and any scenario requiring to localize and staying localized in an environment based on camera data. The image-based localization problem is typically solved by establishing 2D-3D matches between pixels in a camera image and points in a 3D model of the scene, followed by estimating the camera pose from these matches. After briefly introducing the standard pipeline, we mostly focus on two variants of the localization problem: Localization at a large, e.g., city, scale and real-time localization on mobile devices. The first part of the talk focuses on problems of large-scale localization: Feature matching becomes harder due to globally repeating structures, while geometric configurations become less distinctive at the same time. Consequently, advanced metrics need to be developed in order to decide with high precision whether an image has been localized successfully. The second part of the talk concentrates on enabling localization on mobile devices. We show how real-time camera pose tracking can be combined with non real-time localization techniques to adhere to the limited computational capabilities of mobile devices such as tablets or smart phones. In addition, we demonstrate how client-server architectures and compression schemes can be used to handle the memory restrictions of such devices. The talk concludes with an overview over interesting open challenges, including some example scenarios where classical feature matching methods fail.

Biodata:

Torsten Sattler received a PhD in Computer Science from RWTH Aachen University, Germany, in 2013 under the supervision of Bastian Leibe and Leif Kobbelt. Since December 2013, he is a post-doctoral researcher in the Computer Vision and Geometry Group of Prof. Marc Pollefeys at ETH Zurich, Switzerland. His research interests include (large-scale) image-based localization using Structure-from-Motion point clouds, real-time localization and SLAM on mobile devices, image retrieval and efficient spatial verification, camera calibration and pose estimation, Structure-from-Motion, as well as dense, real-time 3D reconstruction on mobile devices. As part of ETH Zurich's efforts, Torsten is contributing to Google's Project Tango and the Trimbot2020 project on building a gardening robot for hedge trimming. In the past, he worked on the V-Charge project on self-driving cars.

Talk 3: Toward Full Autonomy for Vision-Guided Robots: From Self-Calibration to Self-Directed Exploration

Abstract:

Cameras are increasingly used for exteroceptive sensing on robots due to their small size and low cost. In this talk, I will describe my work on the basic building blocks required for vision-guided robots to be autonomous and to be able to execute complex tasks. These building blocks include automatic calibration, visual odometry/SLAM, mapping, and path planning.

Biodata:

Lionel Heng is the Deputy Lab Head and a Senior Member of Technical Staff in the Manned - Unmanned Programme at DSO National Laboratories. Currently, he works at the intersection of 3D computer vision and field robotics. He did his Ph.D. in computer science at the Computer Vision and Geometry Group at ETH Zurich from 2010 to 2014 and under the supervision of Prof. Marc Pollefeys. At ETH Zurich, he did research on autonomous vision-guided robots that included quadrotor/hexacopter MAVs and drive-by-wire cars. His research included self-calibration of multi-camera systems, SLAM with a multi-camera system, stereo-based environment mapping, path planning, exploration, and coverage. He participated in two European-Commission-funded research projects: sFly and V-Charge, that involved vision-guided robots and collaboration across multiple academic and industry partners. He received his undergraduate degree in computer science with an additional major in economics from Carnegie Mellon University in 2006, and his Master's degree in computer science from Stanford University in 2007.

Talk 4: Motion and Pose Estimation with a Multi-Camera System

Abstract:

Multiple cameras that are rigidly fixed onto a body can be set in a configuration that maximizes the field-of-view, which leads to an increase of feature correspondences that are crucial for motion and pose estimation. However, the non-central projection of a multi-camera system makes it inefficient and/or impossible to use existing algorithms from monocular and stereo cameras for motion and pose estimation. In this talk, I will show that by making use of the Plucker coordinates to represent point or line correspondences, and various motion constraints, efficient minimal solvers which are suitable for robust estimation within RANSAC can be derived for motion and pose estimation with a multi-camera system.

Biodata:

Dr. Gim Hee Lee is currently an assistant professor in the Department of Computer Science at the National University of Singapore (NUS). Prior to joining NUS, he was a researcher at the Mitsubishi Electric Research Lab (MERL) in Cambridge Massachusetts. He holds a PhD in Computer Science from ETH Zurich, and his research interest is in the area of Computer Vision and Robotics.