1. Annotation

Course Description

Geometry plays an extremely important role in many computer vision algorithms as certain kinds of
geometric transformations (e.g., projective) form the basis of imaging, estimation, and reconstruction. This
course focuses on obtaining 3D scene geometry from both images and depth sensory data. We will cover
principles of projective geometry and camera models, monocular, stereo, and multi-view vision as well as
the fundamentals of depth sensing and digital geometry processing from range-images. The course only
slightly relies on previous knowledge of deep learning, yet features some deep architectures for 3D data
processing. Most of the material, however, is devoted to more principled topics of computer vision such as
camera calibration, stereo-matching, registration, reconstruction, among others.

Course Prerequisites / Recommendations

Linear algebra, mathematical analysis, algorithms. Intermediate (or
stronger) programming skills are necessary!
First-year Data Science track students aren’t eligible for credits but are
allowed to attend the course.
### Course Academic Level

Master-level course suitable for PhD students

### Number of ECTS credits

3

<table>
<thead>
<tr>
<th>Topic</th>
<th>Summary of Topic</th>
<th>Lectures (# of hours)</th>
<th>Seminars (# of hours)</th>
<th>Labs (# of hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction. Mathematical Background: Linear Algebra</td>
<td>Course overview. Vector Spaces, Linear Transformations and Matrices, Properties of Matrices, Singular Value Decomposition</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Representing a Moving Scene</td>
<td>The Origins of 3D Reconstruction, 3D Space &amp; Rigid Body Motion, The Lie Group SO(3), The Lie Group SE (3), Representing the Camera Motion, Euler Angles Representing a Moving Scene</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Perspective Projection and Camera Calibration</td>
<td>Mathematical Representation, Intrinsic Parameters, Spherical Projection, Radial Distortion, Preimage and Coimage, Projective Geometry, Image Rectification</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Estimating Point Correspondence</td>
<td>From Photometry to Geometry, Small Deformation &amp; Optical Flow, The Lucas-Kanade Method, Feature Point Extraction, Wide Baseline Matching</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Sparse Reconstruction from Two Views: Linear Algorithms</td>
<td>The Reconstruction Problem, The Epipolar Constraint, Eight-Point Algorithm, Structure Reconstruction, Four-Point Algorithm, The Uncalibrated Case</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Sparse Reconstruction from Multiple Views</td>
<td>From Two Views to Multiple Views, Preimage &amp; Coimage from Multiple Views, From Preimages to Rank Constraints, Geometric Interpretation, The Multiple-view Matrix, Relation to Epipolar Constraints, Multiple-View Reconstruction Algorithms, Multiple-View Reconstruction of Lines</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Bundle Adjustment &amp; Nonlinear Optimization</td>
<td>Optimality in Noisy Real World Conditions, Bundle Adjustment, Nonlinear Optimization, Gradient Descent, Least Squares Estimation, Newton Methods, The Gauss-Newton Algorithm, The Levenberg-Marquardt Algorithm</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Dense Reconstruction from Multiple Views</td>
<td>Dense Binocular Stereo, Plane Sweep Framework, Multi-View Stereo, Multi-View Data Association</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Surface Reconstruction</td>
<td>Robust Depthmap Estimation, Fusion and Surface Estimation, Surface Reconstruction</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

### 3. Assignments
<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Assignment Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>Solve theoretical and practical exercises involving SVD, computing inverse and pseudo-inverse, and programming helper functions involving elements of Lie groups and twists</td>
</tr>
<tr>
<td>Test/Quiz</td>
<td>Multiple-choice and short-answer questions on the topics: Mathematical Background: Linear Algebra; Representing a Moving Scene</td>
</tr>
<tr>
<td>Homework</td>
<td>Calibrate a camera with unknown parameters.</td>
</tr>
<tr>
<td>Test/Quiz</td>
<td>Multiple-choice and short-answer questions on the topics: Perspective Projection and Camera Calibration, Estimating Point Correspondence, The Epipolar Constraint</td>
</tr>
<tr>
<td>Computer Labs</td>
<td>In the lab, use an unknown camera and a set of real-world objects to first calibrate the camera, take images of and then compute sparse reconstructions of these real-world objects.</td>
</tr>
<tr>
<td>Homework</td>
<td>Solve practical tasks aimed to improve multi-view reconstruction using non-linear optimization.</td>
</tr>
<tr>
<td>Test/Quiz</td>
<td>Multiple-choice and short-answer questions on the topics: Sparse Reconstruction from Two Views: Linear Algorithms, Sparse Reconstruction from Multiple Views, Bundle Adjustment &amp; Nonlinear Optimization</td>
</tr>
<tr>
<td>Homework</td>
<td>Implement a stereo rectification method.</td>
</tr>
<tr>
<td>Computer Labs</td>
<td>In the lab, use an unknown camera and a set of real-world objects to take images and compute dense reconstructions of these real-world objects.</td>
</tr>
<tr>
<td>Test/Quiz</td>
<td>Multiple-choice and short-answer questions on the topics: Dense Reconstruction from Multiple Views, Surface Reconstruction</td>
</tr>
</tbody>
</table>

### 4. Grading

<table>
<thead>
<tr>
<th>Type of Assessment</th>
<th>Pass/Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Structure</td>
<td>Activity Type</td>
</tr>
<tr>
<td></td>
<td>Attendance</td>
</tr>
<tr>
<td></td>
<td>Homework Assignments</td>
</tr>
<tr>
<td></td>
<td>Test/Quiz</td>
</tr>
<tr>
<td></td>
<td>Computer Labs</td>
</tr>
</tbody>
</table>

#### Grading Scale

<table>
<thead>
<tr>
<th>Pass:</th>
<th>75</th>
</tr>
</thead>
</table>

#### Attendance Requirements

Mandatory with Exceptions
5. Basic Information

Maximum Number of Students

<table>
<thead>
<tr>
<th></th>
<th>Maximum Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall:</td>
<td>40</td>
</tr>
<tr>
<td>Per Group (for seminars and labs):</td>
<td>20</td>
</tr>
</tbody>
</table>

Course Stream

Science, Technology and Engineering (STE)

Course Term (in context of Academic Year)

Term 3

Course Delivery Frequency

Every year

Students of Which Programs do You Recommend to Consider this Course as an Elective?

<table>
<thead>
<tr>
<th>Masters Programs</th>
<th>PhD Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Science</td>
<td>Computational and Data Science and Engineering</td>
</tr>
<tr>
<td>Information Science and Technology</td>
<td></td>
</tr>
</tbody>
</table>

Course Tags

Math
Programming
Machine Learning

6. Textbooks and Internet Resources

Required Textbooks

<table>
<thead>
<tr>
<th>Required Textbooks</th>
<th>ISBN-13 (or ISBN-10)</th>
</tr>
</thead>
</table>

Recommended Textbooks

<table>
<thead>
<tr>
<th>Recommended Textbooks</th>
<th>ISBN-13 (or ISBN-10)</th>
</tr>
</thead>
</table>
### Papers

<table>
<thead>
<tr>
<th>Papers</th>
<th>DOI or URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johannes L. Schönberger, Jan-Michael Frahm. &quot;Structure-from-Motion Revisited&quot;, CVPR 2016.</td>
<td>10.1109/CVPR.2016.445</td>
</tr>
<tr>
<td>Johannes L Schönberger, Filip Radenovi, Ondrej Chum, Jan-Michael Frahm. &quot;From Single Image Query to Detailed 3D Reconstruction&quot;, CVPR 2015</td>
<td>10.1109/CVPR.2015.7299148</td>
</tr>
</tbody>
</table>

### Web-resources (links)

<table>
<thead>
<tr>
<th>Web-resources (links)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://demuc.de/tutorials/cvpr2017/">https://demuc.de/tutorials/cvpr2017/</a></td>
<td>Large-scale 3D Modeling at CVPR 2017</td>
</tr>
<tr>
<td><a href="https://www.youtube.com/watch?v=RDkwklFgMfo&amp;list=PLTBdjV_4f-EJn6udZ34tht9EVIW7Ibeo4">https://www.youtube.com/watch?v=RDkwklFgMfo&amp;list=PLTBdjV_4f-EJn6udZ34tht9EVIW7Ibeo4</a></td>
<td>TUM Class on Multi-View Geometry (Prof. Daniel Cremers)</td>
</tr>
<tr>
<td><a href="https://github.com/sunglok/3dv_tutorial">https://github.com/sunglok/3dv_tutorial</a></td>
<td>An introductory tutorial on 3D vision</td>
</tr>
</tbody>
</table>

### 7. Facilities

#### Equipment

- DLSR cameras and objectives, calibration boards, stereo rig (we have these in our lab)

#### Software

- Unix workstations, python

### Labs for Education

- Computer Lab

### 8. Learning Outcomes

#### Knowledge

- Understand the basic notions of algebraic objects and methods involved in the geometry computer vision, including Lie groups, SVD, rank constraints, etc.
- Understand the principles of image formation in real-world cameras and how the scenes may be represented using cameras.
- Understand the mathematical foundations behind the set of sparse methods for 3D reconstruction.
- Understand the mathematical foundations behind the set of dense methods for 3D reconstruction.
- Understand the principles behind linear and nonlinear methods for multi-view stereo reconstruction and distinctions between them.
<table>
<thead>
<tr>
<th>Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use the mathematical notation to formulate geometric problems in computer vision (e.g., stereo and multi-view reconstruction/estimation, structure-from-motion or simultaneous localization and mapping, etc.)</td>
</tr>
<tr>
<td>Estimate intrinsic parameters of real-world cameras (perform camera calibration).</td>
</tr>
<tr>
<td>Use and develop software to compute sparse and dense point cloud reconstructions of real-world objects.</td>
</tr>
<tr>
<td>Use and develop software to compute surface (mesh) reconstructions of real-world objects.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience in working with real-world cameras (e.g., calibration, taking images, using cameras effectively in the geometric problems).</td>
</tr>
<tr>
<td>Experience in using, implementing and tuning of algebraic and optimization algorithms for sparse and dense point cloud reconstruction.</td>
</tr>
<tr>
<td>Experience in using and tuning of algorithms for surface (mesh) reconstruction.</td>
</tr>
<tr>
<td>Experience in formulating geometric problems in computer vision in terms of optimization tasks.</td>
</tr>
</tbody>
</table>

### 9. Assessment Criteria

#### Input or Upload Example(s) of Assignment 1:

<table>
<thead>
<tr>
<th>Select Assignment 1 Type</th>
<th>Computer Labs</th>
</tr>
</thead>
</table>
| **Input Example(s) of Assignment 1 (preferable)** | During the lab, you will be given a DSRL camera and a set of objects. Your goal during the lab is to compute a set of sparse 3D reconstructions from images obtained using the camera. To this end, you will have to:  
  — Calibrate the camera, as it has unknown intrinsic parameters,  
  — Take images, leveraging your knowledge and intuition about camera poses that should be used,  
  — Compute the reconstructions,  
  — Upload the reconstructions to server,  
  — **At home**, write a report note describing your key decisions made when performing the reconstructions, the difficulties you have encountered and how you tackled them. |

| Assessment Criteria for Assignment 1 | The lab will be assessed according to:  
  — The quantitative measures of quality for camera calibration and reconstructions,  
  — The reasoning behind the image-taking, calibration, reconstruction steps,  
  — The quality of the report explaining the key decisions made when performing the lab |

#### Input or Upload Example(s) of Assignment 2:

| Select Assignment 2 Type | Test/Quiz |
### Q4. What is a rigid-body motion?
A rigid-body motion is any transformation that preserves the volume of the object.
A rigid-body motion is any linear transformation that preserves the cross-product of any two vectors.
A rigid-body motion is any transformation that can be represented by a rotation and translation in its respective space.
**[YES]** A rigid-body motion is any transformation that preserves the norm of any vector and the cross-product of any two vectors.

### Q5. What is the generic projection matrix?
A matrix that can model any projective transformation.
A matrix that is used to estimate the projection transformation.
**[YES]** A matrix that models the standard perspective projection with respect to a normalized coordinate system.
The intrinsic parameter matrix of the camera.

### Assessment Criteria for Assignment 2
Each question is marked with 1 point if the student answers correctly.

### Input or Upload Example(s) of Assignment 3:

**Select Assignment 3 Type**
Homework Assignments

**Input Example(s) of Assignment 3 (preferable)**
Prove or disprove: There exist non-zero vectors $v_1, \ldots, v_4$ from $\mathbb{R}^0$, which are pairwise orthogonal (i.e., for all $i,j$: $(v_i, v_j) = 0$).

### Assessment Criteria for Assignment 3
The assignments will be checked and assessed for correctness.

### Input or Upload Example(s) of Assignment 4:

### Input or Upload Example(s) of Assignment 5:

### 10. Additional Notes