Master in Computer Vision Barcelona

3D recovery of Urban Scenes

Module 4: 3D Vision

Authors

Gloria Haro — gloria.haro@upf.edu
Daniel Ordóñez — daniel.ordonez@upc.edu

Coordinator: Gloria Haro

14 December 2021

Abstract

The goal of this project is to learn the basic concepts and techniques to reconstruct a real world scene given several images (points of view) of it, not necessarily previously calibrated. In this project we focus on 3D recovery of Urban Scenes using images of different datasets, namely images of facades and aerial images of cities.
This project can be useful in different applications where some 3D information has to be inferred from images taken at different points of view. Examples of such kind of applications are: image mosaics or panoramas, augmented reality, depth computation, 3D reconstruction, 3D localization and navigation, new view synthesis.

Important note: the contents of this document may be updated along the sessions. It is highly recommended to regularly check for updates on the course website.

Table of Contents

1

Project Overview

The goal of this project is to learn the basic concepts and techniques to reconstruct a real world scene given several images (points of view) of it. In this project we focus on 3D recovery of Urban Scenes using images of different datasets, namely images of facades and aerial images.
This project can be useful in different applications where some 3D information has to be inferred from images taken at different points of view. Examples of such kind of applications are: image mosaics or panoramas, augmented reality, depth computation, 3D reconstruction, 3D localization and navigation, new view synthesis.
More specificaly, the learning goals are:

- Learn the different image transformations and the basics of projective geometry.
- Learn affine and metric rectification of a single image.
- Compute the homography between two images.
- Learn camera calibration and self-calibration techniques.
- Learn the geometry of two views and how to compute the fundamental matrix.
- 3D inference from two views: triangulation and depth estimation.
- 3D reconstruction from N non calibrated cameras.

The project will be developed in PYTHON: some of the functions will be provided, some others will have to be completed and new functions will have to be implemented as well.

1.1 Project Stages

The whole task of our 3D recovery system is depicted in Figure 1.1, where the four main tasks are considered, namely:
(1) *Image transformations, projective geometry, homography estimation, image mosaics, affine and metric rectification*;
(2) *Camera calibration*;

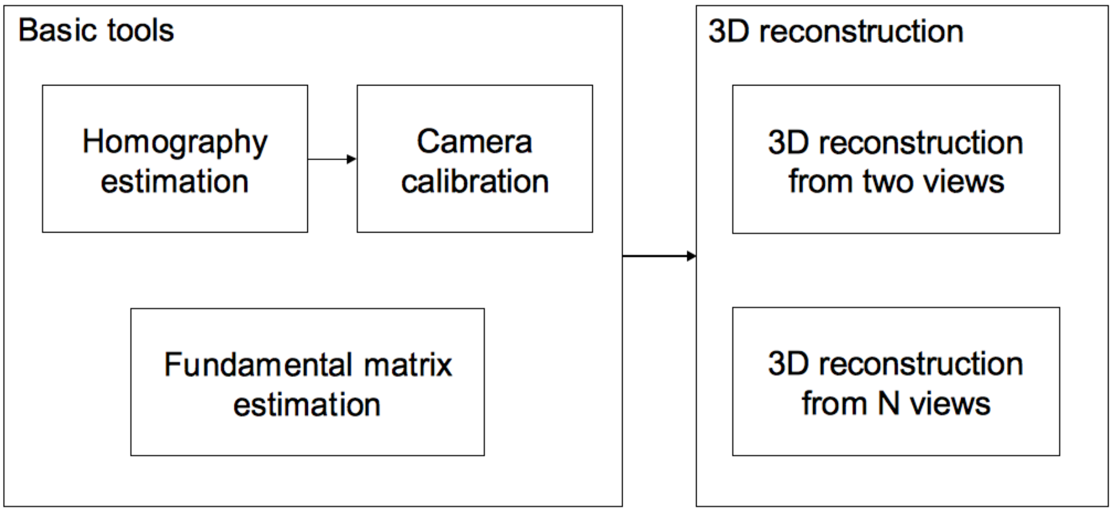

Figure 1.1: Example of project goal

(3) *The geometry of two views and 3D reconstruction or 3D inference from two views*;
(4) *3D reconstruction from N non calibrated cameras*;

1.2 Learning goals

To solve this project you will learn about:

Week 1

- How to tranform an image by a homography.
- The hierarchy of planar transformations.
- Affine rectification.
- Metric rectification.

Week 2

- Homography estimation.
- Build image mosaics.
- Camera models.
- Camera Calibration using a calibration device.
- Calibration and the image of the absolute conic.
- Augmented reality.

Week 3

- Epipolar geometry. Fundamental matrix.
- Algorithms to compute the fundamental matrix.

- Projection of the 3D world onto the image.

Week 4

- The essential matrix.
- Extract camera matrices.
- 3D reconstruction from two views: triangulation methods.
- Depth estimation.
- New view synthesis.
- Depth map fusion.

Week 5

3D reconstruction from N non calibrated cameras (Structure from motion).

- Stratified reconstruction, auto-calibration.
- Resectioning.
- Bundle adjustment.

1.3 Required lectures or concepts

- R. I. HARTLEY, A. ZISSERMAN, Multiple view geometry in computer vision, Cambridge University Press, 2000.
- D. A. FORSYTH, J. PONCE, Computer vision : a modern approach, Prentice Hall, 2003.
- O. FAUGERAS, Q.T. LUONG, The geometry of multiple images, MIT Press, 2001.

2

Assignments

The labs will be organized in 5 weeks and their contents are aligned with the theoretical lectures. More details follow below.

2.1 Week 1. Image transformations

The goal of this week is to understand the different planar transformations that can be applied to an image and how to apply them. Another aspect that will be learnt is how to remove the perspective distortion of an image, that is, how to recover parallel lines and straight angles. This is called affine and metric rectification.

2.1.1 Mandatory tasks

Implement a function that applies a given homography to a given image.
Play with the hierarchy of planar transformations.
Compute a line that joins two points.
Compute vanishing points.
Compute a transformed line.
Affine rectification of an image.
Metric rectification of an affinely rectified image.

2.1.2 Optional tasks

Metric rectification of an image in a single step.

2.1.3 Performance evaluation

To evaluate the image recitification the angles between parallel and orthogonal lines before and after the rectification will be computed.

2.1.4 Deliverable

The deliverable consists of two parts: the completed code in Python (ready to be launched on the provided images) and a short document explaining the results, problems and additional comments.

2.2 Week 2. Homography estimation. Camera calibration

The goal of this week is to implement the module of Homography estimation. As a practical application we will learn how to build an image mosaic (or panorama). Another task of this week is to perform camera calibration using several views of a planar pattern. As a rewards to this task the student could see an easy implementation of a practical application related to augmented reality. Optional tasks are homography estimation with the Gold-Standard algorithm and automatic logo insertion and logo detection in an image.

2.2.1 Mandatory tasks

Implement the DLT algorithm to estimate a homography.
Build an image mosaic.
Camera calibration using several views of a planar pattern (Zhang's algorithm). Then, we will show how to do add a simple virtual geometric object in the scene.

2.2.2 Optional tasks

Implement the Gold-Standard algorithm to estimate the homography.
Detect a logo in an image.
Insert a logo in an image using the DLT algorithm.

2.2.3 Performance evaluation

The geometric error between point correspondences will be computed for each one of the two homography estimation algorithms.

2.2.4 Deliverable

The deliverable consists of two parts: the completed code in Python (ready to be launched on the provided images) and a short document explaining the results, encountered problems, additional comments and answers to the formulated questions.

2.3 Week 3. The geometry of two views

The goal is to estimate the Fundamental matrix that relates a pair of views.

2.3.1 Mandatory tasks

Implement the 8-point algorithm.
Once the fundamental matrix is computed the students will display some epipolar lines on the images.
Apply the theoretical concepts to do photo-sequencing.

2.3.2 Optional tasks

Photo-sequencing with your own images.

2.3.3 Deliverable

The deliverable consists of two parts: the completed code in Python (ready to be launched on the provided images) and a short document explaining the results, encountered problems, additional comments and answers to the formulated questions.

2.4 Week 4. 3D reconstruction from two views

The goal of this week is the 3D reconstruction (point cloud) and the 3D inference from two views.

2.4.1 Mandatory tasks

Compute the projections matrices using the fundamental matrix computed the previous week, the calibration matrix (camera internal parameters) and a set of point correspondences between two images. It is based on computing the Essential matrix and its SVD decomposition.
Compute the 3D points by triangulation (algebraic method).
Depth map computation by a local method (SSD and NCC costs).
New view synthesis using the view morphing technique [Seitz and Dyer 1996].

2.4.2 Optional tasks

Depth map computation by a local method using bilateral weights.
Depth map computation by Belief Propagation.
Depth map computation by Plane Sweep.
Simple depth fusion by averaging truncated signed distance functions.

2.4.3 Deliverable

The deliverable consists of two parts: the completed code in Python (ready to be launched on the provided images) and a short document explaining the results, problems and additional comments.

2.5 Week 5. 3D reconstruction from N non calibrated cameras

The goal of this week is the 3D reconstruction from N non calibrated cameras (Structure from motion) with a stratified method.

2.5.1 Mandatory tasks

Projective reconstruction from two views.
Affine reconstruction locating the plane at infinity.
Metric reconstruction with the image of the absolute conic, IAC.
Bundle adjustment.
Resectioning.

2.5.2 Optional tasks

Apply the whole pipeline to a different set of images.
Projective reconstruction from more than two views.
Any other improvement you may icorporate (e.g. postprocessing of the point cloud).

2.5.3 Deliverable

The deliverable consists of two parts: the completed code in Python (ready to be launched on the provided images) and a short document explaining the results, problems and additional comments.

3

Material

3.1 Programming language

The project will be developed with Python.

3.2 Code

Part of the code will be provided and the students will have to complete some functions or write new ones. In each lab session a file called labX.ipynb (where X indicates the lab session) will be provided; this file contains the main structure of the work to be developed and serves as a guide to the lab session.

3.3 Dataset

Images to be tested in each lab session will be provided to the students at the beginning of each session. They will be extracted from the following this project we use different public datasets:

- EPFL dataset (buildings)

  `http://cvlabwww.epfl.ch/data/multiview/denseMVS.html`

- Oxford dataset (Aerial images and buildings)

  `http://www.robots.ox.ac.uk/~vgg/data/data-mview.html`

- Brown (aerial images from helicopter)

  `http://vision.lems.brown.edu/project_desc/Object-Recognition-in-Probabilistic-3D-Scenes`

- Middlebury

  `http://vision.middlebury.edu/mview/data`

Additionaly, in some specific sessions additional images will be provided so as to better pursue the pedagogical goal of each session.

4

Evaluation

The mark V for module 6 is assigned based on the evaluation of the following three parts.

Project Development (PD) The project will be developed by groups of 4 students. The project is organized in five weeks, where the last lecture (sixth week) is devoted to project presentation and final evaluation. The task of each week is detailed in section 2. As a result of the weekly development of the assignment, each group has to deliver:

Code (Cd) : a working version of the Matlab code developed along the week;

Weekly update (Wu) : a short document (10 pages), summarizing (1) used approach, (2) results, (3) problems and comments, as well as answers to questions raised in the provided lab files.

Both Cd and Wu have to be delivered to the project professor **before 9 a.m. of the day before** the monitoring lecture. **Missing this deliverable will correspond to a zero mark for that week.** Although weekly sessions serve solely to introduce the problem, create discussion and to check for possible common problems along the project development, the code Cd will be evaluated weekly (V_{Cd}), according to the criteria described in chapter 2. The PD mark will be computed as:

$$PD = \frac{1}{5}\sum_{i=1}^{5} V_{Cd_i}.$$

This will account for 80% of the final grade V.

Project Presentation (PP) The final project presentation will be evaluated during 15 minutes + 5 minutes dedicated to questions. A live demonstration of the running code will be asked for.

Each group member is expected to present part of the presentation (either half of the presentation or the questions part). The part to be presented by each member will be sorted randomly at the beginning of the session. This will account for 15% of the final grade V.

Intra-Group Evaluation (IGE) An internal evaluation is also foreseen for each group, and the contribution of each group member should be specified during the final project presentation. Additionally, as mentioned above in the project presentation details, each group member will do part of the presentation. Thus, differences in the marks achieved by each group member may be observed in case that the contribution is clearly judged as non-uniform. This will account for 5% of the final grade V.

According to the described criteria, the final score V will be assigned as:

$$V = 0.8 \cdot PD + 0.15 \cdot PP + 0.05 \cdot IGE.$$