Master in Computer Vision Barcelona

UAB UOC UPC upf

Module: **M4. 3D Vision** **Final exam**
Date: February 15, 2018
Teachers: Antonio Agudo, Coloma Ballester, Josep Ramon Casas, Gloria Haro, Javier Ruiz
Time: 2h

- Books, lecture notes, calculators, phones, etc. are not allowed.
- All sheets of paper should have your name.
- Answer each problem in a separate sheet of paper.
- All results should be demonstrated or justified.

Problem 1 *0.75 Points*

Consider an affine transformation in the 2D projective space.

(a) *(0.25p)* Write the general form of a matrix that represents it. How many degrees of freedom does it have?

(b) *(0.3p)* Show how to decompose it in fundamental transformations such as rotations, scalings and translations. Specify the proper order of these transformations.

(c) *(0.2p)* Enumerate the different invariant properties that are preserved after this type of transformation.

Problem 2 *0.75 Points*

Consider the task of removing projective distortions of flat objects on an image. To this goal, consider an image containing a plane in the 3D world.

(a) *(0.25p)* What is the vanishing line of that plane? What is a vanishing point of a world line?

(b) *(0.5p)* Explain the method of affine rectification of the image via the vanishing line.

Problem 3 *0.75 Points*

(a) *(0.25p)* What is or how is defined the plane at infinity $\mathbf{\Pi}_\infty$ in the 3D projective space $\mathbb{P}^3$?

(b) *(0.5p)* What is the general form of a finite projective camera matrix P? Describe its internal and external parameters.

Problem 4 *1.25 Points*

Calibration with a planar pattern.

(a) *(0.25p)* Define the image of the absolute conic. How many degrees of freedom does it have?

(b) *(0.6p)* Derive the equations that relate the image of the absolute conic and the columns of the homography that relates the planar pattern with an image of it taken from a certain point of view.

(c) *(0.2p)* Assume the camera parameters are completely unkonwn. How many views of the planar pattern do we need to calibrate the camera? Justify your answer.

(d) *(0.2p)* Under which hypothesis is possible to calibrate the camera just by using a single view. Justify your answer.

Problem 5 *2 Points*

Consider two images I and I' taken by the same camera (with intrinsic matrix $K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$) where the motion between them consists in a rotation along the Y axes (the axes defined by the optical center of the camera) and a translation of d_x meters along the X' axes (see corresponding Figure). Remember that a 3D point in the world P (expressed in camera O_c coordinate system) corresponds to point $P' = RP + T'$ (expressed in camera O'_c coordinate system). Answer the following questions:

a) Compute the corresponding translation vector T'.

b) Compute the corresponding rotation matrix R.

c) Compute the essential matrix E.

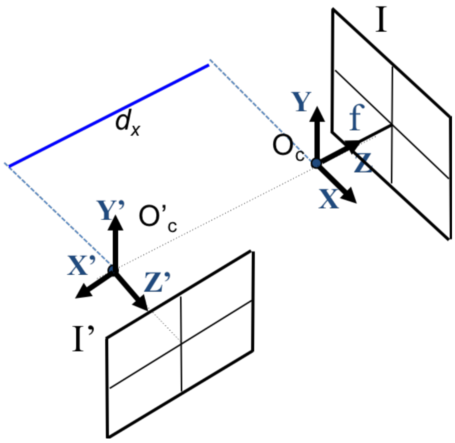
d) Compute the fundamental matrix F.

e) Where is the epipole e in image I?

f) Where is the epipole e' in image I'?

g) What is the main difference between the essential matrix E and fundamental matrix F?

f) Is it possible to recover R and T' from E? with any restriction?



Problem 6 *1.5 Points*

Disparity estimation with local methods.

(a) *(0.25p)* Describe the main steps for estimating the disparity of a pair of stereo rectified images with a local method.

(b) *(0.5p)* Define the Sum of Squares Differences, Sum of Absolute Differences, Normalized Cross Correlation and comment their advantages/disadvantages.

(c) *(0.25p)* Consider a window with uniform weights. Describe the positive and negative effects of the window size.

(d) *(0.25p)* Write the expression of the bilateral weights and comment its benefits.

(e) *(0.25p)* Describe different possible failure cases when estimating the disparity through local methods and two views.

Problem 7 *0.4 Points*

Formulate the projection equation (indicating the corresponding size of every matrix) in terms of a measurement matrix, assuming: 1) a perspective camera, 2) an orthographic camera. To compute shape and motion by rigid factorization, which rank do we have to enforce in every case and how can this be done?

Problem 8 *1.1 Points*

Let us assume a collection of I image frames with extrinsic parameters $\mathbf{P}_i$ with $i = \{1, \ldots, I\}$, where a 3D rigid object composed of P points is observed. Due to lack of visibility and outliers, a few points are not viewed in some frames. Particularly, the corresponding visibility vectors contain 16, 10, 18, and 18 components for every image, respectively. Assuming $P = 18$, for this particular case we always observe the points with smaller indexes $p = \{1, \ldots, P\}$. We want to simultaneously estimate 3D shape $\mathbf{X} \equiv [\mathbf{x}_1, \ldots, \mathbf{x}_P]$ (every $\mathbf{x}_p$ contains the 3D coordinates of the p-th point) and motion solely from 2D annotations by sparse bundle adjustment. For this toy example, represent the corresponding structure of the Jacobian matrix to code the problem and indicate the final matrix size. The intrinsic parameters of the camera can be assumed to be known. (0.8 points)

If the four image frames are a part of a monocular video, could we impose more constraints to sort out the problem? If so, describe them and represent this type of priors in the previous pattern. Could the designed pattern be used to handle continuous non-rigid objects? Explain why, and provide some ideas to modify it if needed. (0.3 points)

Problem 9 *0.5 Points*

We set out with four questions for the 3D sensors/ 3D processing part of this module. Can you provide a short answer for each one?

1) Is "projective vision" a natural way to capture the 3D world?
2) Do we need photometry to get geometry?
3) Does 3D vision mean the same than 3D geometry?
4) Does 2D/3D matter for "Teaching computers to see"?

Problem 10 *1.0 Points*

1) Point cloud data can be "organized" or "unorganized". Define both types.

2) Which type of point cloud data can be straightforwardly converted into RGBD and how? Which type has a disadvantage for processing purposes and why?

3) A convenient feature extracted from pointcloud data should be able to capture the same local surface characteristics in the presence three main variations: transformations, data density and noise. Describe these three variations and state why pointcloud features should be invariant to those

4) Mention two point feature representations and describe what do they aim to capture from pointcloud data characteristics