Module: **M4. 3D Vision** **Final exam**
Date: March 3, 2022
Teachers: Antonio Agudo, Josep Ramon Casas, Gloria Haro, Javier Ruiz, Federico Sukno
Time: 2h

Problem 1 *1.8 Points*

(a) Consider two lines expressed in homogeneous coordinates, demonstrate that if they are parallel, then they intersect at a point at infinity.

(b) Consider two lines which we know are parallel in the real World[1]. When we image them, we obtain lines $\mathbf{l} = (4, 2, 1)^T$ and $\mathbf{m} = (5, 2, 2)^T$.

(i) Can you tell what type of planar transformation there is between the real World and the image we obtained? Justify.

(ii) Based on the knowledge that the two lines $\mathbf{l}$ and $\mathbf{m}$ are parallel in the real World, would you be able to perform metric-rectification to the obtained image? Clearly justify your answer.

(c) Consider the following planar transformation matrix:

$$\mathbf{H} = \begin{bmatrix} 0 & -4 & 5 \\ 4 & 0 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

(i) Indicate the type of the above planar transformation.

(ii) Enumerate at least 3 invariants for this type of transformation

(iii) Describe with as much detail as possible what will be the effect of transformation $\mathbf{H}$ (e.g. if it were a rotation, by what angle, or if it will translate points, by how much, etc).

Problem 2 *0.9 Points*

Consider two images that correspond to different views of the same planar object. We wish to relate the two images by some planar transformation $\mathbf{H}$

(a) How many degrees of freedom will $\mathbf{H}$ have?

(b) Given a pair of corresponding points $\mathbf{x}$ and $\mathbf{x}'$ (in the first and second image, respectively), how many independent equations can we derive from them (to solve for $\mathbf{H}$)? Write down those equations.

(c) How many correspondence pairs $\{\mathbf{x}; \mathbf{x}'\}$ we need in order to fully determine $\mathbf{H}$? Is there any constraint regarding their selection (i.e. beyond the fact that there cannot be repeated points)?

[1] These lines have **no relation** to the lines from the previous item

Problem 3 *1.7 Points*

Camera calibration and pose estimation.

(a) Describe the camera projection matrix and the different elements that form it.

(b) Given an estimation of the camera projection matrix, how would you estimate the internal parameters from it by assuming that the camera has zero skew?

(c) Once you have the internal parameters, how would you estimate the external ones?

(d) In Zhang's algorithm for calibration we need to estimate homographies that relate a template with the different views of it taken by the camera. Consider a generic view of the template, derive the equations that relate the elements of the projection matrix with the columns of the homography associated to that view.

(e) What is the minimum amount of views we need in the Zhang's algorithm. Justify your answer.

(f) Describe the pose estimation problem. How many parameteres do we need to estimate? Which kind of data do the classical methods (non learning-based) use to solve that problem?

Problem 4 *1.8 Points*

Consider the two images I (left) and I' (right) with non-coincident optical centers. Answer the following questions:

a) What is the rank of the corresponding fundamental matrix F between the two images?

b) How many degrees of freedom does the fundamental matrix F have?

c) All corresponding points p and p' in the two images must satisfy the epipolar constraint. What is the mathematical expression for the epipolar constraint?

d) What is the mathematical expression in terms of F for the epipolar line l' in image I' (right) for the point p in image I (left)?

e) What is the mathematical expression in terms of F for the epipolar line l in image I (left) for the points p' in image I' (right)?

f) What is the mathematical relationship between the epipole e on image I (left) and F?

g) What is the mathematical relationship between the epipole e' on image I' (right) and F?

h) Why is the epipolar constraint useful for stereo matching (i.e. finding corresponding points between left and right images)?

i) What happens to the epipoles and the epipolar lines in the rectified images after applying image rectification?

j) What are the advantages of applying image rectification before we do stereo matching?

Problem 5 *1 Point*

Triangulation and depth.

(a) We want to solve the triangulation problem with the geometric method, which is the minimization problem we need to solve? Define all the involved variables.

(b) Define the concept of signed distance function in the context of 3D reconstruction.

(c) How can we obtain the signed distance function of the surface associated to a certain depth map?

(d) Explain the main idea of estimating a 3D reconstruction by depth map fusion.

Problem 6 *1.4 Points*

Consider the structure-from-motion problem: 1) Explain briefly the alternatives that you know to solve the rigid case in terms of processing and algorithms. Provide the most important characteristics of every alternative. 2) Let us consider a toy problem composed of 50 points (full visibility) of a deformable object and 10 images, as well as a low-rank linear subspace model of rank 4 to encode the shape deformations. Obtain the number of unknowns and the equations to solve the problem by bundle adjustment. To do that, you only need to consider the data term.

Problem 7 *1.0 Points*

Range sensors and 3D data.

(a) Describe in a few lines these two technologies for range imaging: structured light (SL) and time-of-flight (ToF).

(b) Mention one aspect that SL and ToF have in common and one main difference.

(c) Why is data acquired with range sensors sparse? Provide an example of dense (non-sparse) 3D data (not acquired with range sensors)

Problem 8 *0.4 Points*

In pointcloud processing, organized data structures may help in vision analysis tasks. Graphs and trees have been proposed to analyze (segment, detect, classify) pointclouds.

a) Explain how adjacency, hierachy, captured primitives (visual or geometric features) and visual boundaries play a role for analysis tasks within the edge and node elements of graphs and trees.

b) What is the reason why graph and tree structures, specially for point clouds, are expected to perform better than raw data?