FACULTÉ DES SCIENCES ET TECHNIQUES SAINT-ÉTIENNE

Master MLDM, AIMA and 3DMT - Computer Vision course

Exam March 2019 - 2h without documents

Part 2 *(4 points – 20 min)*: **Short answers**

Question 5 *(1 point)***:** Express the point where 2-D parallel lines intersect in homogeneous coordinates (assume any notation).

Question 6 *(1 point)***:** How many degrees of freedom are there in perspective projection transformation between two images?

Question 7 *(1 point)***:** Considering that:

$$\mathbf{E}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{M}\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

What are the coefficient of the matrix M in case of orthographic projection?

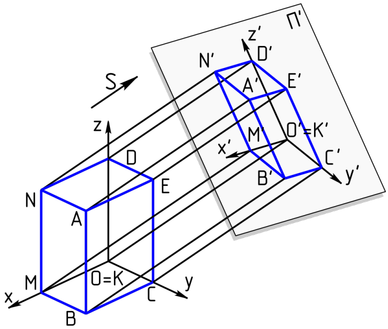
Hint: see Annex 1

Question 8 *(1 point)***:** A projective transformation between two images X and X' can be modelled by transforming each point p in one image to a point p' in a new image using p'= H p, where H is the homography. Briefly explain, why we can't do the same thing with the Fundamental matrix F to transform each point p in the left image to a point p' in the right image from the transform p' = Fp?

Hint: see Annex 2

From Wikipedia:

Orthographic projection (sometimes **orthogonal projection**) is a means of representing three-dimensional objects in two dimensions. It is a form of parallel projection, in which all the projection lines are orthogonal to the projection plane, resulting in every plane of the scene appearing in affine transformation on the viewing surface. The obverse of an orthographic projection is an oblique projection, which is a parallel projection in which the projection lines are *not* orthogonal to the projection plane.



A simple orthographic projection P onto the plane $z = 0$ can be defined by the following matrix:

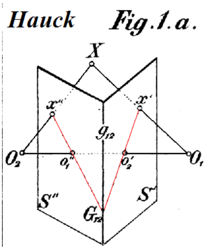
$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

For each point $v = (v_x, v_y, v_z)$, the transformed point would be:

$$Pv = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \\ 0 \end{bmatrix}$$

From Wikipedia:

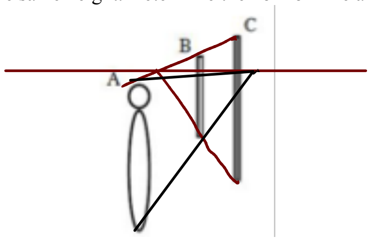
In the field of computer vision, any two images of the same planar surface in space are related by a **homography** (assuming a pinhole camera model). This has many practical applications, such as image rectification, image registration, or computation of camera motion—rotation and translation—between two images. Once camera rotation and translation have been extracted from an estimated homography matrix, this information may be used for navigation, or to insert models of 3D objects into an image or video, so that they are rendered with the correct perspective and appear to have been part of the original scene.



Geometrical setup for homography: stereo cameras O_1 and O_2 both pointed at *X* in epipolar geometry.

Part 4 *(5 points – 20 min)*: **Perspective**

Question 11 *(2 points)*: Consider the image of person A standing on the ground plane and two vertical utility poles B and C on the same ground plane (see image below) as taken by a perspective camera of focal length f. Consider that pole B and C have the same height. Determine the horizon line and draw it on the image after.



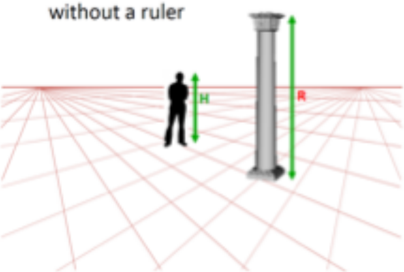
Question 12 *(2 points)*: From the image, is the person A taller than the pole B? Why?

Hint: see Annex 3

Annex 3

From http://www.cs.cornell.edu/courses/cs4670/2015sp/lectures/lec19_svm2_web.pdf

Measuring height without a ruler



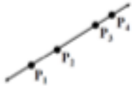
Measuring height without a ruler

Compute Z from image measurements
Actually get a scaled version of z

The cross ratio

- A Projective Invariant
  - Something that does not change under projective transformations (including perspective projection)

The *cross-ratio* of 4 collinear points



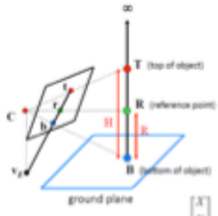
$$\frac{\|P_3 - P_1\|\,\|P_4 - P_2\|}{\|P_3 - P_2\|\,\|P_4 - P_1\|} \qquad P_i = \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

Can permute the point ordering $\quad \dfrac{\|P_1 - P_3\|\,\|P_4 - P_2\|}{\|P_1 - P_2\|\,\|P_4 - P_3\|}$

- 4! = 24 different orders (but only 6 distinct values)

This is the fundamental invariant of projective geometry

Measuring height



$$\frac{\|T - B\|\,\|\infty - R\|}{\|R - B\|\,\|\infty - T\|} = \frac{H}{R}$$

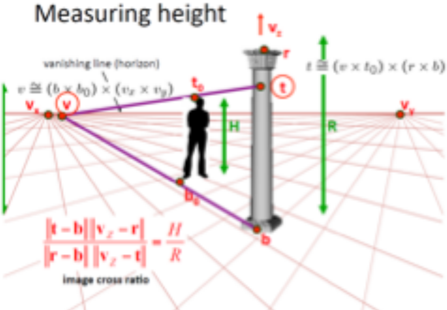
scene cross ratio

$$\frac{\|t - b\|\,\|v_Z - r\|}{\|r - b\|\,\|v_Z - t\|} = \frac{H}{R}$$

image cross ratio

scene points represented as $\quad P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad$ image points as $\quad p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

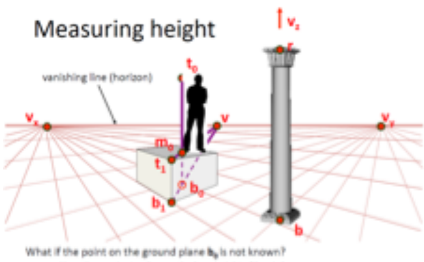
Measuring height



$$\frac{\|t - b\|\,\|v_Z - r\|}{\|r - b\|\,\|v_Z - t\|} = \frac{H}{R}$$

image cross ratio

Measuring height



What if the point on the ground plane b_0 is not known?

- Here the guy is standing on the box, height of box is known
- Use one side of the box to help find b_0 as shown above

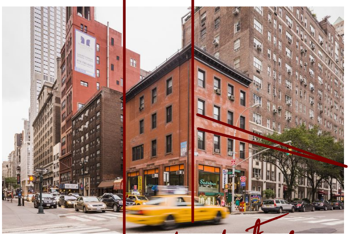
Exam March 2020 - 2h without documents

Part 2 (20 mns)**:** pinhole model

Question 1*(1 point)***:** What is the name of the mathematical equation of a thin lens which models the relationship between the focal f, the distance Z of a point in the scene to the camera frame and the distance z of its 2 projection in the image frame? Write this equation.

Question 2 *(2 points)*: With the pinhole model the most distant objects appear smaller, why? With the pinhole model parallel lines in the scene can intersect in the image, show with a graphic why?

Question 3 *(3 points)***:** For the two images shown below estimate manually the positions of the three major vanishing points, corresponding to the building orientations. Do these vanishing points are located within the image plane? If yes, can we state that we have a weak or a strong perspective?



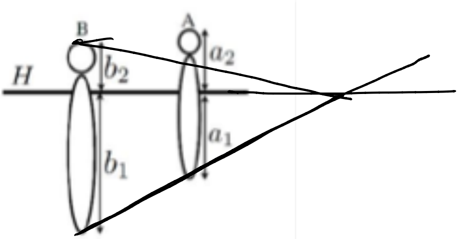
weak perspective

strong perspective.

Exam March 2019 - 2h without documents

Part 1 *(5 points + 4 bonus points – 30 min)***: Comparing heights**

Consider the image of person A and person B standing on the ground plane (see figure below), as taken by a perspective camera of focal length f. H is the horizon line (y=0). a_1 is the distance between the A's toes and the horizon, a_2 is the distance between A's head and the horizon in units of pixels. Similarly for person B. h is the tall of person A. Distances a_1, a_2, b_1, b_2, are expressed in the camera plane. Suppose A's height is h feet.



Question 1 *(3 points)*: From this Figure, who is taller? Briefly explain why?

Question 2 *(2 points)*: How many feet above the ground is the camera? Give the equation which relies the height h_c and the following parameters: a_1, a_2 and h. Briefly explain why?

Hint: see Annex

$$h_c = \frac{a_1}{a_1+a_2} \times h \quad , \quad h_c = \frac{b_1}{b_1+b_2} \times h_b$$

$$\Rightarrow h_b = \left(\frac{b_1+b_2}{b_1}\right) h_c = \left(\frac{b_1+b_2}{b_1}\right) \times h \times \left[\frac{a_1}{a_1+a_2}\right]$$

Question 3 *(bonus question, 2 points)*: How tall is person B (in feet)? Give the equation which relies the height h_b and the following parameters: a_1, a_2, b_1, b_2, and h. Briefly explain why?

Question 4 *(bonus question, 2 points)*: What is the distance (along the z-axis) from the camera to person B (in feet)? Give the equation which relies the height of person B and the following parameters: a_1, a_2, b_1, f, and h. Briefly explain why?

$$x = f\frac{X}{Z}$$

$$\Rightarrow b_1 + b_2 = f \times \frac{h_b}{Z}$$

$$\Rightarrow Z = \frac{f h_b}{b_1+b_2} = \frac{f h a_1}{b_1(a_1+a_2)}$$

Part 3 *(4 points)*: **Projective, Affine, Similarity, and Isometric Transformations**

Question 1 *(2 points)*: Classify each of the following transformations. Fill-in the circle corresponding to the most specific classification.

i. $H_1 = \begin{bmatrix} 3/4 & -1 & 0 \\ 1 & 3/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective ○ Affine ○ **Similarity** ○ Isometric

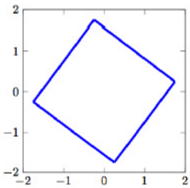
ii. $H_2 = \begin{bmatrix} 3/5 & -4/5 & 1 \\ 4/5 & 3/5 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective ○ Affine ○ Similarity ○ **Isometric**

iii. $H_3 = \begin{bmatrix} 3/16 & -1 & -1/4 \\ 1/4 & 3/4 & 1/2 \\ 1/4 & 1/4 & 1 \end{bmatrix}$ ○ **Projective** ○ Affine ○ Similarity ○ Isometric

iv. $H_4 = \begin{bmatrix} 3/8 & -5/8 & 0 \\ 1/2 & 5/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective ○ **Affine** ○ Similarity ○ Isometric

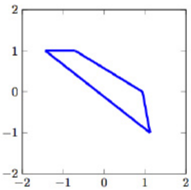
Question 2 *(2 points)*: The figures below shows the outputs of applying one of transformations (projective, affine, similarity, and isometric) to a square with vertices at (1,1), (1,-1), (-1,-1), (-1,1). Fill in the circle corresponding to the most specific transformation used to generate each output.

i.



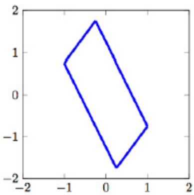
○ Projective ○ Affine ○ **Similarity** ○ Isometric

ii.



○ **Projective** ○ Affine ○ Similarity ○ Isometric

iii.



○ Projective ○ **Affine** ○ Similarity ○ Isometric

Part 4 *(4 points)*: **Camera model and calibration**

Suppose we want to solve for the camera matrix K and that we know that the world coordinate system is the same as the camera coordinate system. Assume the matrix K has the structure outlined below.

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} & K_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{pmatrix}$$

Note that K33 is an unknown. Assume that we are given n correspondences. Each correspondence consists of a world point (X_i; Y_i; Z_i) and its projection (u_i; v_i) for i = 1; ; n.

Question 1 *(1 point)*: What is the minimum number of correspondences needed to solve for the unknowns in the matrix K?

Question 2 *(3 points)*: Set up an equation of the form Ax = 0 to solve for the unknowns in K (where A is a matrix, and x and 0 are vectors). Be specific about what the matrix A and vector x are.

Question 3 *(1 point)*: Explain how to solve for the unknowns in the camera matrix K. Make sure K33 = 1.

1 6 unknowns $\rightarrow$ 6 equations

(5) The equations of two parallel lines are

$$Ax + By + C = 0$$
$$Ax + By + D = 0$$

In homogeneous co-ordinate,

$$(x, y) \longrightarrow (X, Y, W)$$

where $x = \frac{X}{W}$

$y = \frac{Y}{W}$

$$AX + BY + CW = 0$$
$$AX + BY + DW = 0$$

$\therefore (C - D)W = 0 \Rightarrow W = 0$

(6)
$$P = \begin{pmatrix} a_{11} & a_{12} & \sigma_x \\ a_{21} & a_{22} & \sigma_y \\ v_1 & v_2 & v \end{pmatrix}$$

dof = 8

(7)
$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(8)

(11) $\frac{1}{z} + \frac{1}{z} = \frac{1}{f}$

(Fundamental equation of thin lens)

(12) $x = f\frac{X}{Z}$

The length of the object is inversely proportional to the relative distance of the object from the camera.