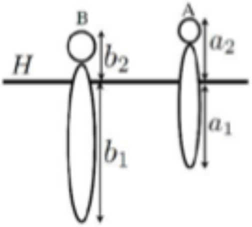
Master MLDM and 3DMT - Computer Vision course

Correction Exam March 2019 - 2h without documents

(6 parts with a total of 15 questions accounting for 25 points (+ 4 bonus points), the exam will be scored for 20 points)

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

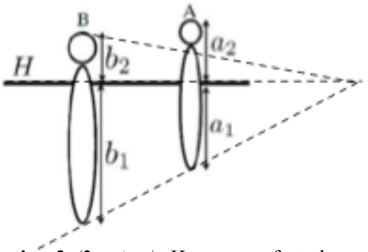
Consider the image of person A and person B standing on the ground plane (see figure below), as taken by a perspective camera of focal length f. H is the horizon line (y=0). a_1 is the distance between the A's toes and the horizon, a_2 is the distance between A's head and the horizon in units of pixels. 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?

Answer:

Person A is taller. Person A is standing behind person B, since you can draw a line from person B's feet to person A's feet and extend to the horizon. If they were the same height, then their feet and their heads would lie on parallel planes. In that case, you could draw a line from person B's head to person A's head to the same point on the horizon. However, if you draw a line from person B's head to person A's head, it actually goes up and doesn't even meet the horizon. Thus, person A is taller.



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?

Answer: see after question 4.

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?

Answer: see after question 4.

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?

Answer:

How many feet above the ground is the camera?	h_c = h * a1 / (a1+a2) The horizon line is straight ahead of the camera. The height of the camera is equal to the percentage of person A which is above the horizon times his height.
How tall is person B (in feet)?	h_b = h_c * (b1+b2)/b1 = h * a1/(a1+a2) * (b1+b2)/b1 We know portion b1 corresponding to camera height h_c. We then add on the portion of Person B's height which is above the horizon line.
Distance (along the z-axis) from the camera to person B (in feet)?	b1+b2 = f * h_b / Z Z = f/(b1+b2) * h_b = f*h/b1 * a1/(a1+a2) Height of person B projects onto b1+b2 in the image plane. Use projection equation and then solve for Z.

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).

Answer: (x, y, 0)

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

Answer:

There are 8 degrees of freedom in a perspective projection transform between two images. Given point p = (x,y,1) and p' = (x,y,1) in a second image, we would like to solve for P' = H*p, where p' is P' normalized over the last coordinate. Matrix H has 9 values, but one must is fixed to prevent arbitrary scaling.

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?

Answer:

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

See https://en.wikipedia.org/wiki/Orthographic_projection

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?

Answer:

There is an ambiguity in this case, since a point in one image maps to the *epipolar line* in the other image.

Part 3 *(2 points – 10 min)*: **Estimating depth**

Question 9 *(2 points)*: Assume that humans can discriminate disparity difference of 6". Further assume that the distance between the eye (b) is 10 cm. What is the depth resolution (d) that humans can discriminate, when an object is fixated at a depth (Z) of:

1. 10 cm (in this case Z » b; tan(F_1) = b/2Z and tan(F_2) = b/2(Z+d))

2. 1 m (in cases 2 to 4, we can assume that Z > > b; $\tan(F_1 - F_2) = bd/2Z^2$)
3. 10 m
4. 100 m

Answer:

For part (1) $Z \approx b$. Use, $tan(\theta_1) = \frac{b}{2Z}$ and $tan(\theta_2) = \frac{b}{2(Z+d)}$. This gives, d = 0.0073mm.
For parts (2-4) we can assume that $Z >> b$. For this, we have, $tan(\theta_1 - \theta_2) = \frac{bd}{2Z^2}$. Making appropriate approximations, we get d = $\frac{\pi Z^2}{54} 10^{-4}$ and substitute Z in cms
(2) 0.0582 cms
(3) 5.82 cms
(4) 582.8 cms

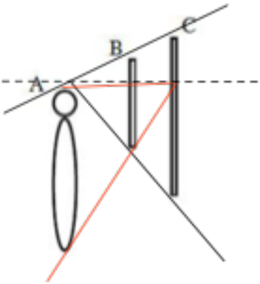
Part 4 *(5 points – 15 min)*: **Perspective**

Question 10 *(1 point)*: Which scenario is most suitable for scale orthographic camera model?

1. Describe the movement of a car in front of another car.
2. Image formation using a laptop.
3. Describe the movement of a plane at very high altitude.

Answer: third answer

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 below.



Answer:

One vanishing point is determined by the two lines passing the top and bottom of B and C. The poles are vertical, thus the horizon line should pass the vanishing point and be orthogonal to the poles.

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

Answer:

Pole B is taller. You can see this geometrically in the figure by drawing the line l_1 joining the feet of A and B, and finding the vanishing point for this line (where it meets the horizon). From this point you can draw a line l_2 passing through the head of A. l_2 pass through the middle of B, hence B is taller.

Part 5 *(3 points – 10 min)*: **Stereo correspondence**

Question 13 *(3 points)*: A stereo system is used on a flying vehicle. It consist of two ccd cameras, each having 512x512 detectors on a 1 cm^2. The lenses used have 16 mm focal lengths (f) (and the focus is fixed at the infinity). For corresponding points (u_1, v_1) in the left image and (u_2, v_2) in the right image, $v_1=v_2$ as the x-axes in the two image planes are parallel to the epipolar lines. The vergence is adjusted so that disparity u_2-u_1 of points at infinity is zero in the two images derived from the cameras. The baseline (b) between the cameras is 1 foot. If the nearest range to be measured is 16 feet (Z), what is the largest disparity that will occur (in pixels)? What is the range resolution, due to the pixel spacing, at 16 feet? What range corresponds to the disparity of one pixel?

$1\,px = \frac{1}{512}\,cm$

Answer :

$d = \frac{bf}{Z} \Rightarrow Z = \frac{bf}{d}$

$= \frac{1 \times 30.48 \times 16 \times 0.1}{\frac{1}{512}}$

$= 819.2\,ft$

(a) 52 px. (The exact answer is 51.2, use $d = \frac{bf}{Z}$).

(b) $\frac{5}{16}$ feet. (Use $\frac{bf}{Z^2}\delta Z = \delta d$).

(c) 819.2 feet. (use, $Z = \frac{bf}{d}$)

Part 6 *(6 points – 20 min)*: **Calibration**

We seek to calibrate a rigid body camera with a planar calibration object. The object possesses M distinct features. The location of those features in the object coordinate frame are unknown (and they are not arranged in a checkerboard). Let be K the number of images acquired and M the number of features on the planar calibration object.

In answering the questions below, you might want to use the following notation/equations

$$\begin{pmatrix} \tilde{X}^C \\ \tilde{Y}^C \\ \tilde{Z}^C \end{pmatrix} = R \begin{pmatrix} X^W \\ Y^W \\ Z^W \end{pmatrix} + \begin{pmatrix} T_X \\ T_Y \\ T_Z \end{pmatrix}$$

$$R = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix}$$

$$\begin{pmatrix} x_{\text{im}} \\ y_{\text{im}} \end{pmatrix} = \left\lfloor \begin{pmatrix} \frac{f}{s_x} \frac{\tilde{X}^C}{\tilde{Z}^C} + o_x + 0.5 \\ \frac{f}{s_y} \frac{\tilde{Y}^C}{\tilde{Z}^C} + o_y + 0.5 \end{pmatrix} \right\rfloor = \begin{pmatrix} \frac{f}{s_x}\frac{1}{Z^c} & 0 & 0 & o_x+0.5 \\ 0 & \frac{f}{s_y}\frac{1}{Z^c} & 0 & o_y+0.5 \end{pmatrix} \begin{pmatrix} \tilde{X}^c \\ \tilde{Y}^c \\ \tilde{Z}^c \\ 1 \end{pmatrix}$$

Question 14 *(2 points)***:** What are the free (i.e. intrinsic and extrinsic) parameters that can be recovered through the calibration process? Compute the number of unknown parameters defined above.

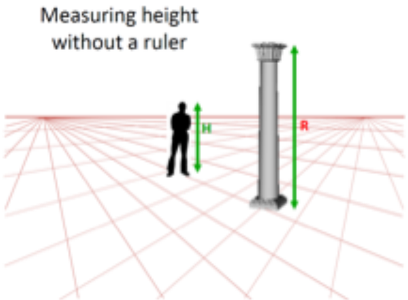
(f, s_x, o_x, s_y, o_y)

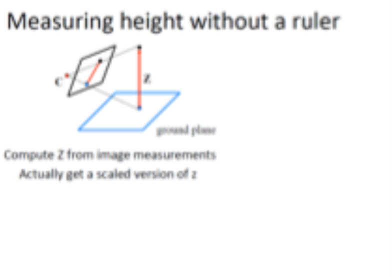
Answer: see similar correction in 2016 exam. (ϕ, φ, ψ)

Question 15 *(4 points)***:** Let us assume that there are 2M parameters for the unknown feature locations $\{X^W_j; Y^W_j\}$ on the planar calibration object and that $Z^W_j = 0$ for all features j. So, we have to add the object parameters $\{X^W_j; Y^W_j\}$ to the set of free parameters. How many parameters have to be estimated to calibrate the camera from this calibration object? How many constraints are being provided by each image of the calibration pattern? What are the lower bounds for M and K? Provide exact formulae, one of the form K 3… and one of the form M 3… For K = 4 images, what is the minimum M? For M = 4 features, what is the minimum K?

Answer: see similar correction in 2016 exam.

Annex *(from http://www.cs.cornell.edu/courses/cs4670/2015sp/lectures/lec19_svm2_web.pdf)*

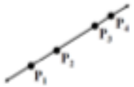




The cross ratio

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

The *cross-ratio* of 4 collinear points



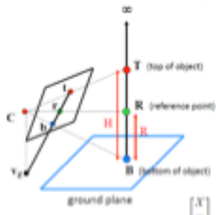
$$\frac{\|\mathbf{P}_3-\mathbf{P}_1\|\,\|\mathbf{P}_4-\mathbf{P}_2\|}{\|\mathbf{P}_3-\mathbf{P}_2\|\,\|\mathbf{P}_4-\mathbf{P}_1\|} \qquad \mathbf{P}_i = \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

Can permute the point ordering
$$\frac{\|\mathbf{P}_1-\mathbf{P}_3\|\,\|\mathbf{P}_4-\mathbf{P}_2\|}{\|\mathbf{P}_1-\mathbf{P}_2\|\,\|\mathbf{P}_4-\mathbf{P}_3\|}$$

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

This is the fundamental invariant of projective geometry

Measuring height



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

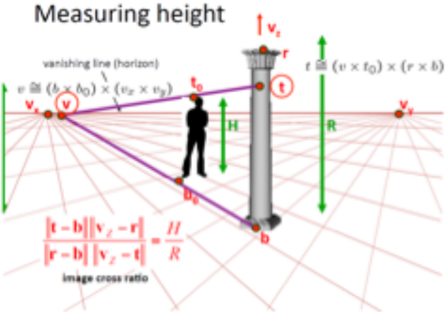
scene cross ratio

$$\frac{\|\mathbf{t}-\mathbf{b}\|\,\|\mathbf{v}_Z-\mathbf{r}\|}{\|\mathbf{r}-\mathbf{b}\|\,\|\mathbf{v}_Z-\mathbf{t}\|} = \frac{H}{R}$$

image cross ratio

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

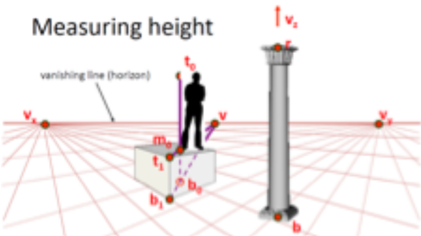
Measuring height



$$\frac{\|\mathbf{t}-\mathbf{b}\|\,\|\mathbf{v}_Z-\mathbf{r}\|}{\|\mathbf{r}-\mathbf{b}\|\,\|\mathbf{v}_Z-\mathbf{t}\|} = \frac{H}{R}$$

image cross ratio

Measuring height



What if the point on the ground plane $\mathbf{b_0}$ 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 $\mathbf{b_0}$ as shown above