Midterm Exam   Fundamentals of Computer Vision

COMP 558   Oct. 13, 2010   Prof. M. Langer
There are a total of 11 points (10 + 1 bonus on Question 4).

1. **(1 point)**

   Consider a thin lens camera with f-number $N = 4$ and focal length 30 mm. Suppose the camera is focused at $Z = 500$ mm. Give an expression for the blur width in mm of a point at $Z = 480$ mm. You must substitute in numerical values for all known quantities in your expression.

2. **(3 points)**

   (a) What is the maximum number of vanishing points that are defined for any scene? Why?

   (b) Give a geometric interpretation of each of the four columns of a 3×4 projection matrix $\mathbf{P}$, namely what scene and/or image properties do each of these columns correspond to?

3. **(3 points)**

   (a) What is the angular width in radians of a typical human head that is seen at a distance of 2 m ? Assume a typical head is a sphere of diameter 10 cm.

   (b) How many pixels does the answer in (a) correspond to? Assume a pinhole camera with $f = 40$ mm, and a sensor with 100 pixels per mm.

   (c) What is the solid angle in steradians of the head in (a) ?

4. **(originally 1 point, but increased to 2 points)**

   Consider a spherical light source of radius r at a height h above a ground plane. Assume $h \gg r$. Give an expression for the irradiance of the ground plane at position (X, Y) that is due to this light source. You may assume the radiance L_{src} of the source is constant in all directions. Your expression must be in terms of the source radiance and the variables X, Y, h, r.

5. **(1 point)**

   Suppose a photographer wishes to increase the range of depths in a scene that are approximately in focus, without changing the field of view angle or the exposure. How can the photographer achieve this?

6. **(1 point)** Let

$$D(x) = \begin{cases} -1, & x = 1 \\ 1, & x = -1 \\ 0, & otherwise \end{cases}$$

$$B(x) = \begin{cases} 1, & -1 \leq x \leq 1 \\ 0, & otherwise \end{cases}$$

   What is $B(x) * D(x)$?

$$(x, y) \;=\; (\frac{X_0}{Z_0}\; f, \frac{Y_0}{Z_0}\; f). \tag{1}$$

$$\mathbf{P} = \mathbf{KR}[\ \mathbf{I}\ |\ \ -\mathbf{C}] \tag{2}$$

$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & p_x \\ 0 & \alpha_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$\frac{1}{Z_o} + \frac{1}{Z_i} = \frac{1}{f} \tag{4}$$

$$\text{f number}, N \equiv \frac{f}{A} \tag{5}$$

$$\Delta X_i = A \mid Z_s \ (\ \frac{1}{f} - \frac{1}{Z_o} \ ) \ -1 \mid \tag{6}$$

$$\text{radiance}, \quad L(\mathbf{X}, \mathbf{l}) \equiv \frac{\text{power}}{\text{cross section area } * \text{ solid angle}} \tag{7}$$

$$\text{irradiance}, \quad E(\mathbf{x}) = \int L(\mathbf{x}, \mathbf{l}_{in})\ \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in}\ d\Omega_{in} \tag{8}$$

$$\text{BRDF} \equiv f(\mathbf{x}, \mathbf{l}_{in}, \mathbf{l}_{out}) \equiv \frac{d\ L(\mathbf{l}_{out})}{L(\mathbf{x}, \mathbf{l})\ \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}\ d\Omega_{in}} \tag{9}$$

$$\text{image irradiance }, \ E(\mathbf{x}) = (\frac{\pi A^2}{4}) \frac{L(\mathbf{x}) \cos^4 \alpha}{Z_s^2} \tag{10}$$

$$E_{R,G,B}(\mathbf{x}) = \int C_{R,G,B}(\lambda)\ E(\mathbf{x}, \lambda) d\lambda \tag{11}$$

$$f(x) * I(x) = \sum_{x'=-\infty}^{\infty} \ f(x')\ I(x - x') \tag{12}$$

Solutions

1.

$$\Delta X_i = \frac{30}{4} \mid (\frac{1}{30} - \frac{1}{500})^{-1} (\frac{1}{30} - \frac{1}{480}) - 1 \mid$$

Marking scheme: 0.5 points for applying eq. 5 to get A. 0.5 points for applying the thin lens formula to get Z_s in eq. 6.

2.

(a) The answer is "infinite", i.e. there is no bound on the number of directions of parallel lines in a scene. Many students said 3 is the maximum number since many scenes have 3 canonical directions e.g. city scenes.

Marking scheme: I gave 0.5 points for saying 3 (provided the justification made sense, namely that man made scenes have 3 canonical direction), and 1 point for saying infinite.

(b) This question was question 10c from Exercises 1.

Many students understood this exam question to mean "what are the variables in Eq. 2", for example, what is **K** and **R** ? There is more to the question that this, though. The question asks for a geometric interpretation of the "columns of **P**".

Let's first deal with the first three columns of **P**. Many students wrote that the first three columns of **P** are the image positions of each of the 3D world coordinate system direction vectors. Such an answer shows that you know that $(1, 0, 0, 0)$ represents a scene direction vector in the X direction, etc.

However, notice that such answer is not quite correct. How can an image *position* correspond to the projection of a 3D *vector* ? You need to say a bit more. For example, $(1, 0, 0, 0)$ is a point at infinity, and so the first column of **P** is the image vanishing point of lines that are in the direction of the scene X axis.

Marking scheme:

I gave 0.5 points if you wrote the first three columns of **P** correspond to the image of the scene axes XYZ, and gave 1 point if you you interpreted the scene axis vectors such as $(1, 0, 0, 0)$ as points at infinity.

Finally, the fourth column of **P** is the image position of the projection of the origin of the world coordinate system. I gave 1 point for writing that.

Some students wrote that the fourth column corresponds to a translation from the world to camera coordinate system. This describes the **C** vector, however, not the last column of **P**. I did not give anything for saying that.

3.

(a) Making a small angle approximation here ($\theta \approx \tan\theta$) gives $\frac{10}{200}$.

(b) $\frac{10}{200}$ $radians * 40\ mm * 100\ pixels/mm = 200\ pixels$.

(c) $\frac{cross\ section\ area}{distance\ squared} = \frac{\pi * .05^2}{2^2}$.

Marking scheme: 1 point for each. For (c), I gave 0.5 if you divided by distance, rather than distance squared.

4.

By assuming $r \ll h$, we can make a small angle approximation to compute Ω, namely solid angle can be computed by the ratio of the cross section area of the disk (πr^2) to the distance squared ($h^2 + X^2 + Y^2$).

$$\begin{aligned} E(X,Y) &= \int L(\mathbf{x}, \mathbf{l}_{in})\ \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in}\ d\Omega_{in} \\ &\approx L_{src}\ \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in}\ \ \Omega_{in} \\ &= L_{src}\ \frac{h}{\sqrt{X^2+Y^2+h^2}}\ \ \frac{\pi r^2}{X^2+Y^2+h^2} \end{aligned}$$

Marking scheme: I gave 1 point for correctly computing $\mathbf{n}(\mathbf{x}) \cdot \mathbf{l}_{in}$. For the first point, several students didn't normalize the $\mathbf{l}$ vector. I gave only 0.5 points for this.

I gave a second point for computing Ω_{in}. The key was to realize that the sphere is not foreshortened, and this makes the situation different from the case described in class. I gave 0.5 points if you did not realize this, but otherwise expressed the answer in terms of X, Y, h, r. (Also, if you did not get the point in Q3c because you divided by distance rather than distance squared, then I did not penalize you again here.)

5.

Decrease the aperture (e.g. by a factor $\frac{1}{\sqrt{2}}$) and increase the exposure time (e.g. by a factor 2). The latter is equivalent to decreasing the shutter speed.

Marking scheme: 1 point for giving the correct direction of changes to aperture and exposure time. I did not consider the specifics of $\frac{1}{\sqrt{2}}$ vs. 2, since the question did not explicitly ask for it.

6.

$$B * I(x) = \begin{cases} -1, & x = 1, 2 \\ 1, & x = -1, -2 \\ 0, & otherwise \end{cases}$$

Marking scheme: I gave 0.5 instead of 1 if you seemed to know what you were doing, but made a non-trivial mistake in the calculation.