Questions

1. For the camera calibration matrix $\mathbf{K}$,

   (a) Describe what is the effect on the appearance of an image if $\alpha_x < \alpha_y$, that is, $m_x < m_y$.

   (b) What are typical values of the pixel position p_x and p_y of the optical axis? (Your answer must be in terms of other camera parameters.)

2. The projection matrix

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

   is a 3×4 matrix, i.e. $\mathbf{I}$ is 3×3 and $\mathbf{C}$ is 3×1.

   (a) As a 3×4 matrix, you might think that it has 12 degrees of freedom. But in fact it has only 11. Why?

   (b) Can we relate the 11 degrees of freedom to the factorization into **K, R, C**? If so, how? If not, why not?

   (c) What is the null space of $\mathbf{P}$? ($\mathbf{P}$ is not invertible since it has more columns than rows.)

   (d) What are the four columns of $\mathbf{P}$? Geometrically, what do they represent?

   (e) What are the three rows of $\mathbf{P}$? Geometrically, what do they represent?

Answers

1. (a) $\alpha_x = fm_x$ and $\alpha_y = fm_y$ are the number of pixels per mm in the x, y directions. (**??**) If $\alpha_x < \alpha_y$ then the density of pixels is less in the x direction. When the image is displayed on a square grid of pixels on a monior it will appear squished in the x (horizontal) direction.

   (b) (p_x, p_y) is supposed to be in the center of the image, so it should be about $(N_x/2, N_y/2)$ where (N_x, N_y) are the number of pixels in the x, y directions.

2. (a) We can multiply the matrix $\mathbf{P}$ by any constant $a > 0$ but this will not change the projection, since $\mathbf{P}$ maps vectors that are represented in homogeneous coordinates, so $\mathbf{P}(a\mathbf{X})$ represents the same point as $\mathbf{PX}$. But, $\mathbf{PX}$ represents the same point as $a\mathbf{PX}$.

   (b) Yes. $\mathbf{K}$ has 5 parameters $\alpha_x, \alpha_y, s, p_x, p_y$. $\mathbf{R}$ has 3 parameters, namely 2 parameters to specify the unit direction in the first row, 1 parameter to specify the unit direction in the second row (only one parameter rather than 2, since this row vector must be orthogonal to the row vector in the first row), and 0 parameters to specify the unit direction in the third row (since it must be orthogonal to the first two rows, and so it can be specified by a cross product of the first two rows). Finally $\mathbf{C}$ has 3 parameters.

   (c) The null space of $\mathbf{P}$ is the set of vectors $\mathbf{X}$

   $$\mathbf{PX} = \mathbf{0}.$$

   By inspection, the camera position $\mathbf{X} = \mathbf{C}$ is a null vector since

   $$[\mathbf{I} \mid -\mathbf{C}]\mathbf{C} = \mathbf{O}$$

   that is,

   $$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 1 & -c_z \end{bmatrix} \begin{bmatrix} c_x \\ c_y \\ c_z \\ 1 \end{bmatrix}$$

   This is the only null vector since both $\mathbf{K}$ and $\mathbf{R}$ are invertible, so multiplication by $\mathbf{KR}$ cannot null out any other vectors.

   (d) The first three columns of $\mathbf{P}$ are, by definition, the image projection of the of the world coordinate axes $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, $(0, 0, 1, 0)$, which are direction vectors. The last column of $\mathbf{P}$ is the image projection of the world origin. Note that the world origin is not the same thing generally as the camera origin (and if it is, then fourth column will be $(0, 0, 0)^T$).

   (e) We begin with the third row of $\mathbf{P}$. If we treat this row as the set of coefficients of the 3D plane,

   $$P_{3,1}X + P_{3,2}Y + P_{3,3}Z + P_{3,4} = 0$$

   then any point (X, Y, Z) on this plane will map either to a point at infinity in the image, i.e. to some point $(*, *, 0)$. Actually there is one exception to this. Since the camera center $\mathbf{C}$ lies in the null space of $\mathbf{P}$, the camera center must lie on this plane as well.

What scene points project to image points at infinity? These are the scene points that lie in a plane containing the camera's center of projection and whose normal is the optical axis. This is called the *principal plane* of the camera, i.e. $Z = 0$ where Z is depth variable in the camera's coordinate system. In particular, the 3-vector $(P_{3,1}, P_{3,2}, P_{3,3})$ must be in the direction of the optical axis of the camera, since it is normal to the principal plane.

The first row of $\mathbf{P}$ can be interpreted as the set of coefficients of a different plane,

$$P_{1,1}X + P_{1,2}Y + P_{1,3}Z + P_{1,4} = 0$$

whose points are projected to the line $x = 0$ in the image plane. Note that since the camera center $\mathbf{C}$ lies in the null space of $\mathbf{P}$, the camera center must lie on this plane as well.

Does the optical axis lie in this plane? Generally no. Remember that $\mathbf{P}$ maps to pixel coordinates. Typically pixels start indexing at $x = 1$ on the left edge of the image.

By similar reasoning, the second row of $\mathbf{P}$ corresponds to a plane containing the camera center and the $y = 0$ plane. Again, this does not contain the optical axis.

