Camera Model and Triangulation

Instructor : Gabriel Taubin*

In these notes we will discuss the parameters used to model the camera projection operator from 3D object space to 2D image space, and camera calibration procedures.

1 COORDINATE SYSTEMS

The coordinates of 3D points are measured with respect to the *object coordinate system* (in meters or inches), are called *object coordinates*, and are denoted

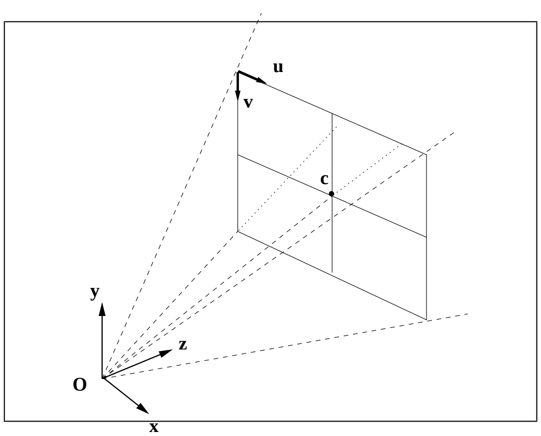
$$p = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} .$$

The projection of such points are measured with respect to the *camera coordinate system* (in pixels), are called *image coordinates*, and are denoted

$$p^I = \begin{pmatrix} p^I_u \\ p^I_v \end{pmatrix} .$$

When two or more cameras are considered, we will use additional indices to differentiate the different projections. For example, when a pair of stereo cameras used, we will denote the *left image coordinates* and *right coordinates*

$$p^L = \begin{pmatrix} p^L_u \\ p^L_v \end{pmatrix} \quad \text{and} \quad p^R = \begin{pmatrix} p^R_u \\ p^R_v \end{pmatrix} .$$



2 CAMERA MODEL

The function $p \mapsto p^I$ that let us compute the image coordinates as a function of the space coordinates is the *camera transformation* or projection. The camera transformation can be decomposed as the composition of an *extrinsic transformation* and an *intrinsic transformation*.

3 EXTRINSIC TRANSFORMATION

The extrinsic transformation is an Euclidean transformation $p \mapsto R\,(p - O)$ determined by the position and orientation of the camera with respect to the object coordinate system. The matrix R is a rotation with respect to the camera center of projection O of the camera. It transforms object coordinates to camera aligned coordinates. After the extrinsic transformation, the *center of projection* of the camera is at the origin, the image plane is parallel to the $x - y$ plane, the *camera axis* is the z axis, and the camera is looking along the camera axis in the direction of positive z, so that points with negative z coordinates are behind the camera. This is the *canonical* camera position and orientation.

4 INTRINSIC TRANSFORMATION

In this section we assume that the position and orientation of the camera is canonical. The intrinsic transformation is determined by the camera design and construction, including image sensor resolution, and optical components. We use a *radial distortion* model in which we decompose the intrinsic transformation into three factors. A first *pinhole projection* which produces unit-less *normalized coordinates*

$$a = \begin{pmatrix} a_u \\ a_v \end{pmatrix} = \begin{pmatrix} p_x/p_z \\ p_y/p_z \end{pmatrix} ,$$

a *radial distortion correction* factor, which is a first order approximation of the distortion introduced by the lens and also produces unit-less coordinates

$$b = (1 + k\,|a|^2)\,a = \begin{pmatrix} b_u \\ b_v \end{pmatrix} ,$$

and a final *viewport* transformation that converts the unit-less coordinates into pixel coordinates

$$p^I = \begin{pmatrix} p^I_u \\ p^I_v \end{pmatrix} = \begin{pmatrix} f_u\,b_u + c_u \\ f_v\,b_v + c_v \end{pmatrix} .$$

The parameter k that appears in the radial distortion factor is the *radial distortion coefficient*, and has typical values from 0.01 to 0.1. The parameters f_u and f_v are the *focal lengths* of the camera. Their unit of measurement is pixels, and include compensation for non-square camera pixels and proper scalling determined by the space to pixel unit conversion. The ratio $|f_u/f_v|$ is equal to the pixel aspect ratio of the camera, so that it the camera has square pixels, then $|f_u| = |f_v|$. The sign of the focal lengths is used to revert the orientation reversal produced by the top-down row scanning order in the camera. Usually $f_u > 0$ and $f_v < 0$. The parameters c_u and c_v measure the intersection of the optical axis with the image plane in pixels from the image origin (usually upper left corner).

The camera model normally used in computer graphics, which has no radial distortion ($k = 0$), has a simpler expression

$$p^I = \begin{pmatrix} f_u\,(p_x/p_z) + c_u \\ f_v\,(p_y/p_z) + c_v \end{pmatrix}$$

In addition to the radial distortion correction factor, the two focal lengths, and the image coordinates of the optical axis, the intrinsic

*taubin@brown.edu

parameters of the camera include the *width* and *height* of the image sensor measured in pixels.

5 INVERSE PROJECTION

The camera transformation assigns straight lines in space to pixels in the image plane. To reconstruct the position of a point in space from the image coordinates in two or more projections we need to compute the equation of the line in space coordinates corresponding to each image pixel. The problem is straightforward without lens distortion ($k = 0$), because the lens distortion factor is the only nonlinear part of the transformation. The parametric equation of the line is

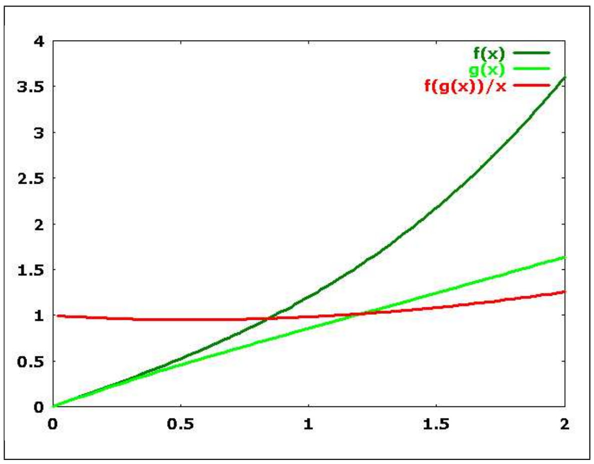
$$p = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \lambda \begin{pmatrix} (p_u - c_u)/f_u \\ (p_v - c_v)/f_v \\ 1 \end{pmatrix}$$

where λ is the parameter. Points in front of the image plane correspond to $\lambda \geq 1$.

An approximation to the inverse of the radial distortion factor is given by

$$a \approx \frac{b}{1 + k\ |b/(1 + k\,|b|^2)|^2}.$$

The following figure illustrates the approximation with $k = 0.2$, $f(x)$ the radial distortion function, and $g(x)$ the inverse approximation



6 CAMERA CALIBRATION

The camera calibration procedure consists in estimating the extrinsic parameters R and O, and the intrinsic parameters f_u, f_v, c_u, c_v, and k of the camera by taking pictures of objects with known dimensions, and finding the parameters that minimize the differences between the image coordinates measured in the images, and those predicted by the parameters. The standard method is to use a 3D structure such as a checker board pattern or a box, to simplify the feature detection and correspondence procedure.

7 STEREO TRIANGULATION

Ignoring the radial distortion, the image coordinates p^I of a point define a vector

$$\omega^I = R^t \begin{pmatrix} (p_u^I - c_u^I)/f_u^I \\ (p_v^I - c_v^I)/f_v^I \\ 1 \end{pmatrix}$$

and a line which we can describe in parametric form as

$$l^I = \{O^I + \lambda^I \omega^I : \lambda^I\}$$

Given the left and right image coordinates of a point, the problem is to compute the location of the point in object coordinates. Since due to measurement errors the two lines l^L and l^R may not intersect, we solve the problem in the least squares sense. We look for $\lambda = (\lambda^L, \lambda^R)^t$ minimizing

$$|(O^L + \lambda^L \omega^L) - (O^R + \lambda^R \omega^R)|^2$$

If we write $b = O^L - O^R$ and $A = [\omega^L| - \omega^R]$, the solution is

$$\lambda = -(A^t A)^{-1} A^t b$$