Robotics

Lecture 7: Camera Measurements

See course website
http://www.doc.ic.ac.uk/~ajd/Robotics/ for up to date information.

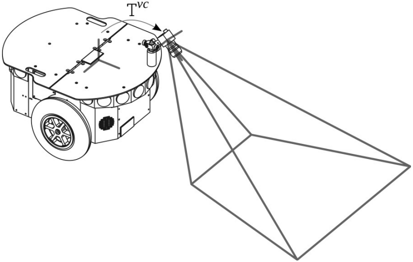
Andrew Davison
Department of Computing
Imperial College London

A Camera: a Projective Sensor

- A camera is a *projective* sensor. Each pixel captures the intensity and colour of the light arriving at the camera centre from one particular vector direction.
- In general, a camera does not capture geometric information directly because the depth of the surface emitting light is unknown. From one image, we do not know if objects are "small and close" or "large and far away" unless we have some other information — such as priors or multiple views
  `https://makezur.github.io/SuperPrimitive/`.

Using a Camera as a Planar Sensor

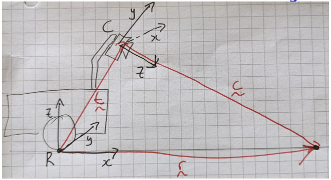
- A useful specific case of extra scene knowledge is if we know that the camera is observing the *ground plane*. Then every pixel in the image captured corresponds to a specific point on the ground plane.



The correspondence between points on the ground plane *specified in the robot coordinate frame* and pixels in the camera image is *fixed*, and depends on:

- The position and orientation of the camera relative to the robot (sometimes called *extrinsic calibration parameters*).
- The *intrinsic calibration parameters* of the camera (focal length and principal point).

Camera Geometry



The vector from the centre of the camera to a point on the ground, expressed in the coordinate frame C of the camera, is:

$$\mathbf{c}^C = \mathtt{R}^{CR}\left(\mathbf{r}^R - \mathbf{t}^R\right) \ .$$

Here $\mathtt{R}^{CR}$ is the rotation matrix relating robot frame R and camera frame C; $\mathbf{t}^R$ is the vector from the robot centre to the camera centre; and $\mathbf{r}^R$ is the vector from the robot centre to the point (both expressed in the robot frame).

A perspective camera projects a 3D point at vector $\mathbf{c}^C$ to image coordinates (u, v) via the *camera calibration matrix* $\mathtt{K}$:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathtt{K}\mathbf{c}^C \ .$$

Homogeneous Coordinates

A quick note about the notation $\begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$: this is a homogenous (or projective) vector. The idea is that we use a 3-vector to represent something which really has two elements, with the interpretation that to get the two elements we are actually interested in we divide each of the first two elements by the 3-vector by the third one.

Ground Plane Homography

Putting this together: the projected image coordinates (u, v) of a point at position $\mathbf{r}^R$ are given by:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathtt{KR}^{CR} \left(\mathbf{r}^R - \mathbf{t}^R\right)$$

If we know that the point is on the ground plane, then we can write:

$$\mathbf{r}^R = \begin{pmatrix} x_R \\ y^R \\ 0 \end{pmatrix}$$

And therefore:

$$\mathbf{r}^R - \mathbf{t}^R = \begin{pmatrix} x^R - t_x^R \\ y^R - t_y^R \\ -t_z^R \end{pmatrix} = \mathtt{T} \begin{pmatrix} x^R \\ y^R \\ 1 \end{pmatrix},$$

where

$$\mathtt{T} = \begin{bmatrix} 1 & 0 & -t_x^R \\ 0 & 1 & -t_y^R \\ 0 & 0 & -t_z^R \end{bmatrix}$$

Ground Plane Homography Continued

So we can write:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathtt{K}\mathtt{R}^{CR}\mathtt{T} \begin{pmatrix} x^R \\ y^R \\ 1 \end{pmatrix}$$

The three 3×3 matrices $\mathtt{K}$ (the camera calibration matrix), $\mathtt{R}^{CR}$ (rotation matrix of the orientation of the camera relative to the robot frame), and $\mathtt{T}$ (encodes the translation of the camera relative to the robot frame) all have unknown elements, but rather than estimating all of these elements separately, we can roll all of them up into a single matrix:

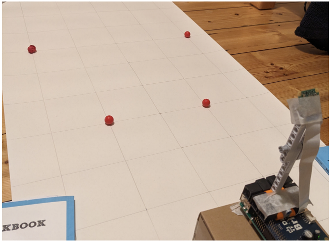
$$\mathtt{H} = \mathtt{K}\mathtt{R}^{CR}\mathtt{T}$$

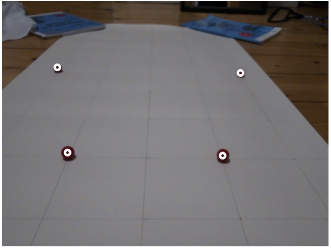
$\mathtt{H}$ is also a 3×3 matrix called the *ground plane homography*. Therefore: So we can write:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathtt{H} \begin{pmatrix} x^R \\ y^R \\ 1 \end{pmatrix}$$

Direct Calibration of a Ground Plane Homography

- Instead of measuring all of the individual parameters which the ground plane homography depends on and calculating the matrix explicitly, we can estimate the homography directly from *correspondences*.
- We mark a minimum of four points on the ground plane in carefully measured positions, then capture an image and record the image coordinates of the points.





(Note that this could be done with more than 4 points for extra accuracy.)

Calibrating the Ground Plane Homography from 4 Correspondences

Homography definition:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \propto \mathtt{H} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$\mathtt{H}$ is defined up to scale, so let us write:

$$\mathtt{H} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$$

Writing it with an explicit scale factor k, from one correspondence $(u, v) \rightarrow (x, y)$ we get three equations:

$$\begin{pmatrix} ku \\ kv \\ k \end{pmatrix} = \begin{pmatrix} ax + by + c \\ dx + ey + f \\ gx + hy + 1 \end{pmatrix} .$$

Calibrating the Homography from 4 Correspondences

From which we deduce:

$$(gx + hy + 1)u = ax + by + c \implies ax + by + c - uxg - uyh = u$$

and

$$(gx + hy + 1)v = dx + ey + f \implies dx + ey + f - vxg - vyh = v$$

If we have four corresponding points, we can use each one to produce two lines of the following matrix equation which we can solve to find the elements $a \dots h$ of $\mathtt{H}$:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1x_1 & -v_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -u_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2x_2 & -v_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -u_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3x_3 & -v_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4x_4 & -u_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4x_4 & -v_4y_4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{bmatrix}$$

We can solve for $a \dots h$ with for instance `np.linalg.lstsq()`.

Using the Ground Plane Homography

- Having estimated the homography $\mathtt{H}$, we can use its inverse to find the ground plane coordinates of a point for which we know the image coordinates:

$$\begin{pmatrix} x^R \\ y^R \\ 1 \end{pmatrix} \propto \mathtt{H}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

- But remember that this is only valid for points which we know to lie on the ground plane. So how do we use it in practice?
- One situation is where there are point-like obstacles or markers on the ground. We can detect these with some image segmentation and blob detection, then use the inverse homography to find the ground plane coordinates of the obstacles.
- In principle we could carry out systematic experiments with markers in known positions to investigate the uncertainty in the ground plane measurements.

Using the Ground Plane Homography

- As an alternative use, consider a scene where the camera observes the ground in front of the robot and the intersection of the ground with some walls or obstacles. Using some image processing we can attempt to *segment* the image into ground plane and obstacle regions.
- In simple scenes, basic colour-based segmentation will be good enough to do this. In more complex scenes, a trained classifier (such as a small U-Net neural network) could be used.
- The region at the bottom of the image, closest to the robot, will usually lie in the ground plane. We are interested in the *boundary* between the ground plane and obstacles. These boundary points lie in the ground plane, so our homography is valid.
- By finding the ground plane coordinates of a number of points on the boundary, we turn our camera into a kind of simple laser scanner which can simultaneously measure the distance to a set of points on a wall.