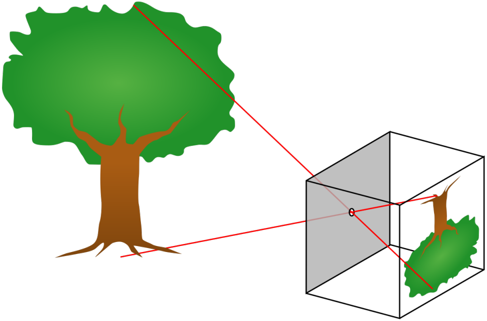
Lecture 2: Single view geometry

6 February, 2020

Recall: Geometric description of a camera

In the previous lecture we started with a simple model of a camera, given by projecting through a pinhole onto a screen.



In many applications this is a reasonable model, although for some types of images (e.g. a picture taken with a fisheye lens) we would need a more complicated model that incorporates the distortion caused by the lens.

In these lectures we will focus on the pinhole camera model.

Recall: The projective line $\mathbb{R}P^1$

Recall that last time we defined the projective line $\mathbb{R}P^1$ as the set of lines through the origin in $\mathbb{R}^2$.

On the projective line we have homogeneous coordinates, where $[x : y]$ represents the line through the origin and the point $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

Note. At least one of the coordinates must be non-zero. The projective coordinate $[0 : 0]$ does not make sense, since there is no unique line through the origin and $(0, 0)$.

Since there are infinitely many points on this line, and they all have coordinates $\lambda(x, y)$ for some $\lambda \in \mathbb{R} \setminus \{0\}$, then we see that

$$[x : y] = [\lambda x : \lambda y] \quad \text{for all } \lambda \in \mathbb{R} \setminus \{0\}.$$

This says that the coordinates $[x : y]$ and $[\lambda x : \lambda y]$ represent the same line through the origin in $\mathbb{R}^2$, or equivalently the same point in $\mathbb{R}P^1$

The projective plane $\mathbb{R}P^2$

Now we can move up one dimension and extend this concept to the projective plane, denoted $\mathbb{R}P^2$. This is the set of all lines in $\mathbb{R}^3$, defined using the equivalence relation

$$\mathbb{R}P^2 = (\mathbb{R}^3 \setminus \{(0,0,0)\})/\sim$$

where $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ if and only if $(x_1, y_1, z_1) = \lambda(x_2, y_2, z_2)$ for some $\lambda \in \mathbb{R} \setminus \{0\}$.

Again, we can use homogeneous coordinates, where the line through the origin and (x, y, z) is represented by $[x : y : z]$. These are subject to the relation

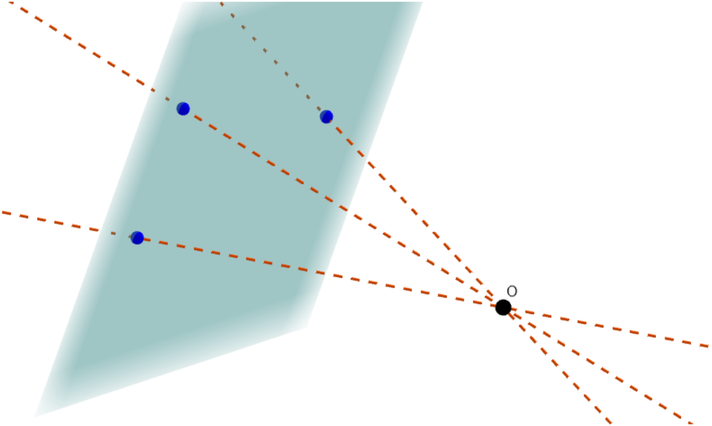
$$[x : y : z] = [\lambda x : \lambda y : \lambda z] \quad \text{for any } \lambda \in \mathbb{R} \setminus \{0\}.$$

Note that this makes sense, since $[x : y : z]$ and $[\lambda x : \lambda y : \lambda z]$ both represent the same line through the origin in $\mathbb{R}^3$.

The projective plane $\mathbb{R}P^2$

Given any plane in $\mathbb{R}^3$ (not passing through the origin) we can describe it using homogeneous coordinates.

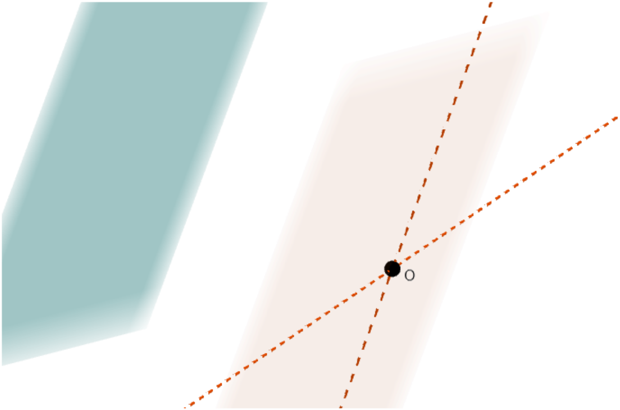
Each point on the plane corresponds to a line through the origin.



The projective plane $\mathbb{R}P^2$

Any line through the origin parallel to the plane corresponds to a "point at infinity" on this plane.

(Compare with the point at infinity in the projective line.)



The projective plane $\mathbb{R}P^2$

We can think of these lines through the origin parallel to the plane as "points on the horizon".

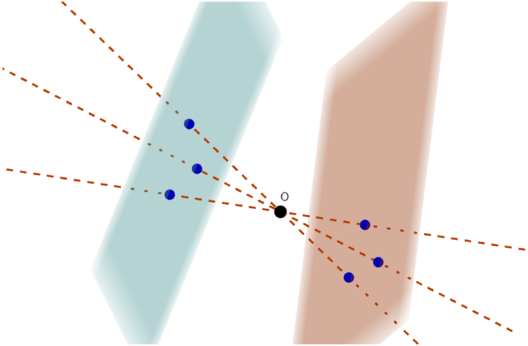
This is motivated by our previous picture of the airport runway. Any line of sight parallel to the plane of the runway will appear to us as a point on the horizon.

The projective plane $\mathbb{R}P^2$

Given two different planes in $\mathbb{R}^3$, we can define a correspondence using lines through the origin.

A point on one plane corresponds to a point on the other if they both lie on the same line through the origin.

Equivalently, a point on one plane corresponds to a point on the other if they both have the same homogeneous coordinates.



Realising the projective plane $\mathbb{R}P^2$

Example. We will use projective coordinates to describe the plane $\{z = 1\}$ in $\mathbb{R}^3$.

Consider a line through the point $(a, b, c) \in \mathbb{R}^3 \setminus \{0\}$. In homogeneous coordinates, this is represented by $[a : b : c]$.

Suppose that $c \neq 0$ so that the line intersects the plane $\{z = 1\}$.

Then the line intersects the plane $\{z = 1\}$ in the point $\left(\frac{a}{c}, \frac{b}{c}, 1\right)$.

Conversely, any point $(x, y, 1)$ in the plane $\{z = 1\}$ corresponds to a line through the origin with homogeneous coordinates $[x : y : 1]$ (i.e. the last coordinate is not zero).

Therefore, we have seen that there is a one-to-one correspondence between points in the plane $\{z = 1\}$ and the points in $\mathbb{R}P^2$ with homogeneous coordinates $[x : y : 1]$.

Realising the projective plane $\mathbb{R}P^2$ (cont.)

What about the horizon line (points at infinity)?

Any line through the origin and a point $(a, b, 0)$ will be parallel to the plane $\{z = 1\}$.

In homogeneous coordinates, this line is represented by $[a : b : 0]$.

Therefore the "line at infinity" (or the "horizon line") corresponds to all the points in $\mathbb{R}P^2$ with homogenous coordinates $[a : b : 0]$ (i.e. the last coordinate is zero).

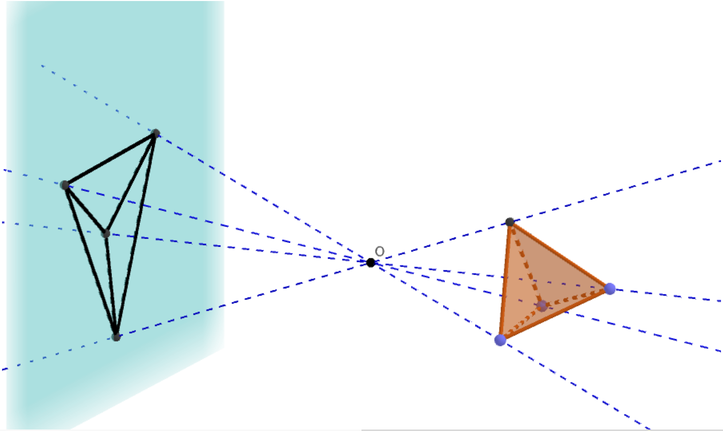
We can relate this to the "airport runway" example, where the runway is on the plane $\{z = 1\}$ and we are viewing the runway from the point $(0, 0, 0)$. If we look in the direction of the point $(a, b, 0)$ then we will be looking at the horizon.

Understanding this example will be important later when we model the pinhole camera.

Algebraic description of a camera

Now we are ready to use the geometry of projections to study our model of a pinhole camera.

First consider the following simple model. We are projecting points from $\mathbb{R}^3$ through a pinhole placed at the origin onto the blue image plane.



Algebraic description of a camera (cont.)

How do we describe this in coordinates? If a point on the object has coordinates (x, y, z), then what are the coordinates in the image plane?

First note that all points on the same line through the origin will be projected to the same point in the image plane. Therefore the points (x, y, z) and $(\lambda x, \lambda y, \lambda z)$ will be projected to the same point.

Therefore, we should study the lines through the origin in $\mathbb{R}^3$ and see how to project them onto the image plane.

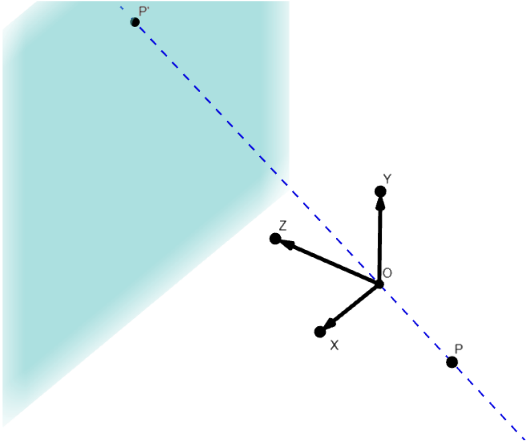
Each point in the image plane corresponds to a unique line in $\mathbb{R}^3$.

Not all lines in $\mathbb{R}^3$ correspond to a point in the image plane. Recall that any line parallel to the image plane will project to the horizon line, or a point on the line at infinity.

Therefore we should think of this projection as a projection onto $\mathbb{R}P^2$, since we are projecting onto the plane as well as the horizon line.

Algebraic description of a camera (cont.)

In coordinates, suppose that the image plane corresponds to $\{z = c\}$.



The point $P = (x, y, z)$ projects to $P' = \left(\frac{cx}{z}, \frac{cy}{z}, c\right)$.

Algebraic description of a camera (cont.)

In homogenous coordinates, any line which is not on the horizon will project to the point

$$[x : y : z] \mapsto \left(\frac{cx}{z}, \frac{cy}{z}, c\right) \in \mathbb{R}^3$$

In the 2-dimensional image plane, we ignore the last coordinate, and use the coordinates

$$P' = \left(\frac{cx}{z}, \frac{cy}{z}\right).$$

In our projection, we have to consider points on the horizon (corresponding to lines parallel to the image plane). Therefore we should use homogeneous coordinates in the image plane.

The image point $P' = \left(\frac{cx}{z}, \frac{cy}{z}\right)$ then corresponds to

$$P' = \left[\frac{cx}{z} : \frac{cy}{z} : 1\right] = [cx : cy : z].$$

Algebraic description of a camera (cont.)

Summary. We started with a pinhole camera centred at the origin.

This projects points in $\mathbb{R}^3$ onto a screen (or camera film) corresponding to the plane $\{z = c\}$.

We wrote the point $P = (x, y, z) \in \mathbb{R}^3$ in homogeneous coordinates as $P = [x : y : z]$.

The camera projects this to the point $P' = \left(\frac{cx}{z}, \frac{cy}{z}\right)$ on the image plane.

In the homogeneous coordinates of the image plane, this corresponds to the point

$$P' = [cx : cy : z] = \left[\frac{cx}{z} : \frac{cy}{z} : 1\right]$$

Therefore, our pinhole camera can be represented by a linear map which takes (x, y, z) to $[cx : cy : z]$.

Equivalently, we can write this as a matrix.

Algebraic description of a camera (cont.)

The matrix that maps (x, y, z) to $[cx : cy : z]$ is

$$\begin{pmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

In our model, we used the origin $(0, 0, 0)$ as the location of the pinhole (or "camera centre"). Later on we will want to generalise this, so we will write the point $(x, y, z) \in \mathbb{R}^3$ in homogeneous coordinates as $[x : y : z : 1]$.

In these coordinates, the matrix becomes

$$\mathrm{M} = \begin{pmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

This is the camera matrix for our model of the pinhole camera. It is a 3×4 matrix representing the projection onto the image plane.

Properties of the camera matrix (camera centre)

Now we will study some properties of the camera matrix.

The matrix M represents the projection induced by taking lines through the origin in $\mathbb{R}^3$. All points in $\mathbb{R}^3$ project to the image plane and its horizon line, except for the origin $(0, 0, 0)$ (equivalently, $[0 : 0 : 0 : 1]$ in projective coordinates).

Let C be the column vector representing $[0 : 0 : 0 : 1]$. This is the camera centre.

Then the vector C is in the kernel of M.

$$\mathrm{MC} = \mathrm{M}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}c & 0 & 0 & 0\\0 & c & 0 & 0\\0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = 0.$$

In general, our camera centre may not be at the origin in $\mathbb{R}^3$. If we know the camera matrix M, then we can find C as a vector in the kernel of M.

Properties of the camera matrix (rank)

Note also that the camera defines a projection from $\mathbb{R}^3$ onto a 2-dimensional plane, the image plane.

In homogeneous coordinates, we are projecting onto $\mathbb{R}P^2$ (which includes the horizon line).

This means that the camera matrix must have rank 3, so that the image is all of the projective plane.

If the rank of the camera matrix is less than 3, then the image will be a line or a point, and therefore won't represent a real camera.

Example. If the camera matrix is

$$\mathrm{M} = \begin{pmatrix} c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

then the homogeneous coordinates of the image will be $[cx : 0 : z]$, and the image will be a line in the plane $\{z = c\}$.

Properties of the camera matrix

Summary. We have seen that

1. The kernel of the camera matrix tells us the camera centre.

2. The camera matrix must have rank 3.

We want to extend our model to a general pinhole camera with

(a) an arbitrary centre (c_1, c_2, c_3), and

(b) an arbitrary image plane $\{ax + by + cz = d\}$.

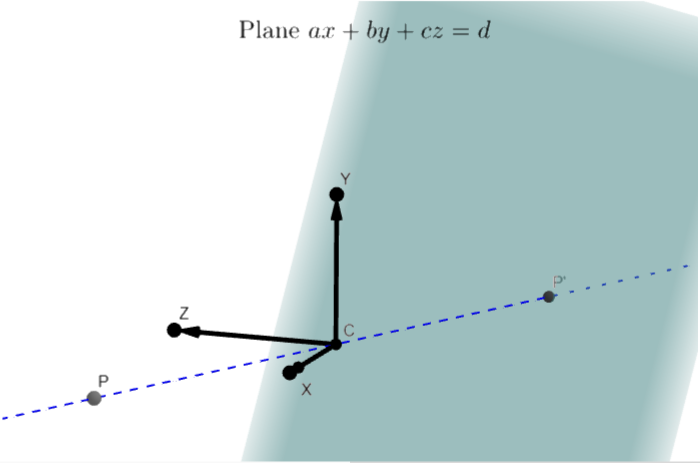
Previously, the centre of the coordinate system on the image plane was at $[0 : 0 : 1]$. We may also want to consider

(c) a more general coordinate system on the image plane.

Moving the image plane

First consider a camera with centre C at the origin, and image plane

$$\{ax + by + cz = d\}.$$



Moving the image plane (cont.)

We can relate this to the previous problem by rotating the coordinate system.

The equation for the plane is $ax + by + cz = d$. We can rewrite this as

$$((x, y, z) - (x_0, y_0, z_0)) \cdot \widehat{\mathbf{n}} = 0$$

where (x_0, y_0, z_0) is a point on the plane and $\widehat{\mathbf{n}}$ is a unit vector normal to the plane.

We want find new coordinates (x', y', z') such that the equation of the plane is now $\{z' = D\}$ for some constant D depending on a, b, c, d.

The new coordinates (x', y', z') are related to the old coordinates (x, y, z) by

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where R is a 3×3 rotation matrix.

Rotation matrices

Definition. A rotation matrix in $\mathbb{R}^3$ is a 3×3 matrix such that the rows are mutually orthonormal.

$$R = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Equivalently, each row vector $\mathbf{r_i} = (a_{i1} \ a_{i2} \ a_{i3})$ satisfies $\mathbf{r_i} \cdot \mathbf{r_j} = \delta_{ij}$, where δ_{ij} is the *Kronecker delta* ($\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$).

Equivalently, the rows are all mutually orthogonal unit vectors.

Example. The matrix R below corresponds to anticlockwise rotation by an angle θ around the z-axis.

$$R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Moving the image plane (cont.)

In the coordinates (x', y', z'), the image plane is $\{z' = D\}$. Therefore, in these coordinates, the camera matrix is

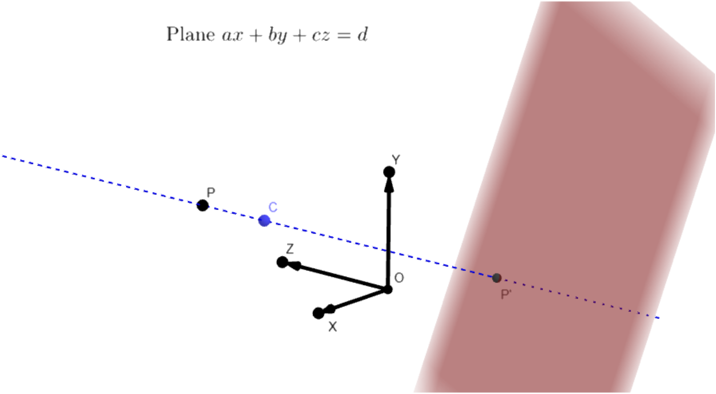
$$\mathrm{M}' = \begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Now we want to change coordinates back to (x, y, z) (equivalently, we are changing the basis of the coordinate system). We can do this by conjugating by the rotation matrix.

$$\mathrm{M} = R\,\mathrm{M}' \begin{pmatrix} R^{-1} & 0 \\ 0 & 1 \end{pmatrix}.$$

Moving the camera centre

Now consider what happens if we move the camera centre.



Once again, we can solve this problem by translating the coordinates so that the camera centre is the origin in the new coordinate system.

Moving the camera centre (cont.)

Let (c_1, c_2, c_3) be the coordinates of the camera centre C.

First change coordinates $(x', y', z') = (x, y, z) - (c_1, c_2, c_3)$, so that the camera centre is mapped to the origin.

Now do our previous rotation $\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}$ so that the image plane is mapped to $\{z'' = D\}$.

Now the camera matrix is the same as in our first example.

$$\mathrm{M}'' = \begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Now we need to apply the inverse matrices to undo all the previous coordinate transformations to get back to the original coordinate system again.

Moving the camera centre (cont.)

Let (c_1, c_2, c_3) be the coordinates of the camera centre C.

The first coordinate change (translation by $-(c_1, c_2, c_3)$) is given in homogeneous coordinates by

$$\mathsf{T}_{-C}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -c_1 \\ 0 & 1 & 0 & -c_2 \\ 0 & 0 & 1 & -c_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x - c_1 \\ y - c_2 \\ z - c_3 \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix}$$

The inverse of this matrix is

$$\mathsf{T}_{C}\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & c_1 \\ 0 & 1 & 0 & c_2 \\ 0 & 0 & 1 & c_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} x' + c_1 \\ y' + c_2 \\ z' + c_3 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Moving the camera centre (cont.)

Therefore, our camera matrix M is determined by the following process.

First apply T_{-C} so that the camera centre is moved to the origin.

Now apply $\begin{pmatrix} R^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ so that the image plane is now of the form $\{z = D\}$.

Now apply the matrix $\begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ which corresponds to projecting to the plane $\{z = D\}$ using the origin as the camera centre.

Now apply the matrix R to rotate back to the original coordinates.

Now translate the image plane back to the original image plane.

Next time

We will study the process of viewing the same picture with two different cameras, and see how to relate the two images using geometry.

We will also see how using more than one camera adds depth and three dimensionality to the image.