Computer Vision and Pattern Recognition

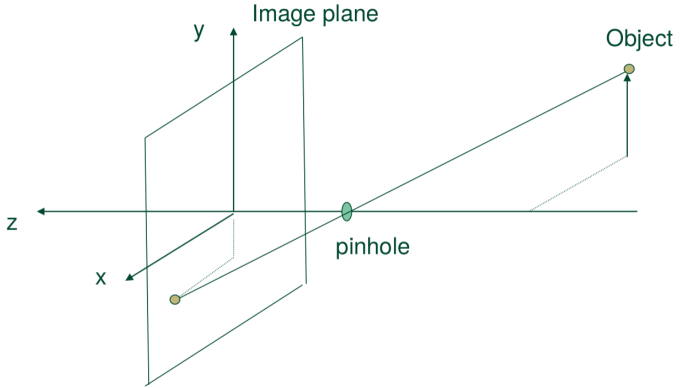
Krystian Mikolajczyk & Ad Spiers

Department of Electrical and Electronic Engineering

Imperial College London

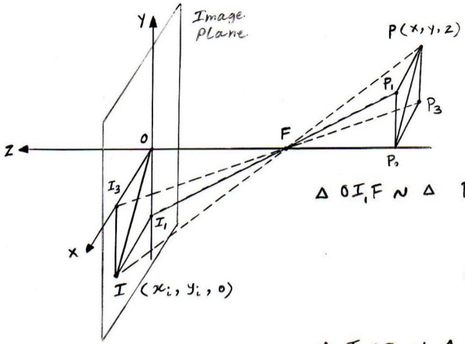
Part 1: Camera model and geometry

- Pinhole camera model
- Homogeneous coordinates



f - focal length of the camera

Pinhole camera model



$$\triangle\, OI_1F \sim \triangle\, P_2P_1F \Rightarrow \frac{OI_1}{OF} = \frac{P_2P_1}{P_2F} \Rightarrow \frac{-y_i}{f} = \frac{y}{-z-f}$$

$$\Rightarrow y_i = \frac{yf}{z+f} \qquad (1)$$

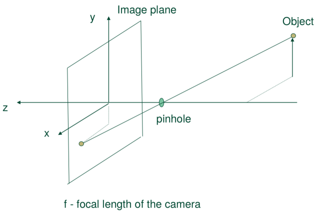
$$\triangle\, I_3OF \sim \triangle\, P_3P_2F \Rightarrow \frac{I_3O}{OF} = \frac{P_3P_2}{P_2F} \Rightarrow \frac{x_i}{f} = \frac{-x}{-z-f}$$

$$\Rightarrow x_i = \frac{xf}{z+f} \qquad (2)$$

Equations (1) and (2): Given position of object point, return position of the corresponding image point.

$$x = \frac{x_i(z+f)}{f} \quad (3) \qquad y = \frac{y_i(z+f)}{f} \quad (4)$$

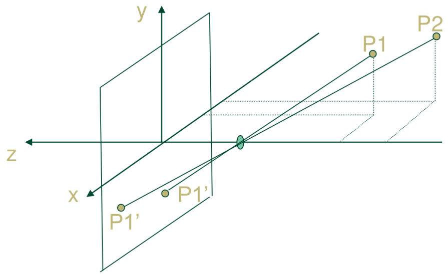
Pinhole camera model



- Position of a 3D object point: (x, y, z)
- Homogeneous coordinates of the same point: (kx, ky, kz, k)
  - To recover actual coordinates divide first three coordinates by k
- Augmented position vector of image point: $\mathrm{x}_i = (x, y, z, k)$
- Augmented position vector of object point in 3D: $\mathrm{x} = (x, y, z, 1)$
- We now have a linear relationship between the two: $\mathrm{x}_i = C \cdot \mathrm{x}$ where
  - Let: $k = \frac{z+f}{f}$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{bmatrix}, \qquad \begin{bmatrix} x \\ y \\ z \\ k \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \implies \begin{aligned} x &= x \\ y &= y \\ z &= z \\ k &= \frac{z}{f} + 1 \end{aligned}$$

Pinhole camera model



- To obtain point coordinates in the image :
  $(x_i, y_i, z_i) = (\frac{xf}{z+f}, \frac{yf}{z+f}, \frac{zf}{z+f},)$
  - disregard z_i

Exercise

- $P1 : x = 10,\ y = 5,\ z = 20$
- $P2 : x = 20,\ y = 10,\ z = 10$
- $f = 2$
- Task: Calculate the distance between $P1'$ and $P2'$ in the image

Pinhole camera model

- 3D to image projection

$$\begin{bmatrix} kx_i \\ ky_i \\ k \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & C_{34} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Camera Calibration
  - The process of estimating values C_{ij}
- Textbook: Visual Geometry, Hartley and Zisserman

Homogeneous coordinates

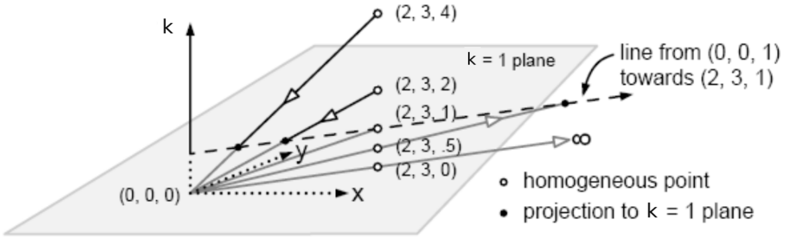
Allow various image transformations to be easily represented by a matrix

- Projective space : *ordinary plane* augmented with points at infinity is known as the *projective plane*
- Practical aspect: unification of the translation, scaling and rotation of geometric objects.
- $(x, y, z, \ldots, k)$ is a vector of length $n + 1$, other than $(0, 0, 0, \ldots, 0)$, where n is the number of dimensions
- Two sets of coordinates that are proportional and denote the same point of projective space for any non-zero scalar c i.e. $(cx, cy, cz, \ldots, ck)$ denotes the same point
- The plane at infinity is usually identified with the set of points with $k = 0$.

Homogeneous coordinates

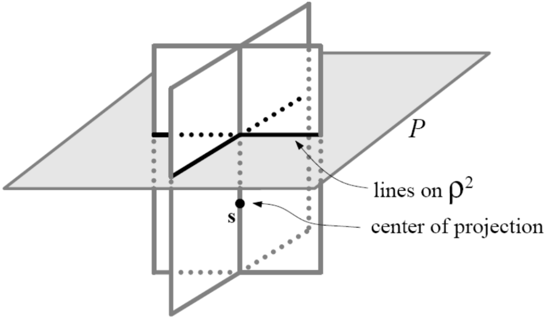
Homogeneous points represent the projection, and can also represent points at infinity.

- Consider a homogeneous point as k approaches 0;
- As k approaches 0, the projected Euclidean points move away from the origin in the $(2, 3)$ direction.
- At $k = 0$, the point is infinitely far and may be treated as a positionless vector.
  - Point $(2, 3, k)$ is shown for $k = \{4, 2, 1, 1/2, 0\}$.



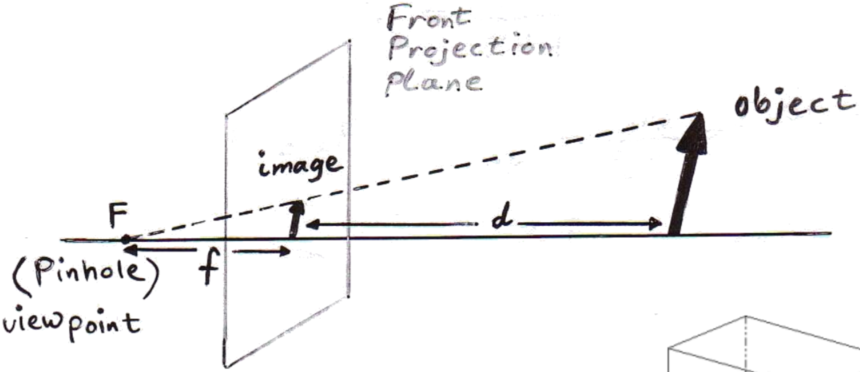
Homogeneous coordinates: Projective plane

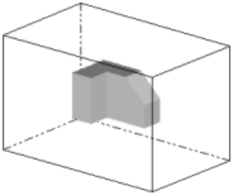
- Consider all lines and planes passing through a given point s;
- If they are intersected by a plane P that does not pass through s, then each point (or line) on P may be associated with a line (or plane) through s
- The division by k means that the conversion of a homogeneous point to its Euclidean equivalent is inherently a projection of the homogenous point onto the $k = 1$ plane.



Homogeneous coordinates: Orthographic projection

- Representing a three-dimensional object in two dimensions





orthographic projection f>>d