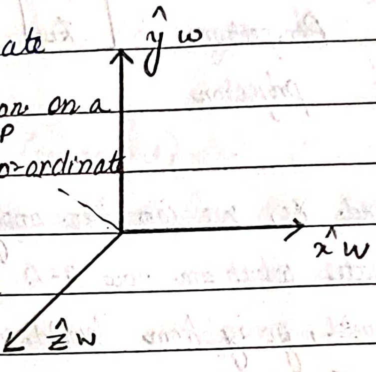
- the third channel is very similar to 'a' channel, tells about the spectrum from blue to yellow. the more blue the pixel is, the lower the value of this particular pixel in the channel. the higher the value, the more yellow we have

Camera Model and Camera Calibration:

→ Lets develop a comprehensive model of the camera so that we can calibrate the camera for this.

↳ Forward Imaging Model: (3D to 2D)

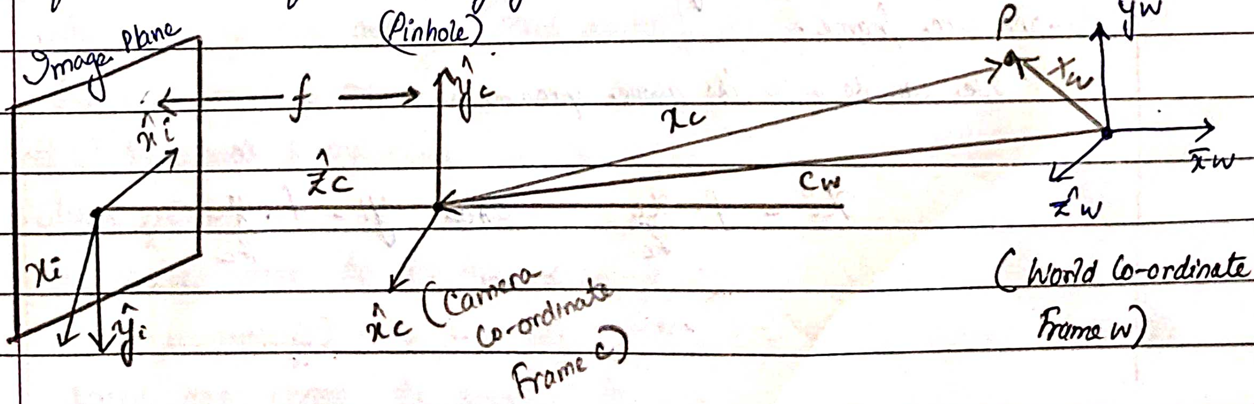
→ A world co-ordinate frame, w, we focus our attention on a single point in this world co-ordinate frame. That point is p.



→ In this World co-ordinate frame lies your camera & the camera is defined by its own co-ordinate frame c, where the z-axis of the camera co-ordinate frame is aligned with the optical axis of the camera.

→ Lets assume that the effective focal length (distance between the effective central projection & the image plane of the camera, is F right here)

If we know the Relative position & the orientation of the camera co-ordinate frame with respect to the world co-ordinate frame, then we can write an expression that takes you all the way from the point B in the world coordinate frame to its projection x_i on the image plane. That complete mapping is what we refer to as the forward Imaging model.



↳ We start with a point in the world co-ordinate frame X_W, and then, we're going to model its transformations to the camera coordinate System frame.

3D – 3D transformation from the point X_W to X_c

Image co-ordinates — Camera co-ordinates — World-Co-ordinates

$$X_P = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \xleftarrow[\text{Perspective projection}]{} X_c = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \xleftarrow[\text{(Co-ordinate Transformation)}]{} X_W = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$$

↳ Once we obtained X_c, we can then apply perspective projection to end up with the Image co-ordinates which are now 2-D co-ordinates on the Image plane X_i.

↳ This whole model, going from World-coordinates to Image co-ordinates 3D - 2D is called as Forward Imaging Model.

↳ This Model will be used to develop a comprehensive linear Model for the Camera.

Let's start with Perspective projection:

Equation of Perspective projection:

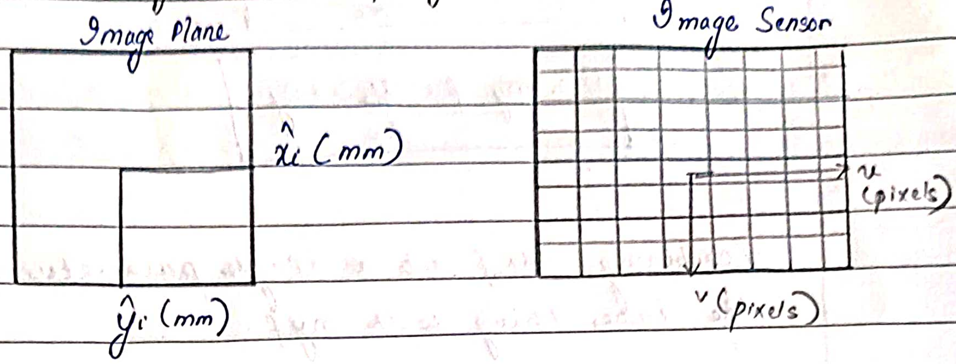
$$\frac{x_i}{f} = \frac{x_c}{z_c} \quad \text{and} \quad \frac{y_i}{f} = \frac{y_c}{z_c}$$

In this, we are assuming that the point has been defined in the camera Co-ordinate frame.

· We can Re-write the above equation as →

$$x_i = f \cdot \frac{x_c}{z_c} \quad \text{and} \quad y_i = f \cdot \frac{y_c}{z_c}$$

Image plane to Image Sensor Mapping:



Given the 2 co-ordinates in terms of mm, but in reality what we have is an image sensor which is used to capture the image. These Image Sensors have pixels.

↳ We have pixel co-ordinates (u and v).

↳ We have to figure out the mapping from the Image co-ordinates here in mm to pixels. — (Task)

* Pixels can be rectangular as well, not just square

If m_x and m_y are the pixel densities (pixels/mm) in x and y directions respectively, then pixel co-ordinates are:

$$u = m_x x_i \qquad v = m_y y_i$$

$$= \boxed{m_x \cdot f \frac{x_c}{z_c}} \qquad = \boxed{m_y \cdot f \frac{y_c}{z_c}}$$

↳ Now, the pixel co-ordinates of the projection of the point → P, R, U & V where u and v are given above. Now, we have gone from mm to pixels.

↳ The value of m_x and m_y (pixel densities) are unknown. They're part of the calibration process.

↳ (0,0) corresponds to the center of the Image where the Optical axis pierces the Image plane.

↳ We usually treat the top-left corner of the image sensor as its origin. (easier for indexing). If pixel (O_x, O_y) is the principle point, where the optical axis pierces the sensor, then →

Date

$$\boxed{u = m_x \cdot f \cdot \frac{x_c}{z_c} + O_x}$$

$$\boxed{v = m_y \cdot f \cdot \frac{y_c}{z_c} + O_y}$$

Combining $m_x f$ into a single parameter f_x, we can do the same thing with $m_y f$.

$$\boxed{u = f_x \cdot \frac{x_c}{z_c} + O_x} \qquad \boxed{v = f_y \cdot \frac{y_c}{z_c} + O_y}$$

where $(f_x, f_y) = (m_x f, m_y f)$ are the focal lengths in pixels in x & y directions.

- Camera has only one focal length, effective focal length, the distance b/t the center of projection & the image plane.
- In order to accomodate for Non-equal pixel density, as in the X-Y directions, the fact that pixels may be rectangular
- $(f_x, f_y, O_x, O_y) \rightarrow$ Intrinsic parameters of the camera. They represent the camera's Internal geometry.
- Equations for perspective projection are Non-Linear. It is convenient to express them as linear equation to make the estimation process easier.

Q: How do we go from Non-Linear Model to Linear Model?
$\rightarrow$ Using homogeneous co-ordinates.
Let us see about this $\rightarrow$

Homogeneous Co-ordinates:

The homogeneous representation of a 2D point $u = (u, v)$ is a 3D point $\tilde{u} = (\tilde{u}, \tilde{v}, \tilde{w})$. The third co-ordinate $\tilde{w} \neq 0$ is fictitious such that —

$$u = \frac{\tilde{u}}{\tilde{w}} \quad , \quad v = \frac{\tilde{v}}{\tilde{w}}$$

$$u = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{w}u \\ \tilde{w}v \\ \tilde{w} \end{bmatrix} \equiv \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} = \tilde{u}$$

The third co-ordinate is used for scaling or normalization process.

Visual Representation of u,v:

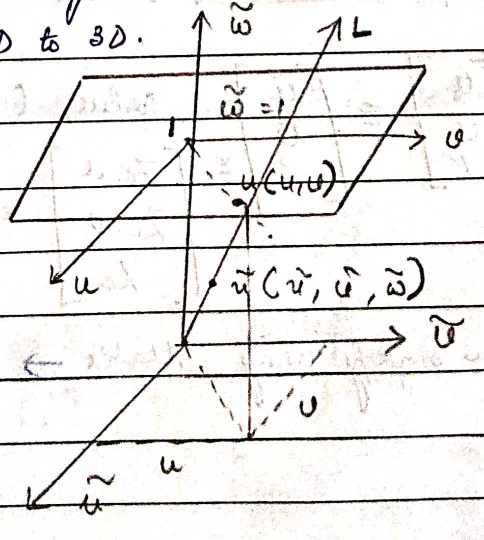
→ Imagine uv plane, and in Relation to this plane, we can define a co-ordinate frame $\tilde{u}, \tilde{v}, \tilde{w}$ such that uv plane lies at ($\tilde{w} = 1$)

→ All the points go from the origin, but not including the origin, go from the origin through the point to u.

→ All of the point on 'L' corresponds to homogeneous co-ordinates corresponding to this point to u. They're all equivalent to each other

→ given any point on this line 'L', you can find from that point u, so 'u' can be expressed in homogeneous co-ordinates as ~~all~~ $u = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$

→ So that's going from 2D to 3D.



Page No.	
Date	

Homogeneous co-ordinates, given 3D point:

↳ The homogeneous Representation of a 3D point $x = (x, y, z) \in R^3$ is a 4D point $\tilde{x} = (\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}) \in R^4$.

↳ The 4th co-ordinate $\tilde{w} \neq 0$ is fictitious such that:

$$x = \frac{\tilde{x}}{\tilde{w}}, \quad y = \frac{\tilde{y}}{\tilde{w}}, \quad z = \frac{\tilde{z}}{\tilde{w}}$$

$$x \equiv \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{w}x \\ \tilde{w}y \\ \tilde{w}z \\ \tilde{w} \end{bmatrix} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ \tilde{w} \end{bmatrix} = \tilde{x}$$

- 3D-point can be directly presented as 4D point
- $\tilde{x}$, where now the fictitious co-ordinate is again $\tilde{w}$ are right here.
- We are going to use in our linear camera model.

Perspective projection:

$$\boxed{u = f_x \cdot \frac{x_c}{z_c} + O_x} \qquad \boxed{v = f_y \cdot \frac{y_c}{z_c} + O_y}$$

Multiplying by z_c

Homogeneous co-ordinates of (u, v):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} \quad \text{where, } (u, v) = (\tilde{u}/\tilde{w}, \tilde{v}/\tilde{w})$$

$$\equiv \begin{bmatrix} z_c \cdot u \\ z_c \cdot v \\ z_c \end{bmatrix} \equiv \begin{bmatrix} f_x \cdot x_c + z_c O_x \\ f_y\, y_c + z_c O_y \\ z_c \end{bmatrix}$$

further simplify using Matrix →

Page No.	
Date	

3/4

$$\begin{bmatrix} f_x & 0 & O_x & 0 \\ 0 & f_y & O_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \leftarrow \text{(Intrinsic Matrix)}$$

[it has all internal parameters)

→ This includes all the Internal parameters of the camera effects F, Y, O, X, Y multiplyed by the homogeneous co-ordinate of the 3-D point defined on the camera.

→ co-ordinate frame xy, y, c, zc

X ⟶ Linear model for perspective projection — X

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} = \begin{bmatrix} f_x & 0 & O_x & 0 \\ 0 & f_y & O_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

Calibration Matrix:

$$K = \begin{bmatrix} f_x & 0 & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix}$$ (upper Right triangular Matrix) (below diagonals = 0)

Intrensic Matrix, Mint

→ is a concatenation of K, this K matrix & column of zeroes right here.

$$M_{int} = [K|0] = \begin{bmatrix} f_x & 0 & O_x & 0 \\ 0 & f_y & O_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Page No.
Date

Mapping of a point from the world coordinate to the camera co-ordinates: (3D - 3D)

This can be done by using the position & orientation of the camera co-ordinate frame. With respect to the world co-ordinate frame, the position, c_w, a vector & the orientation is given by a matrix R.

Extrinsic parameters:

Position c_w & Orientation R of the camera in the world co-ordinate frame w are the camera's Extrinsic parameters.

R is the Rotation Matrix given by :

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