COMS30030
Image Processing and Computer Vision

Part II: Stereo and Motion

Andrew Calway

andrew@cs.bris.ac.uk

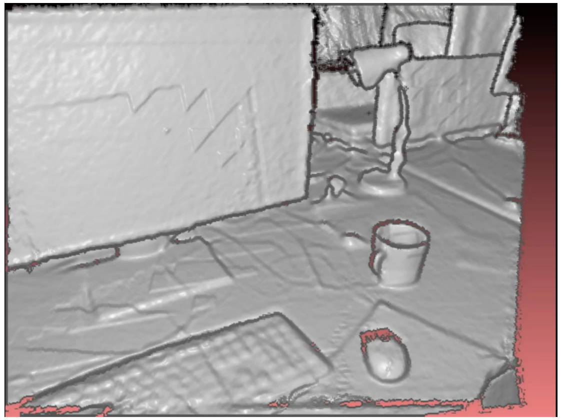
Andrew Calway

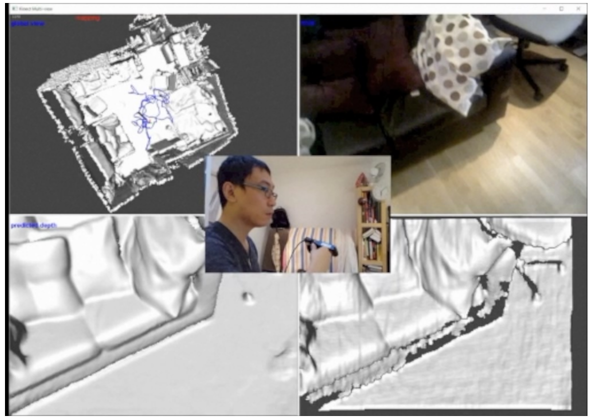
Professor of Computer Vision

www.cs.bris.ac.uk/~andrew

andrew@cs.bris.ac.uk

Computer Vision and Robotics

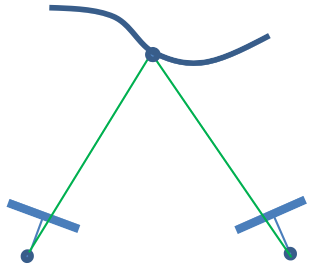




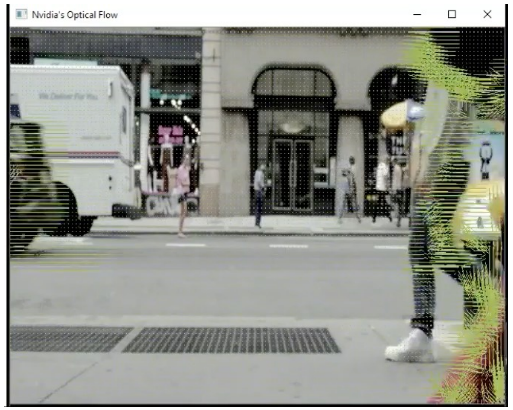
Vision-Based Simultaneous Localisation and Mapping (SLAM)

Stereo and Motion

- **Stereo**
  - Determining scene depth information from 2 (or more) images captured from different viewpoints
  - Geometry, correspondence matching and 3-D reconstruction

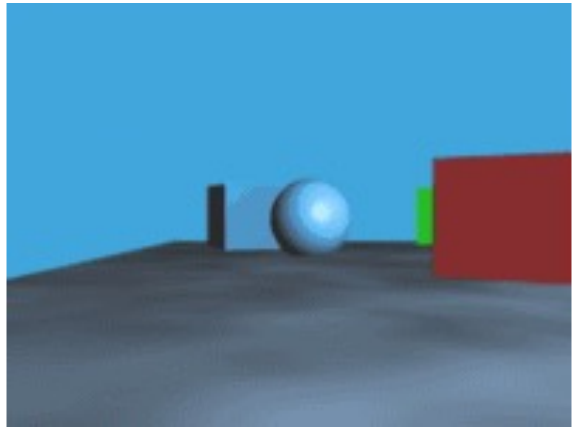


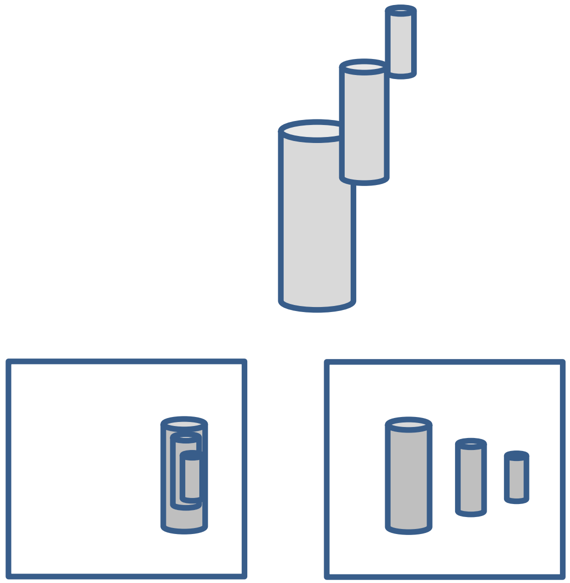
- **Motion**
  - Determining 2-D motion in video frames
  - Modelling, optical flow and motion estimation



Stereo Vision

- Stereo vision - 3-D from two images taken from different viewpoints.
- Objects appear in different positions in each viewpoint – **parallax**.
- Position of object in each image depends on its depth.
  - position difference (**disparity**) inversely proportional to depth
- If we know disparity & viewpoints
  - 3-D scene structure



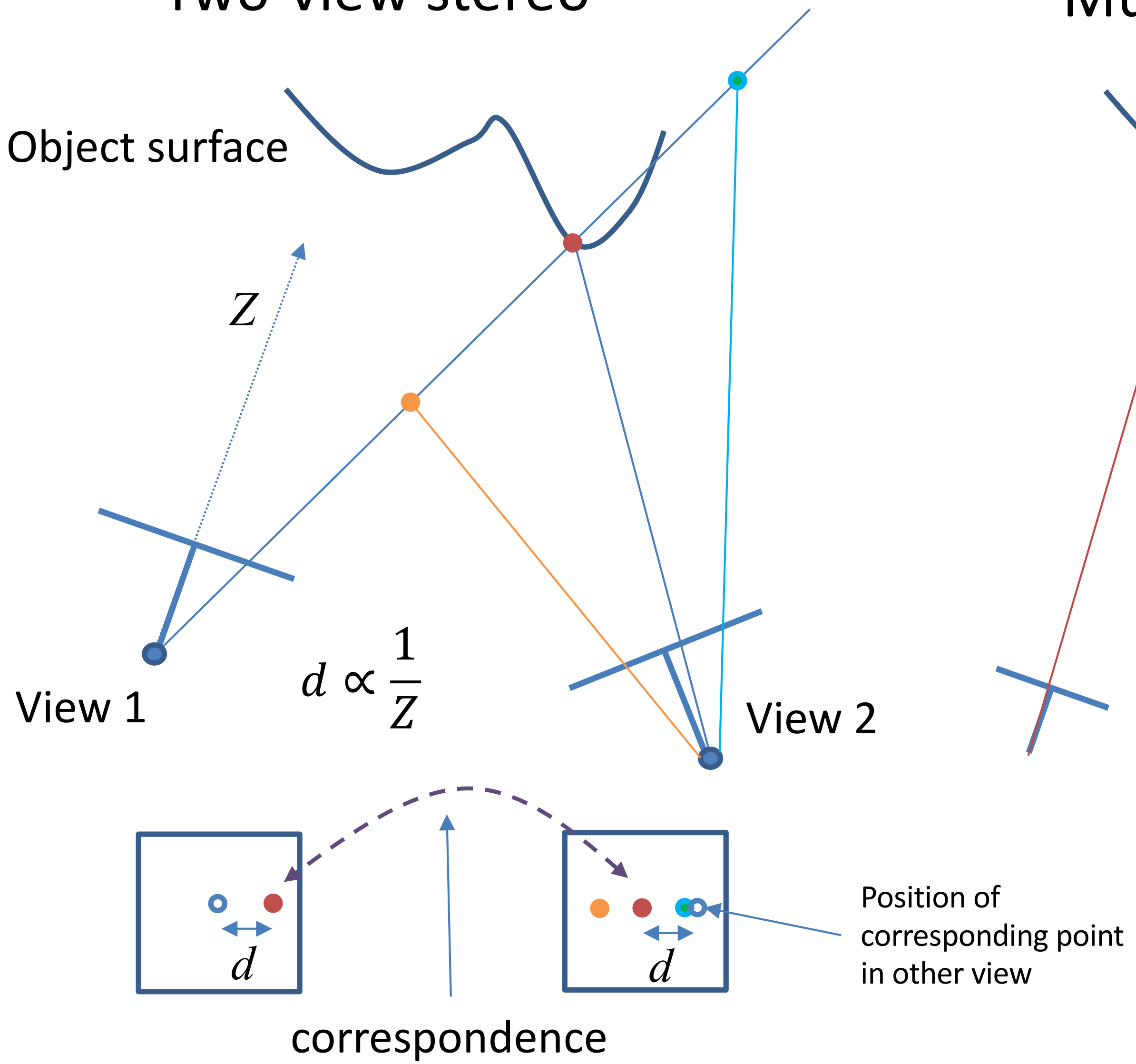


Colour animated gif by **Nathaniel Domek**
https://en.wikipedia.org/wiki/Parallax#/media/File:Parallax.gif

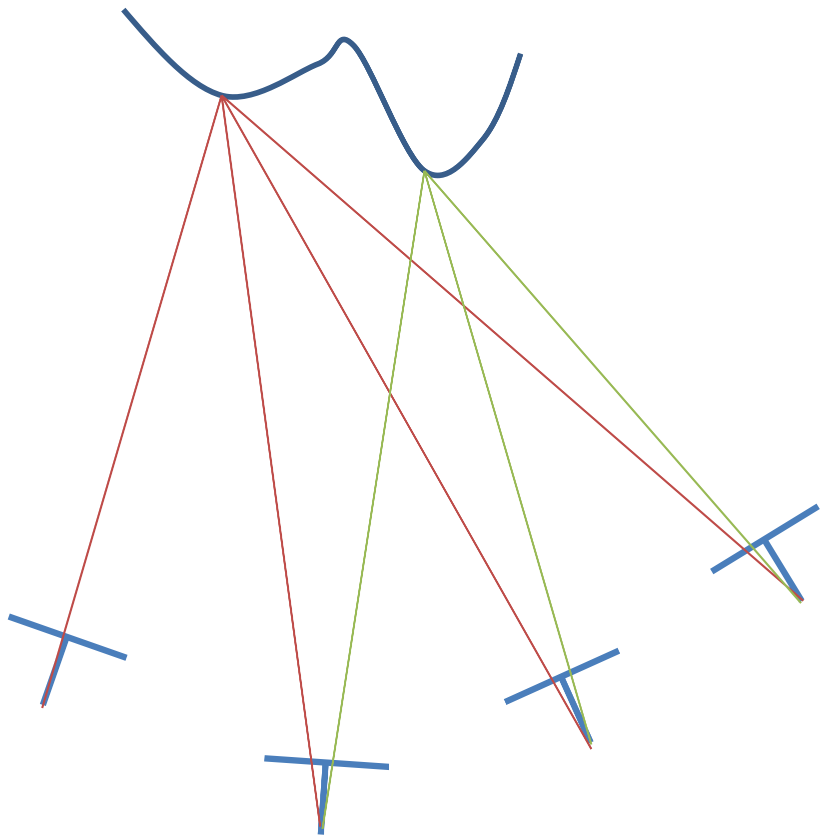
Two-View Stereo Examples

Stereo Computer Vision

Two-view stereo



Multi-view stereo



Similar to SLAM and Structure-from-Motion (SFM)

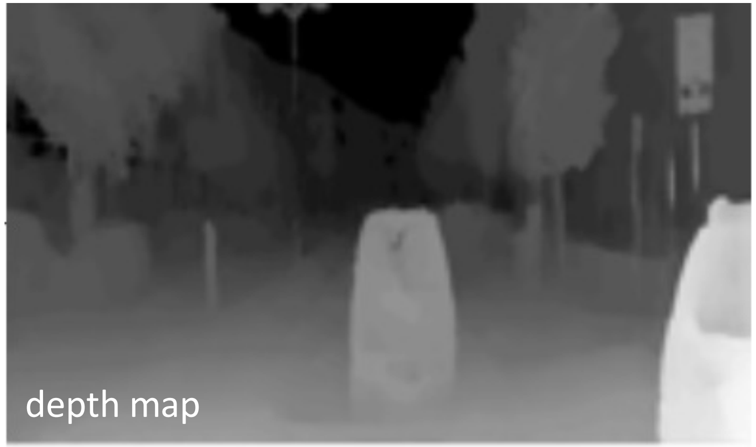
Three Problems of Stereo

- **Geometry** – determine relative position and orientation of the cameras
- **Correspondence** – determine matching points in the stereo views
- **Reconstruction** – determine 3D location in scene of matched points via triangulation

all interrelated

Stereo Vision – SOTA Examples

Two view



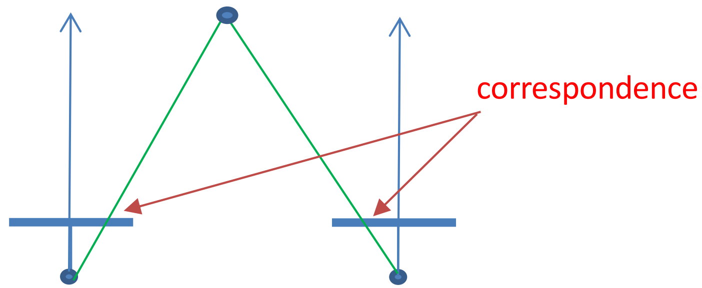
Group-wise Correlation Stereo Network,
Guo *et al*, CVPR 2019

Multi-view

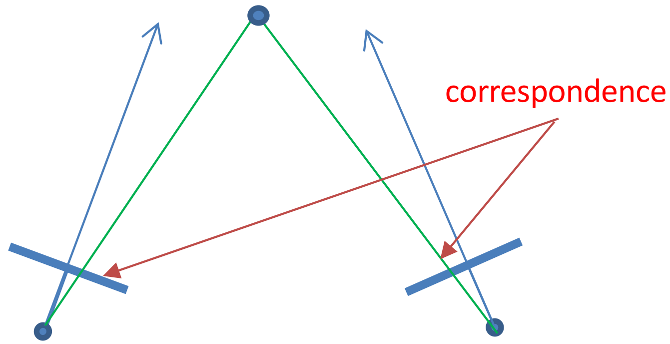
Depth from Gradients in Dense Light Fields for
Object Reconstruction, Yucer *et al*, 3DV 2016

Stereo Geometry

- Need to understand geometric relationship between cameras to allow 3-D reconstruction from correspondences
- Simple two-view stereo - coplanar image planes – geometry defined by similar triangles

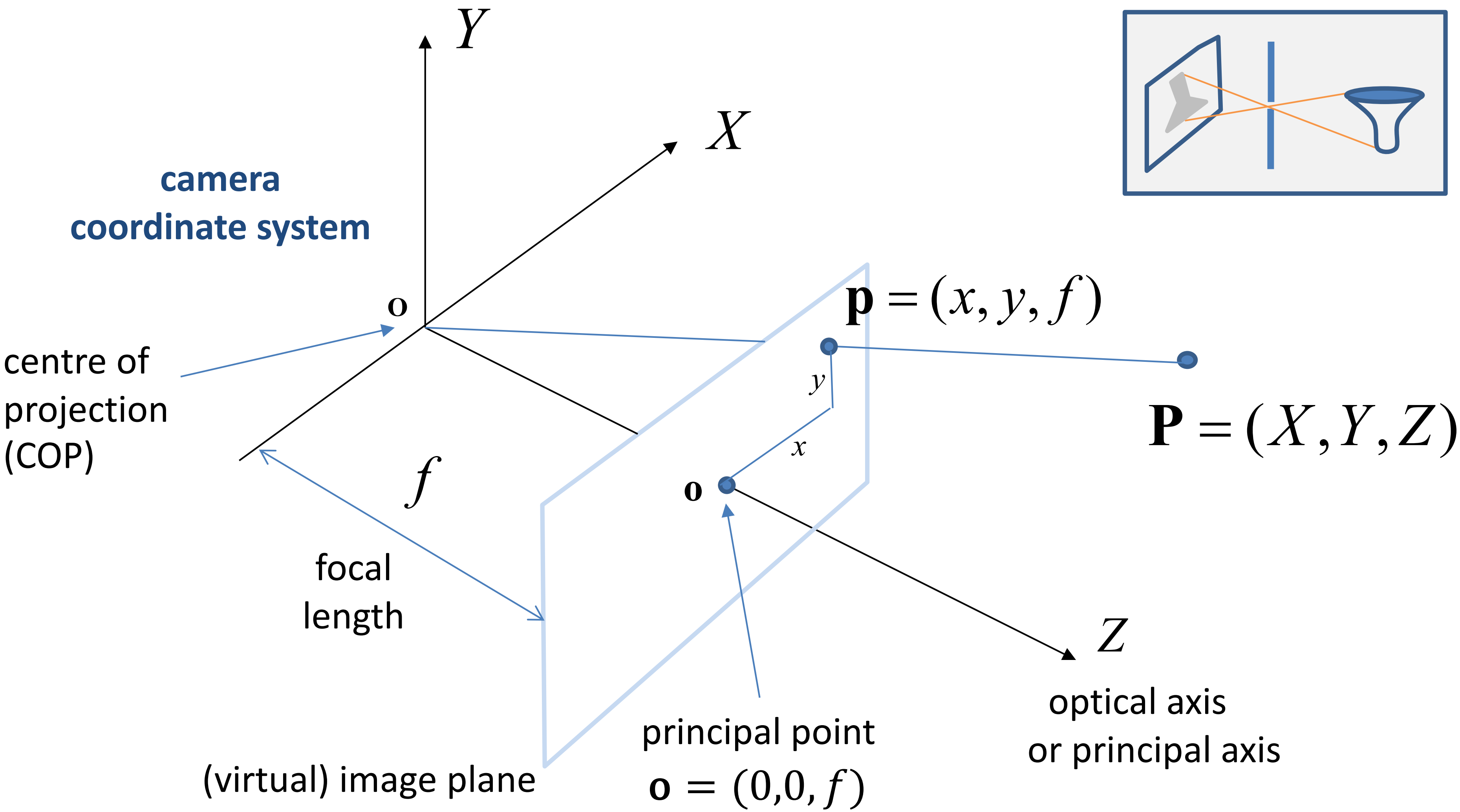


- General stereo – geometry depends on **position** and **orientation** of cameras
  - **epipolar geometry**



- But we also a need camera model …..

Pin Hole Camera Model

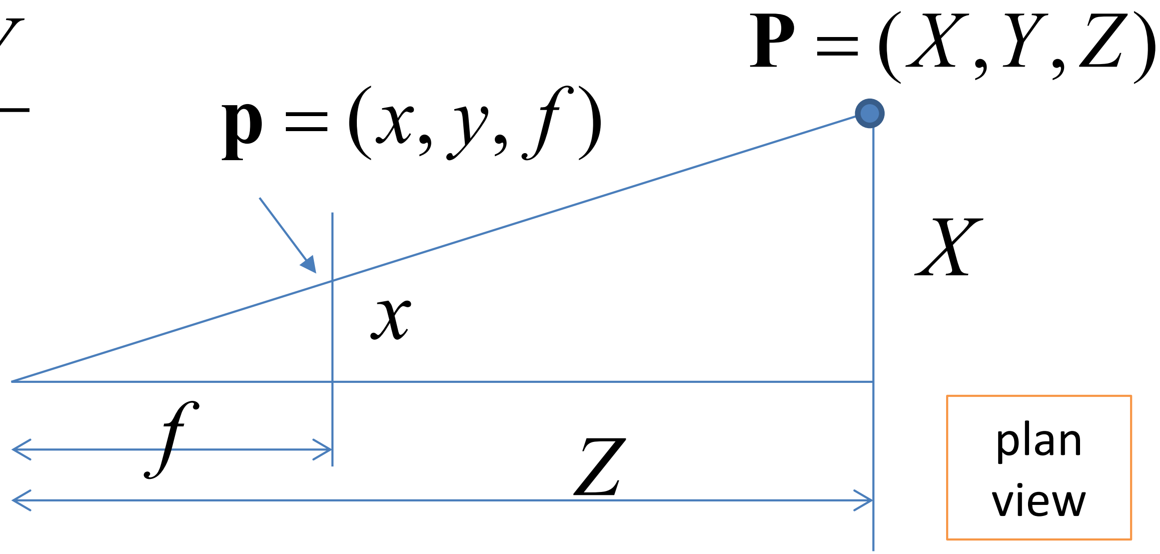


Perspective Projection Equations

- 3D point: $\mathbf{P} = (X, Y, Z)$ (on surface of object)
- Projects to 2D point: $\mathbf{p} = (x, y, f)$ (in image)
- Using similar triangles (pinhole model):

$$x = \frac{fX}{Z} \qquad y = \frac{fY}{Z}$$

$$\mathbf{p} = f\,\mathbf{P}/Z$$



Simple Two-View Stereo

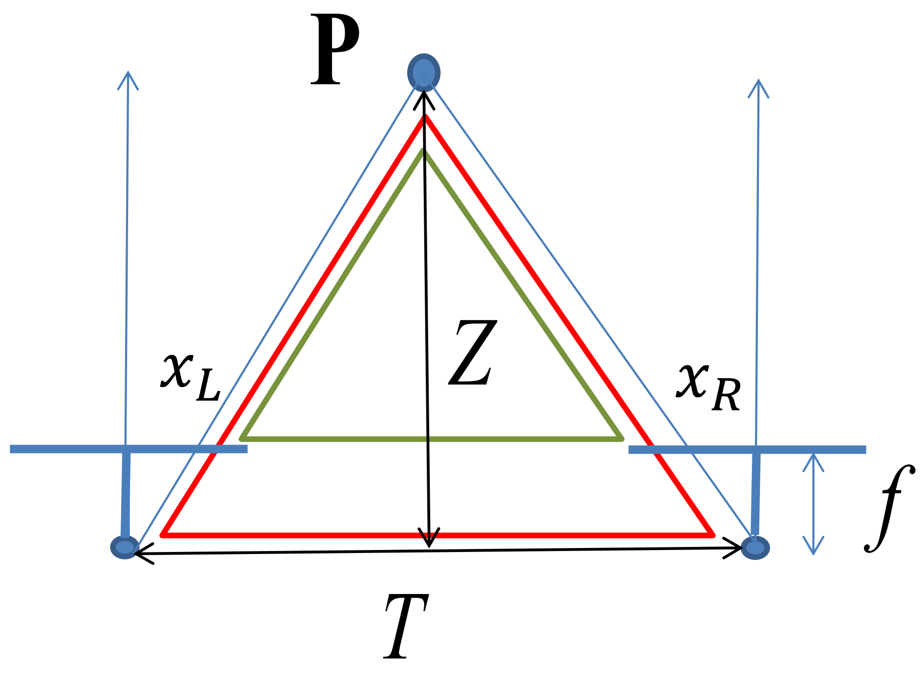
- Coplanar image planes, COPs in X-Z plane
- T – **baseline**, distance between COPs
- Similar triangles:

Point P has depth Z

$$\boxed{\frac{T}{Z}} = \boxed{\frac{T - x_L + x_R}{Z - f}}$$

- Rearrange for depth:

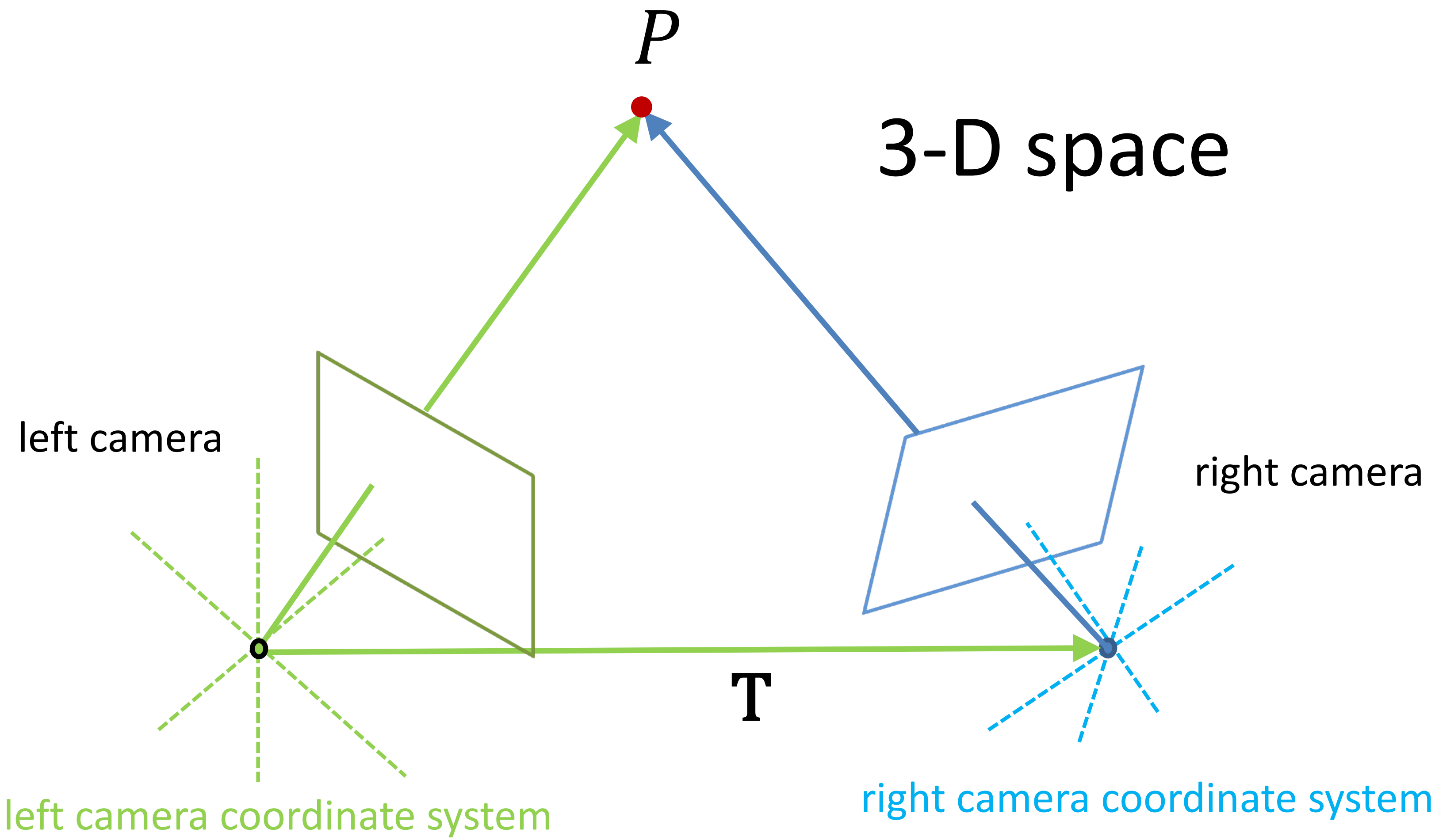
$$\boxed{Z = \frac{fT}{x_L - x_R} = \frac{fT}{d}}$$



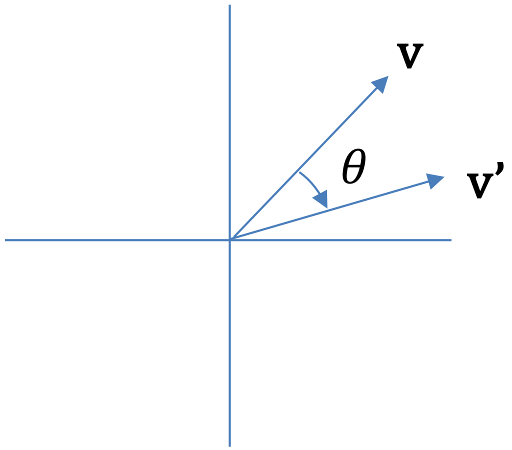
d - disparity

focal length f

General Two-View Stereo



Rotation Matrices

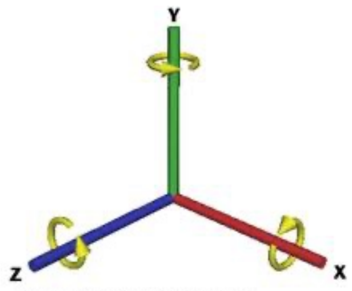


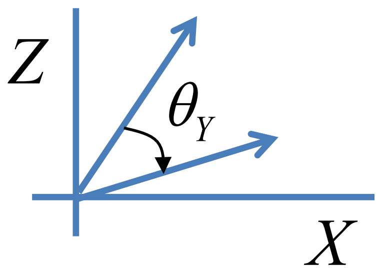
2-D clockwise rotation: 2x2 matrix

$$\mathbf{v}' = \begin{bmatrix} v'_x \\ v'_y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = R(\theta)\mathbf{v}$$

3-D rotation composed of rotations around X, Y and Z axes:

3x3 matrix: $R = R_X R_Y R_Z$

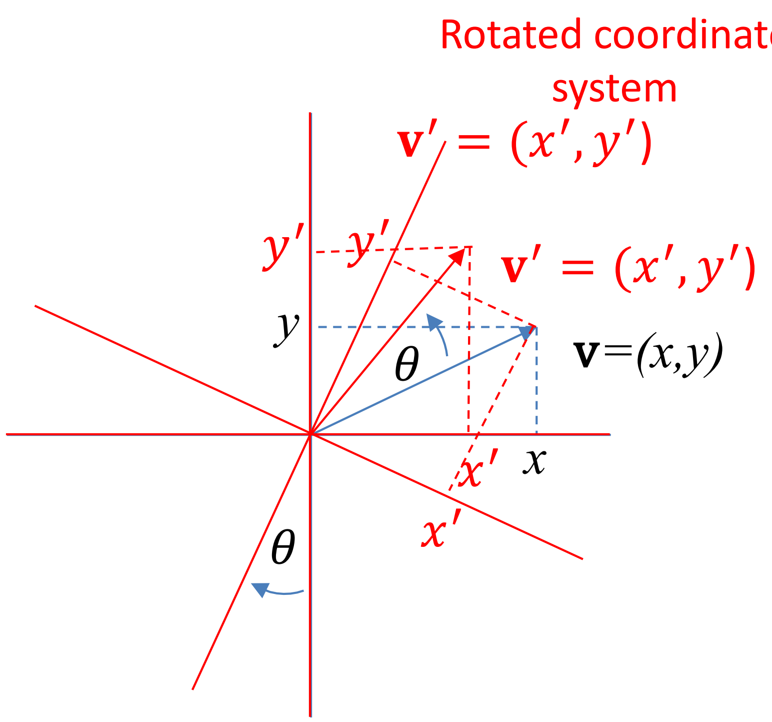




$$R_Y\mathbf{P} = \begin{bmatrix} \cos\theta_Y & 0 & \sin\theta_Y \\ 0 & 1 & 0 \\ -\sin\theta_Y & 0 & \cos\theta_Y \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Clockwise rotation about Y axis

2-D Coordinate Transformations



$$\mathbf{v}' = R(-\theta)\mathbf{v}$$

Vector representation in rotated coordinate system