COMS30030
Image Processing and Computer Vision

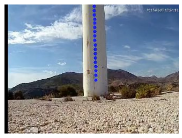
Part II: Stereo and Motion

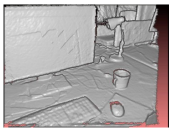
Andrew Calway

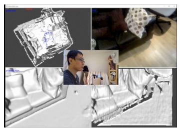
andrew@cs.bris.ac.uk

1

Andrew Calway
Professor of Computer Vision
www.cs.bris.ac.uk/~andrew
andrew@cs.bris.ac.uk
Computer Vision and Robotics





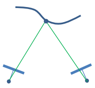


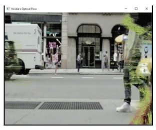
Vision-Based Simultaneous Localisation and Mapping (SLAM)

2

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

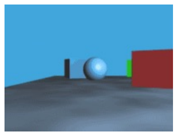


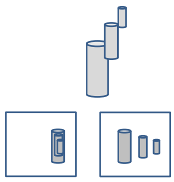


3

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





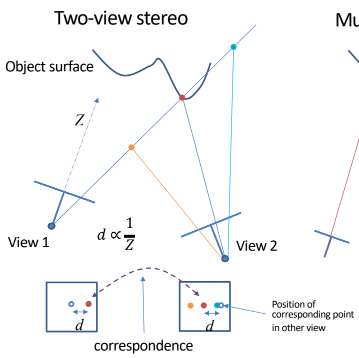
Colour animated gif by **Nathaniel Domek**
https://en.wikipedia.org/wiki/Parallax#/media/File:Parallax.gif

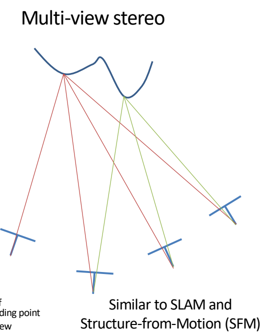
4

Two-View Stereo Examples

5

Stereo Computer Vision





6

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

7

Stereo Vision – SOTA Examples

Two view

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

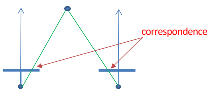
Multi-view

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

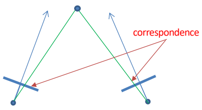
8

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

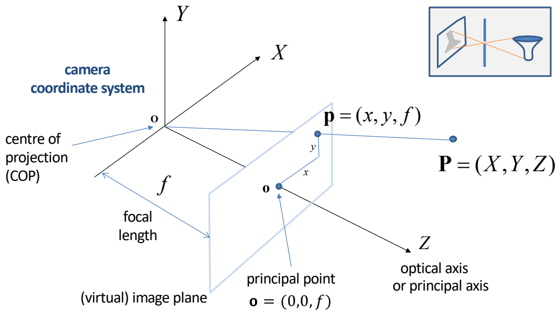


- General stereo – geometry depends on **position** and **orientation** of cameras
  – **epipolar geometry**
- But we also a need camera model .....



9

Pin Hole Camera Model



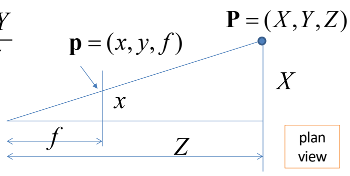
10

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):

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

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



11

Simple Two-View Stereo

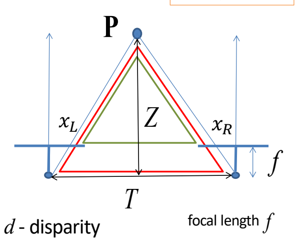
- Coplanar image planes, COPs in X-Z plane
- T – **baseline**, distance between COPs
- Similar triangles:

Point P has depth Z

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

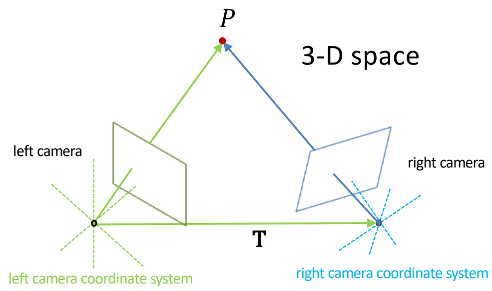
- Reorganising for depth:

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



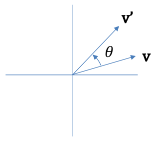
12

General Two-View Stereo



13

Rotation Matrices



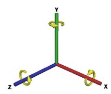
2-D counter-clockwise rotation: 2x2 matrix

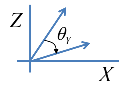
$$\mathbf{v}' = \begin{bmatrix} v'_x \\ v'_y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = R(\theta)\mathbf{v}$$

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

3x3 matrix: $R = R_X R_Y R_Z$

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

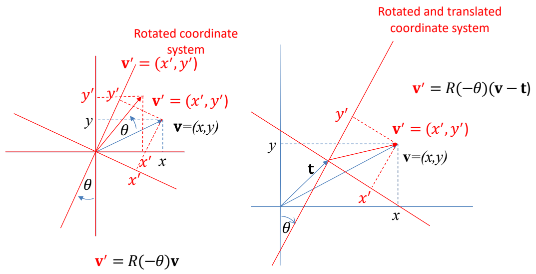




Clockwise rotation about Y axis

14

2-D Coordinate Transformations



Vector representation in rotated coordinate system

15

3-D Coordinate Transformations

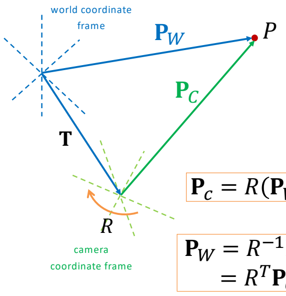
$\mathbf{P}_W$ Vector defining P in world coordinates

$\mathbf{P}_C$ Vector defining P in camera coordinates

$\mathbf{T}$: 3-D camera position vector

R : 3-D camera rotation matrix

R defines rotation to be applied to camera coordinate system to align it with world coordinate system



$\mathbf{P}_C = R(\mathbf{P}_W - \mathbf{T})$

$$\mathbf{P}_W = R^{-1}\mathbf{P}_C + \mathbf{T} = R^T\mathbf{P}_C + \mathbf{T}$$

Rotation matrices: $R^T = R^{-1}$

16

Homogeneous Coordinates

- Homogeneous coordinates allow coordinate transformations to be defined by 4x4 matrices:

$$\mathbf{P}'_W = \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{P}_W \\ 1 \end{bmatrix} = \begin{bmatrix} R^T & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P}_C \\ 1 \end{bmatrix} = H_{CW}\mathbf{P}'_C \Rightarrow \boxed{\mathbf{P}_W = R^T\mathbf{P}_C + \mathbf{T}}$$

$$H_{CW} = \begin{bmatrix} R_{00} & R_{10} & R_{20} & T_x \\ R_{01} & R_{11} & R_{21} & T_y \\ R_{02} & R_{12} & R_{22} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{P}'_C = H_{CW}^{-1}\mathbf{P}'_W = H_{WC}\mathbf{P}'_W$$

$$H_{WC} = \begin{bmatrix} R & -R\mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \Rightarrow H_{WC}H_{CW} = I$$

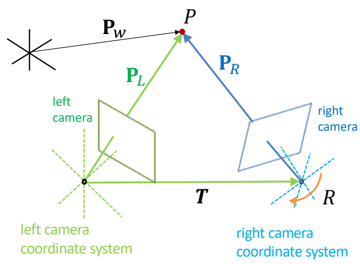
I - Identity matrix

$$\Rightarrow \boxed{\mathbf{P}_C = R\,(\mathbf{P}_W - \mathbf{T})}$$

H_{cw} : camera to world coordinate transformation matrix

17

Stereo Coordinate Systems



$$\mathbf{P}'_L = H_{WL}\mathbf{P}'_W$$
$$\mathbf{P}'_R = H_{WR}\mathbf{P}'_W$$
$$\mathbf{P}'_L = H_{WL}H_{WR}^{-1}\mathbf{P}'_R$$
$$\mathbf{P}'_L = H_{RL}\mathbf{P}'_R$$

$$H_{RL} = \begin{bmatrix} R^T & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}$$

$\boxed{\mathbf{P}_L = R^T\mathbf{P}_R + \mathbf{T}}$

$\boxed{\mathbf{P}_R = R(\mathbf{P}_L - \mathbf{T})}$

R defines rotation to be applied to right camera coordinate system to align it with left coordinate system

18