Course Review

Week 12

Announcements

- **Drop-in Sessions:** Please bring your assignment/exam questions to the following locations
  - 09:00-10:30 Friday **24 May**, Melville Hall (Dylan)
  - 15:00-17:00 Friday **24 May**, 1.23 Hanna Neumann (Yiran)

Semester 1 SELT is live!

21 – 24 May
Food and prizes just for completing your SELT!

Join us for
free food and prizes
21 – 24 May
11.00am – 2.00pm
CSIT Lab entrance or lawn
(weather dependent)

20 May – Survey opens
Check your email or Wattle page for available surveys

Survey runs for 4 weeks
Please provide constructive and respectful feedback *(your teacher can't identify you)*

16 June - Survey closes
IR team perform screening of comments for welfare concerns

8 July
SELT feedback is made available to teachers and course convenors to improve future course delivery

Australian National University

Semester 1 SELT - survey journey

The Student Experience of Learning & Teaching survey allows students to give feedback on their courses and teachers. It is **voluntary** and **confidential**, and run by the Institutional Research (IR) team.

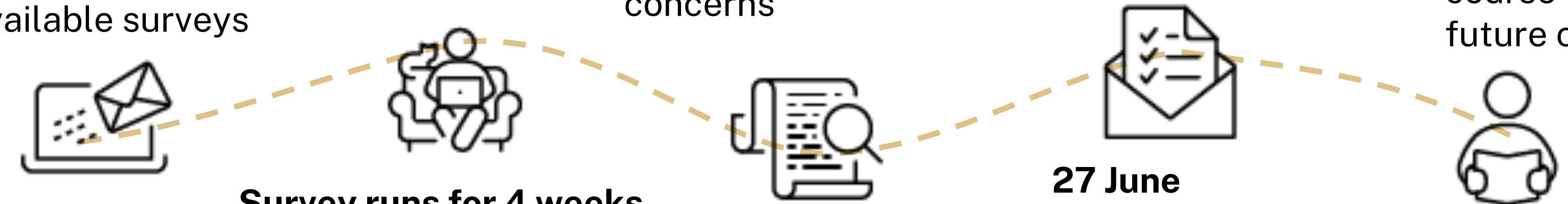
20 May – Survey opens
Check your email or Wattle page for available surveys

Survey runs for 4 weeks
Please provide constructive and respectful feedback *(your teacher can't identify you)*

16 June - Survey closes
IR team perform screening of comments for welfare concerns

27 June
Grades are released to students

8 July
SELT feedback is made available to teachers and course convenors to improve future course delivery



Australian National University

Find out more on the *Info for Students* webpage:

https://services.anu.edu.au/learning-teaching/education-data/student-experience-of-learning-teaching-selt/information-for

SELT - Frequently asked questions

What kind of feedback is helpful?

Think about your experience of the course and teaching, and what worked or didn't work for you.

When writing feedback, focus on respectful and constructive language – if you were a teacher, what type of feedback would help you improve the class?

Can teachers see who left specific feedback?

SELT is confidential, and teachers cannot see, or ask to see, the identity of a respondent. Unless you self-identify, for example by using names or describing specific events, teachers cannot identify you.

Australian National University

Find out more on the *Info for Students* webpage:

https://services.anu.edu.au/learning-teaching/education-data/student-experience-of-learning-teaching-selt/information-for

Weekly Study Plan: Overview

Wk	Starting	Lecture	Lab	Assessment
1	19 Feb	Introduction	✗	
2	26 Feb	Low-level Vision 1	1	
3	4 Mar	Low-level Vision 2	1	
		Mid-level Vision 1		
4	11 Mar	Mid-level Vision 2	1	CLab1 report due Friday
		High-level Vision 1		
5	18 Mar	High-level Vision 2	2	
6	25 Mar	High-level Vision 3[1]	2	
	1 Apr	Teaching break	✗	
	8 Apr	Teaching break	✗	
7	15 Apr	3D Vision 1	2	CLab2 report due Friday
8	22 Apr	3D Vision 2	3	
9	29 Apr	3D Vision 3	3	
10	6 May	3D Vision 4	3	
		Mid-level Vision 3		
11	13 May	High-level Vision 4	✗	CLab3 report due Friday
12	20 May	Course Review	✗	

Weekly Study Plan: Part B

Wk	Starting	Lecture	By
7	15 Apr	3D vision: introduction, camera model, single-view geometry	Dylan
8	22 Apr	3D vision: camera calibration, two-view geometry (homography)	Dylan
9	29 Apr	3D vision: two-view geometry (epipolar geometry, triangulation, stereo)	Dylan
10	6 May	3D vision: multiple-view geometry	Weijian
		Mid-level vision: optical flow, shape-from-X	Dylan
11	13 May	High-level vision: self-supervised learning, detection, segmentation	Dylan
12	20 May	Course review	Dylan

Monocular Scene Reconstruction @ 10FPS

Outline

1. Course Review (cont’d)
2. Practice exam questions
3. Q&A
4. Drop-in session

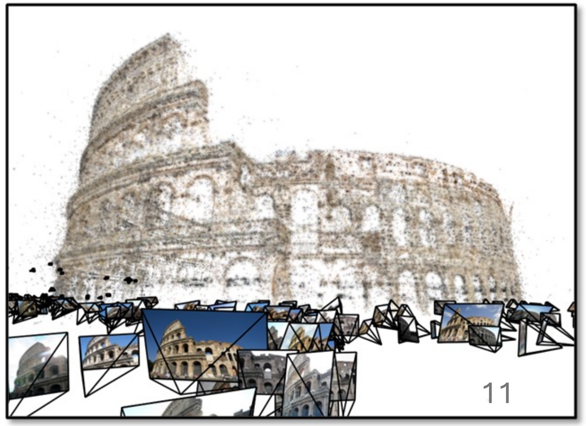
Course Review

Continued

Structure From Motion

- Core Steps of SfM

Feature detection → **Feature matching**: Find scene points seen by multiple cameras → **Reconstruction**: Robustly estimate camera poses and triangulation → **Refinement: BA**: Refine camera poses **R, T** and scene structure $\{X_i\}$



Optical Flow

- **Problem:** Compute a flow field that takes pixels in the first image to their location in the second image.

$$(dx, dy) = f\big(I(t), I(t + dt)\big)_{(x,y)}$$

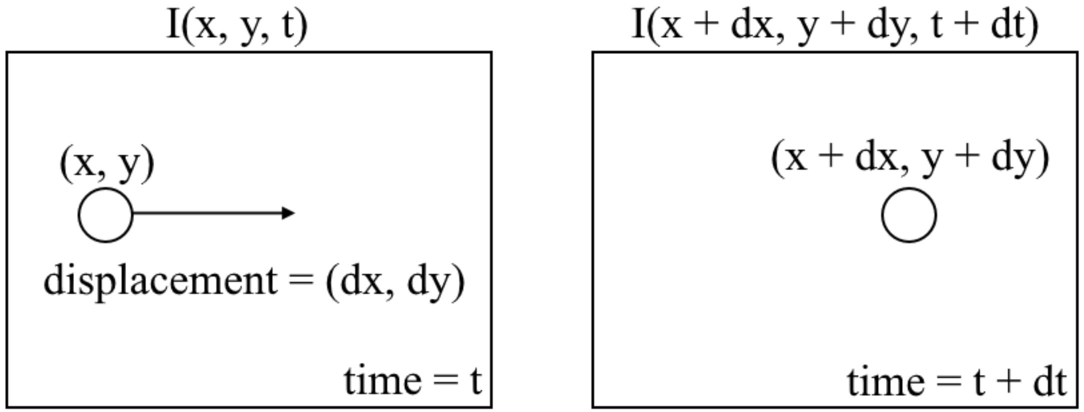


Image Credit: Chuan-en Lin

Brightness/Colour Constancy Assumption

$$I_x u + I_y v = -I_t \;\leftrightarrow\; \frac{\partial I}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial I}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{\partial I}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}t}$$

- From first-order Taylor expansion of $I(x, y, t) = I(x + \mathrm{d}x, y + \mathrm{d}y, t + \mathrm{d}t)$
  - Note that all partial derivatives are evaluated at (x, y, t), e.g., $I_x(x, y, t)$
- We have one equation and two unknowns (u, v): the optical flow
- Assumes a smooth visual gradient over the area of motion

Lucas–Kanade Optical Flow Algorithm

$$\begin{bmatrix} I_x & I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -I_t$$

- Least squares solution: $\boldsymbol{u}^\star = \operatorname{argmin}_{\mathbf{u}} \|A\boldsymbol{u} - \boldsymbol{b}\|^2 = (A^\top A)^{-1} A^\top \boldsymbol{b}$
- $A^\top A = \begin{bmatrix} \sum_{\boldsymbol{p}\in\mathcal{P}} I_x I_x & \sum_{\boldsymbol{p}\in\mathcal{P}} I_x I_y \\ \sum_{\boldsymbol{p}\in\mathcal{P}} I_y I_x & \sum_{\boldsymbol{p}\in\mathcal{P}} I_y I_y \end{bmatrix} \in \mathbb{R}^{2\times 2}$ (autocorrelation matrix)
- $A^\top \boldsymbol{b} = -\begin{bmatrix} \sum_{\boldsymbol{p}\in\mathcal{P}} I_x I_t \\ \sum_{\boldsymbol{p}\in\mathcal{P}} I_y I_t \end{bmatrix} \in \mathbb{R}^{2\times 1}$
  where the sum is over pixels $\boldsymbol{p}$ in patch $\mathcal{P}$
- You saw this in Harris corner detector!

Lucas & Kanade, 1981

Lucas–Kanade Optical Flow Algorithm

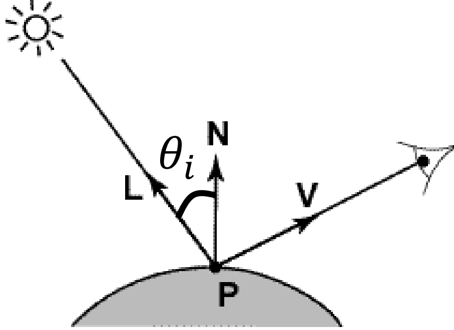
- When is $A^\top A\boldsymbol{u} = A^\top \boldsymbol{b}$ solvable?
  - $A^\top A$ invertible:
    - Determinant non-zero
  - $A^\top A$ not too small, otherwise estimate is sensitive to noise:
    - Eigenvalues λ_1 and λ_2 of $A^\top A$ should not be too small
  - $A^\top A$ well-conditioned:
    - ${}^{\lambda_1}/_{\lambda_2}$ should not be too large (λ_1: larger eigenvalue)
- Implications:
  - Harris corners are where λ_1 and λ_2 are both big; this is also when Lucas–Kanade optical flow estimation works best
    - Corners are good places to compute optical flow

Lucas & Kanade, 1981

Visual Cues: Shape-from-

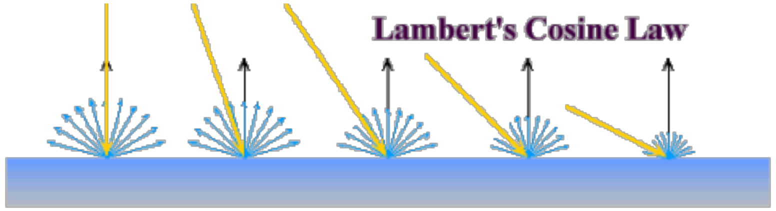
- Shading
- Texture
- Focus
- Motion
- Perspective distortion
- Colour
- Size
- Occlusion
- Stereo
- Specular highlights
- Inter-reflections
- Symmetry
- Light Polarisation
- Structured light (active)
- Time-of-flight (active)
- Shadow
- Silhouette

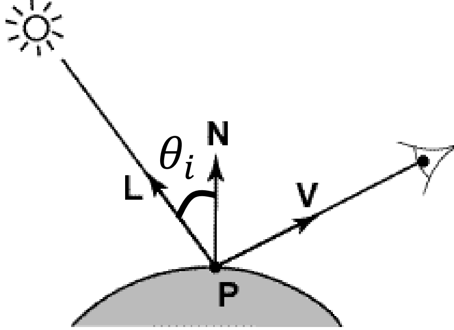
Shape-from-Shading



- Lambertian reflectance assumption gives us $I = N^\top L = \cos\theta$
- Not enough information to compute normal: 1 equation, 2 DoF
- Add additional information: e.g., smoothness
  - Variational shape-from-shading
- Then, convert normal maps to depth maps via integration

Lambertian Reflection





$$R_r = \rho N \cdot L R_i$$
$$I = \rho N \cdot L$$
$$\boxed{I = N \cdot L = \cos \theta_i}$$

- R_i: incident light intensity
- R_r: observed light intensity
- L: unit illuminant direction
- N: unit normal direction
- I: image intensity at point **P**
- ρ: material albedo

- Simplifying assumptions
  - $I = R_r$: camera response function f is the identity function
    - Can also assume linearly proportional
    - (If required, can perform radiometric calibration)
  - $R_i = 1$: light source intensity is 1
    - Can achieve this by dividing each pixel in the image by R_i
  - $\rho = 1$: material albedo is 1

Photometric Stereo

$$\underbrace{[I_1 \;\; I_2 \;\; I_3]}_{\substack{I \\ 1\text{x}3}} = \underbrace{\rho N^\top}_{\substack{G \\ 1\text{x}3}} \underbrace{[L_1 \;\; L_2 \;\; L_3]}_{\substack{L \\ 3\text{x}3}}$$

Image intensity matrix I is known

Light source matrix L is known

$$G = IL^{-1}$$

- ρ and N are unknowns (ρ may differ across surface)
- Surface normal: $N = \frac{G}{||G||}$
- Albedo: $\rho = ||G||$
- When is L invertible? >= 3 light directions are linearly independent
- More than 3 lights? Solve using least squares

Practice Exam Questions

Q&A + Drop-In Session

Thank you for taking the Computer Vision course!

What next?

Advanced Topics in Deep Learning for Computer Vision (COMP8536)

Advanced Topics in Computer Vision (COMP8539) [S2 2025]

Semester 1 SELT is live!

Join us for

free food and prizes

21 – 24 May

11.00am – 2.00pm

CSIT Lab entrance or lawn

(weather dependent)

20 May – Survey opens
Check your email or Wattle page for available surveys

21 – 24 May
Food and prizes just for completing your SELT!

Survey runs for 4 weeks
Please provide constructive and respectful feedback *(your teacher can't identify you)*

16 June - Survey closes
IR team perform screening of comments for welfare concerns

8 July
SELT feedback is made available to teachers and course convenors to improve future course delivery

Australian National University