COMP9517 Computer Vision

Introduction

What is Computer Vision?

Every picture tells a story

Computer vision automates and integrates many information processing and representation approaches useful for visual perception

What is computer vision?

Computer science perspective

Computer vision is the **interdisciplinary field** that develops **theories and methods** to allow computers **extract relevant information from digital images or videos**

Computer engineering perspective

Computer vision is the **interdisciplinary field** that develops **algorithms and tools** to **automate perceptual tasks** normally performed by the human visual system

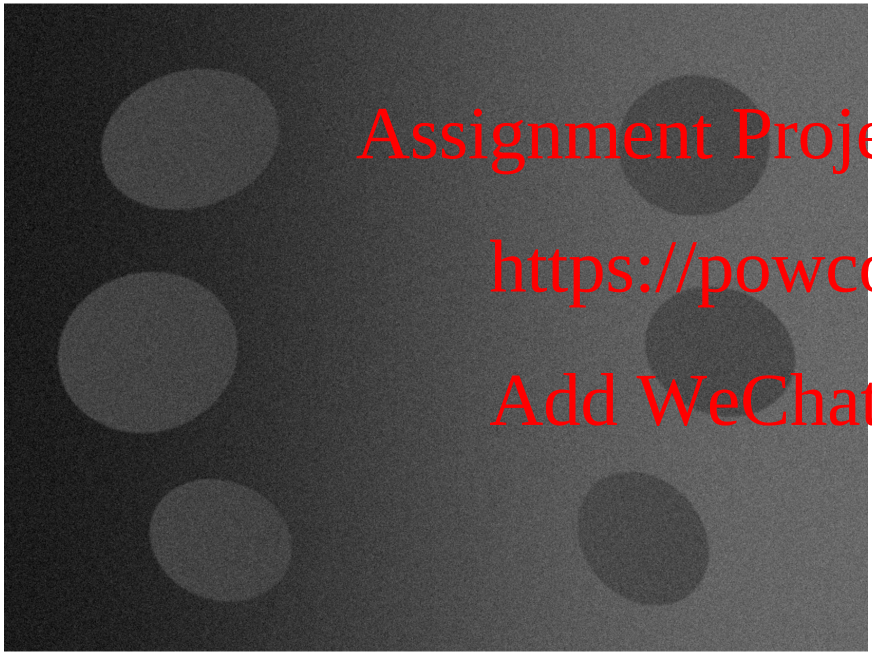
Can computers match (or beat) human vision?

Yes and no (but mostly no)

- Humans are much better at “hard” tasks

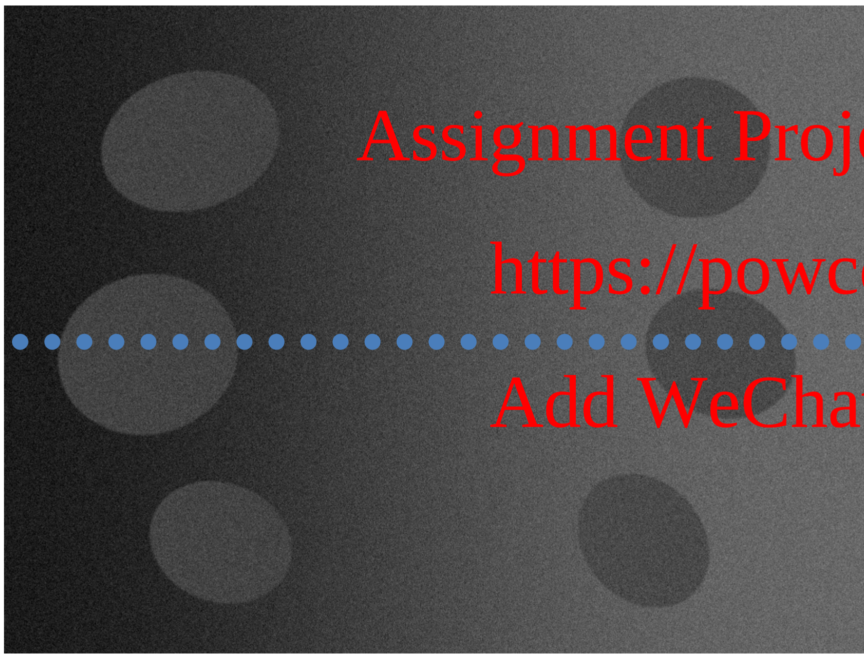
- Computers can be better at “easy” tasks

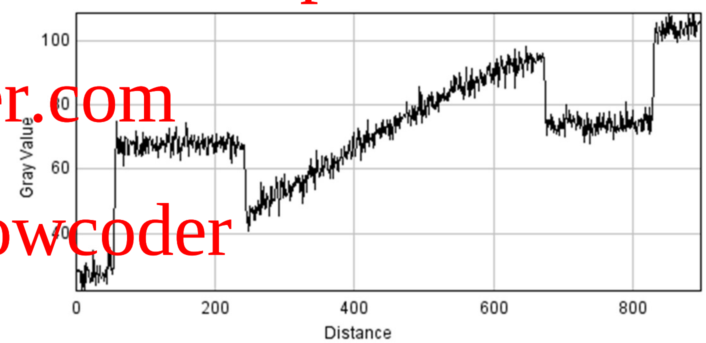
Human vision has its limitations…



Which objects are brighter?

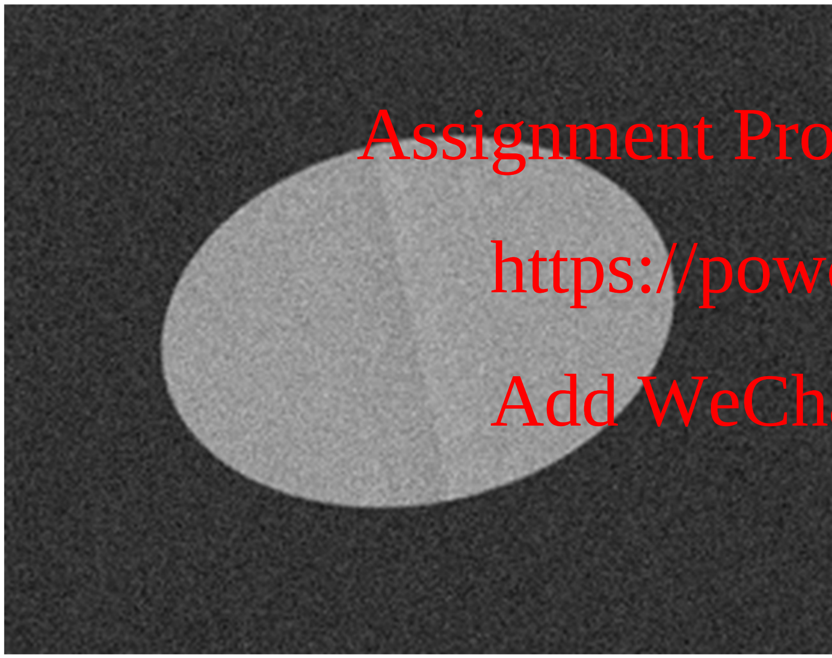
Human vision has its limitations…





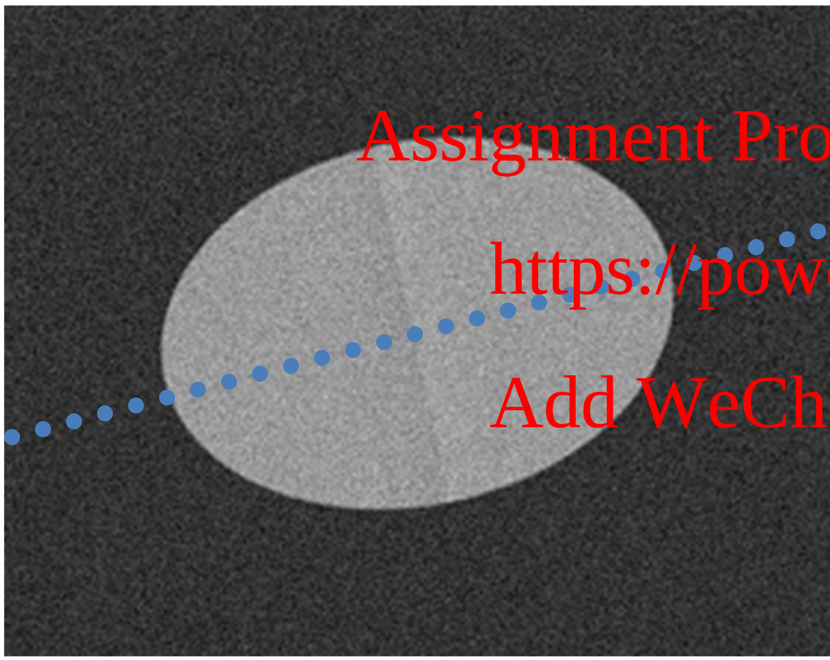
Which objects are brighter?

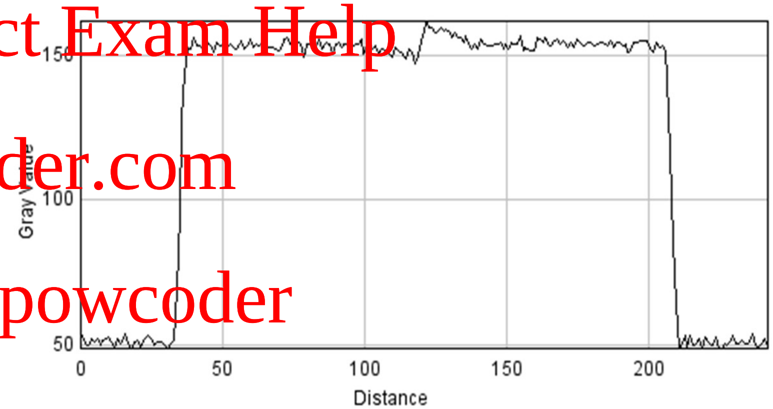
Human vision has its limitations…



Which side of this object is brighter?

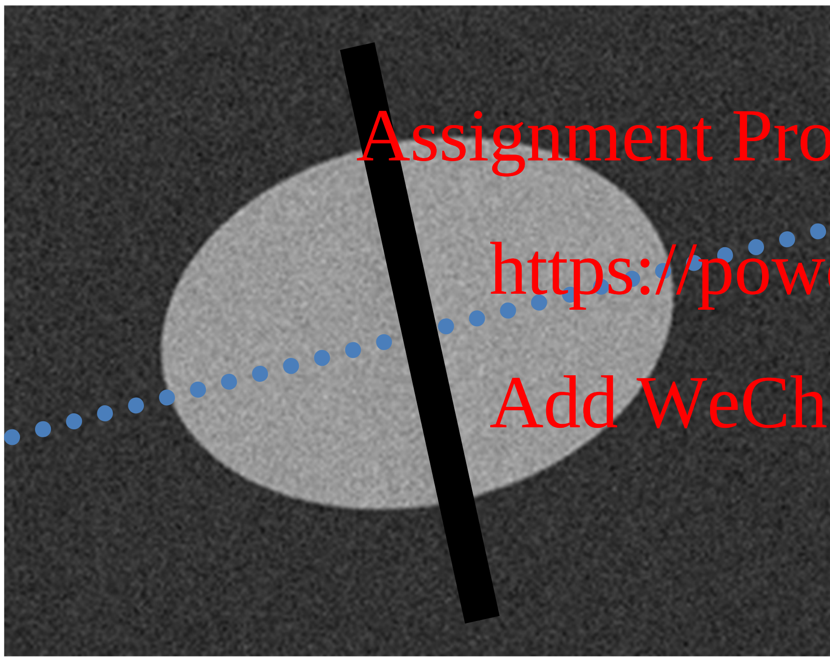
Human vision has its limitations…

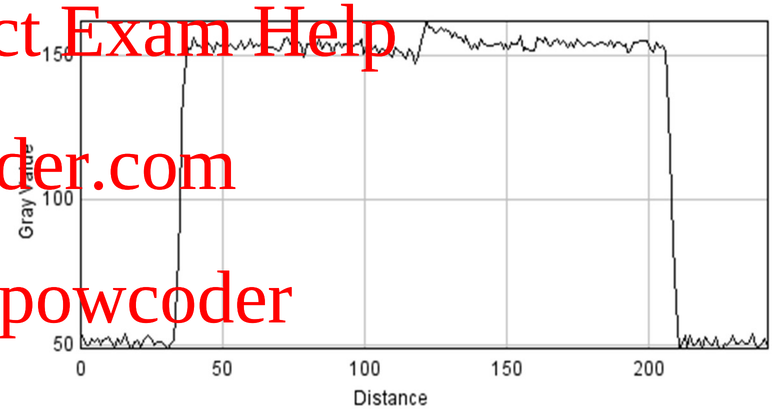




Which side of this object is brighter?

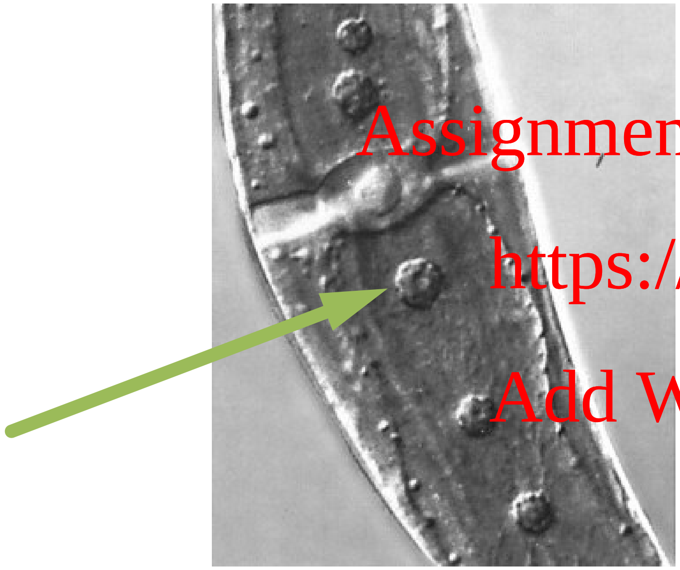
Human vision has its limitations…





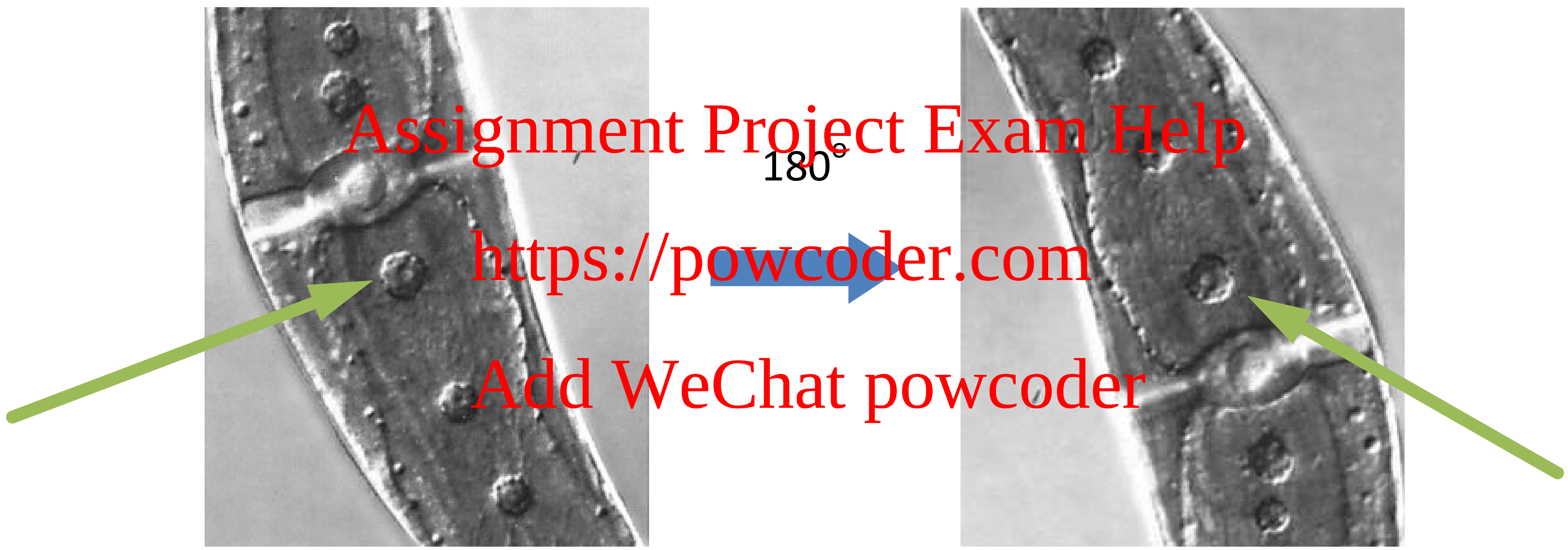
Which side of this object is brighter?

Human vision has its limitations…



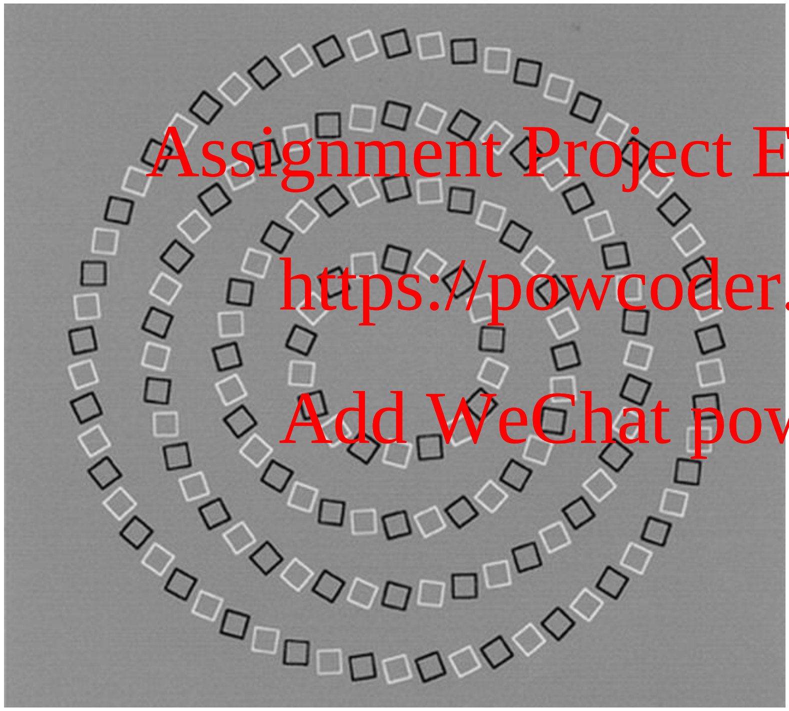
Are the cells popping in or out?

Human vision has its limitations…



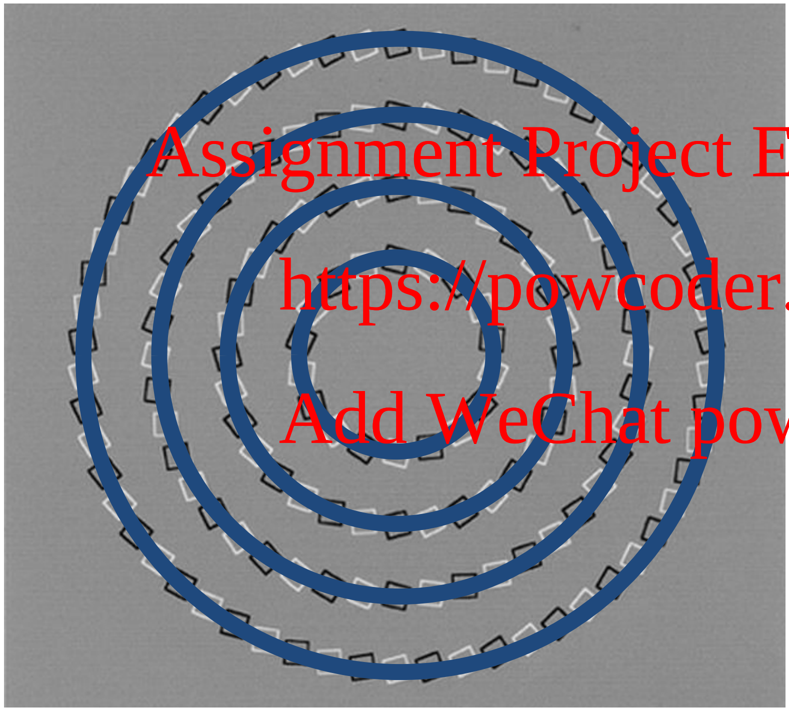
Are the cells popping in or out?

Human vision has its limitations…



What is the pattern you see?

Human vision has its limitations…



What is the pattern you see?

Human vision has its limitations…

What object do you see in this image?

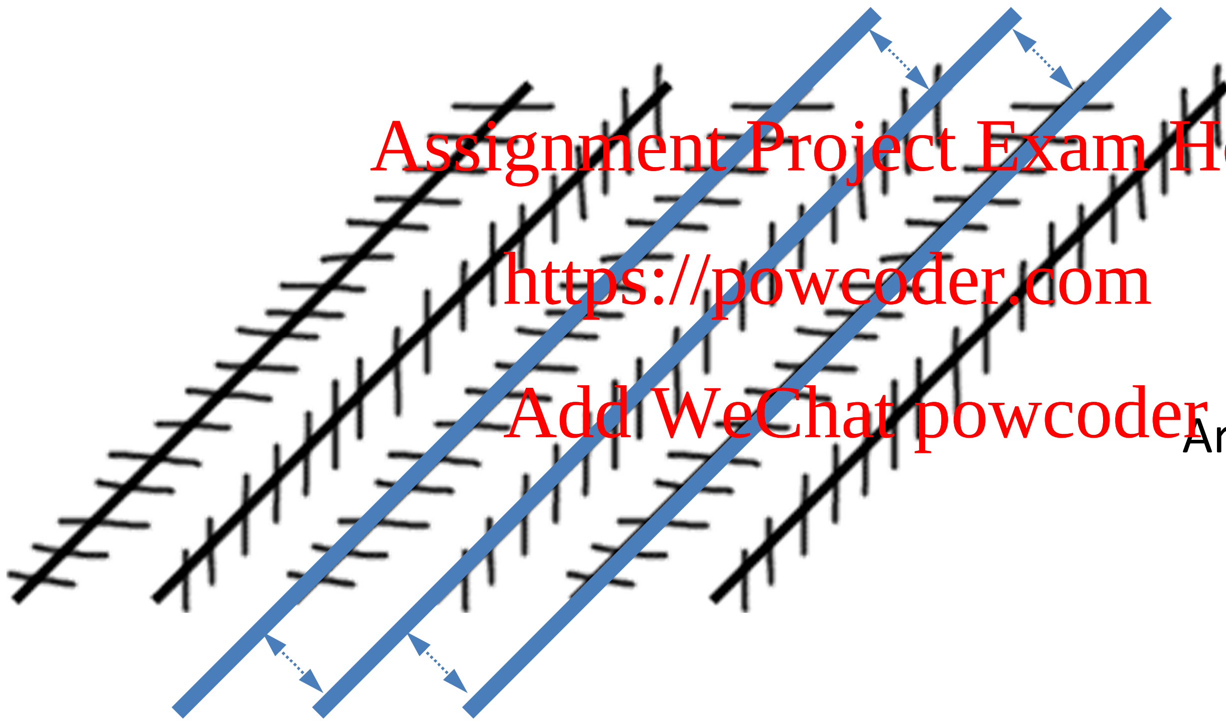
Human vision has its limitations…

What object do you see in this image?

Human vision has its limitations…

Are the main lines parallel??

Human vision has its limitations…

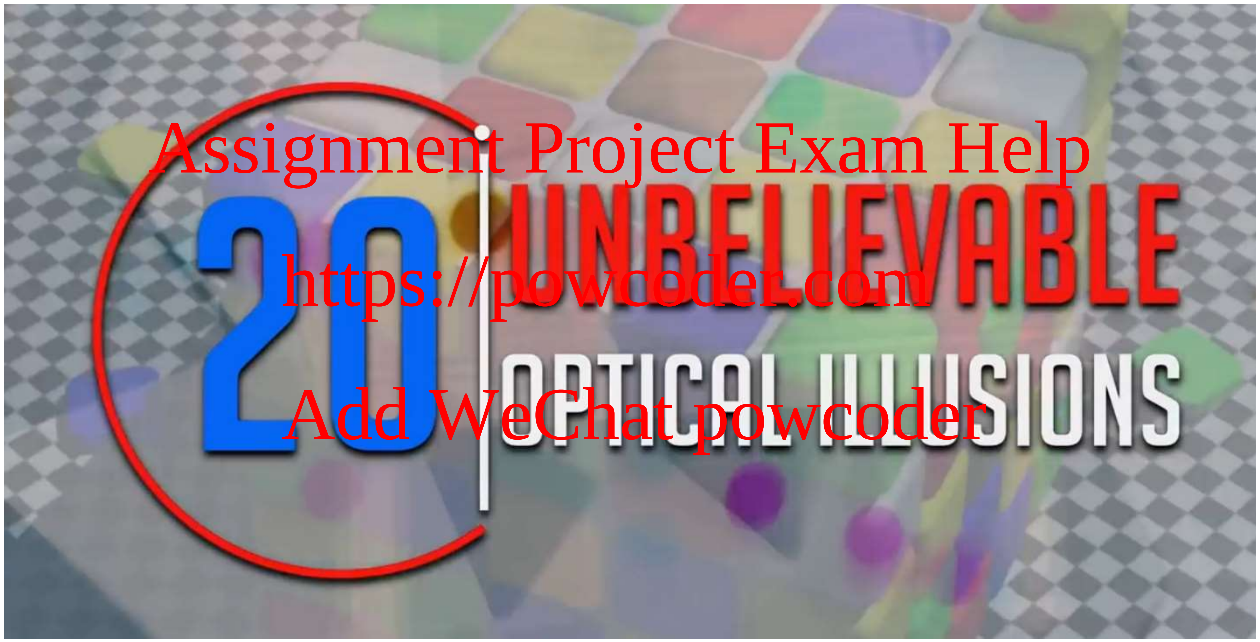


Are the main lines parallel?

Human vision has its limitations…

In which direction are these particles moving ?

Human vision has its limitations…

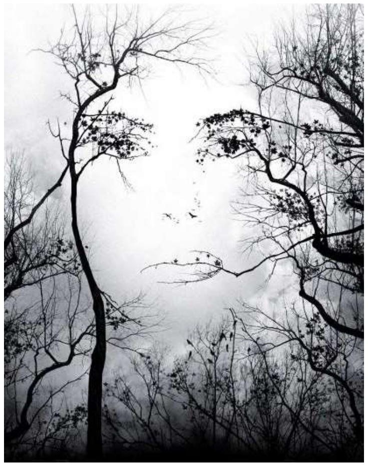


https://www.youtube.com/watch?v=a7efEqgpIrE

Course rationale

Human vision has its limitations

- intensities, shapes, patterns, motions can be misinterpreted
- it is labor intensive, time-consuming, subjective, error-prone



Computer vision can potentially improve this

- work day and night without getting tired
- analyse information quantitatively and objectively
- potentially more accurate, precise, reproducible

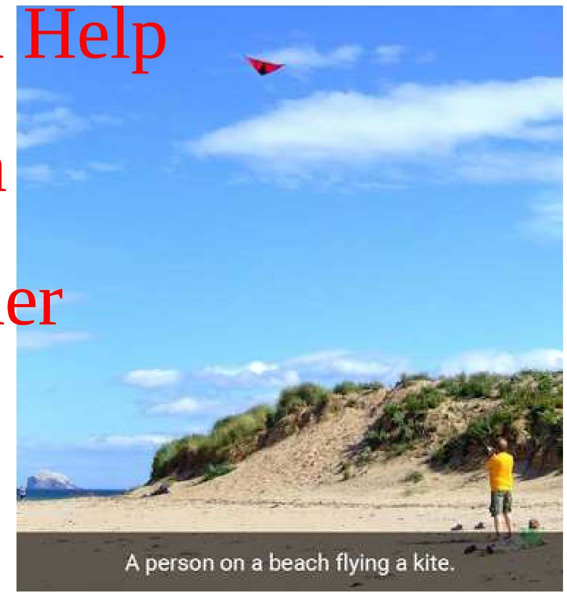
If the methods and tools are well designed!

Application: 3D shape reconstruction

Project VarCity recreates 3D city models using social media photos

Application: image classification and captioning

[Google's Show and Tell](#) open-source image captioning model in TensorFlow

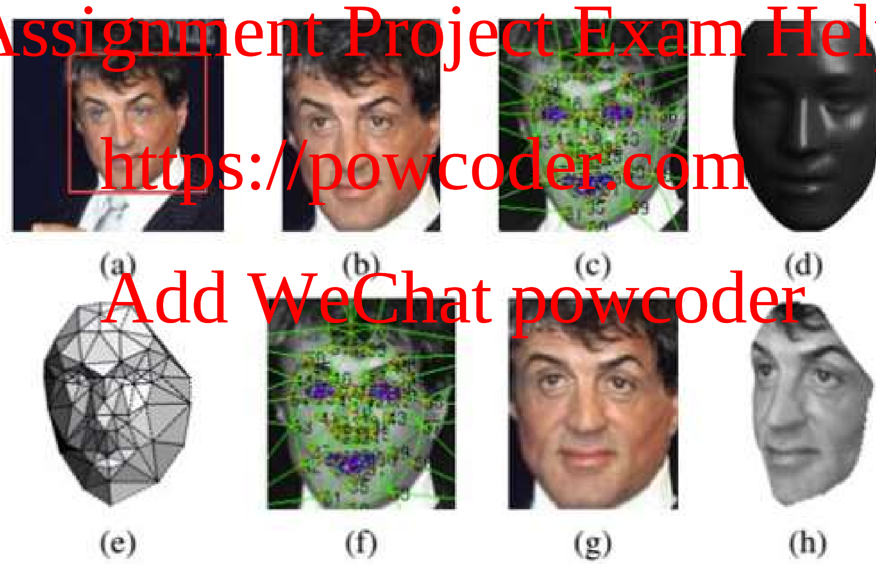


Application: intelligent collision avoidance

Iris Automation provides safer drone operation with intelligent collision avoidance

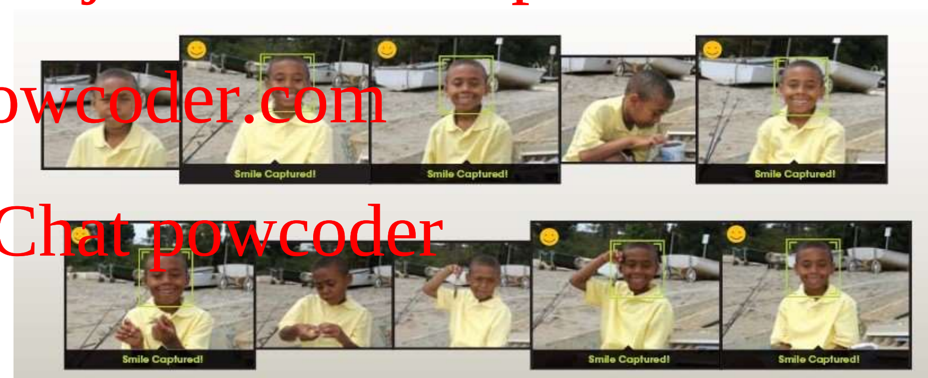
Application: face detection and recognition

Facebook's DeepFace project nears human accuracy in identifying faces

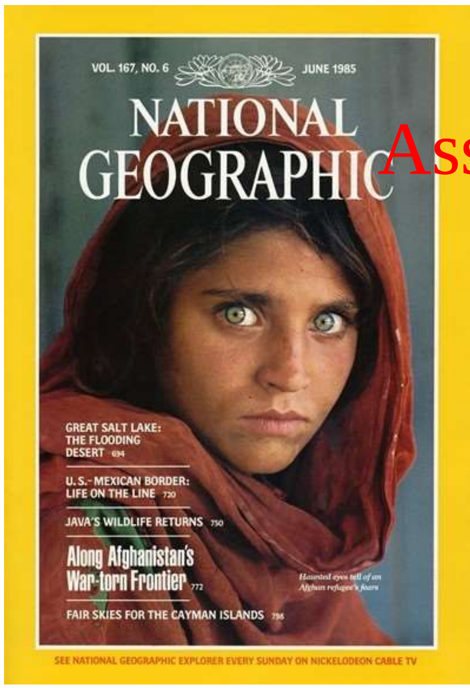


Application: face detection and recognition

For improving image capture on digital cameras

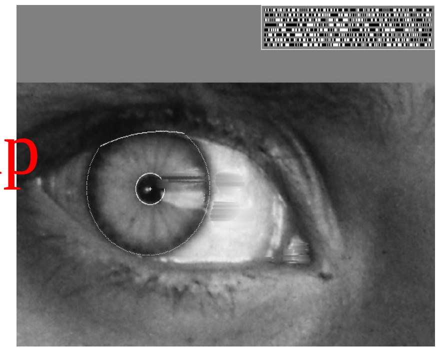


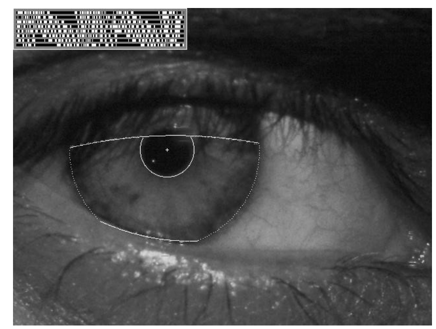
Application: vision-based biometrics



[How the Afghan girl was identified by her iris patterns]

The remarkable story of Sharbat Gula, first photographed in 1984 aged 12 in a refugee camp in Pakistan by National Geographic photographer Steve McCurry, and traced 18 years later to a remote part of Afghanistan where she was again photographed by McCurry…





Who is she?

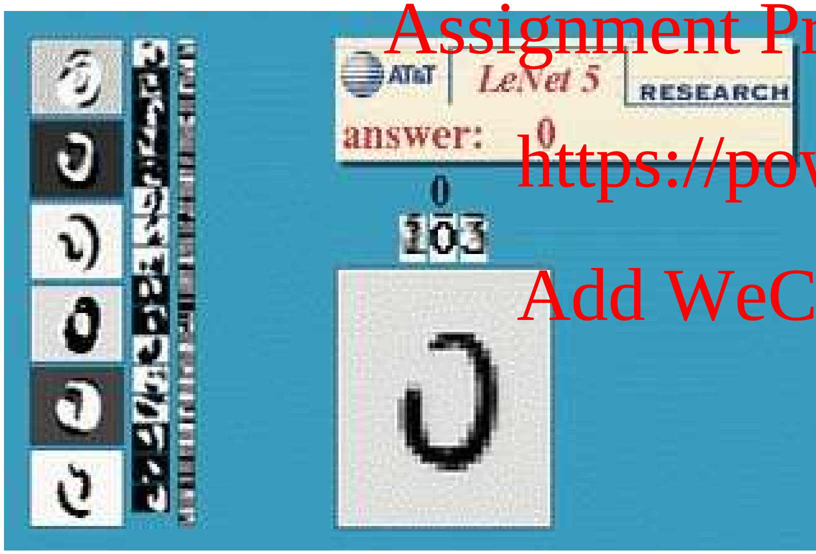
Application: logging in without a password

Fingerprint scanners on modern laptops and other devices

Windows Hello makes logging in as easy as looking at your PC

Application: optical character recognition (OCR)

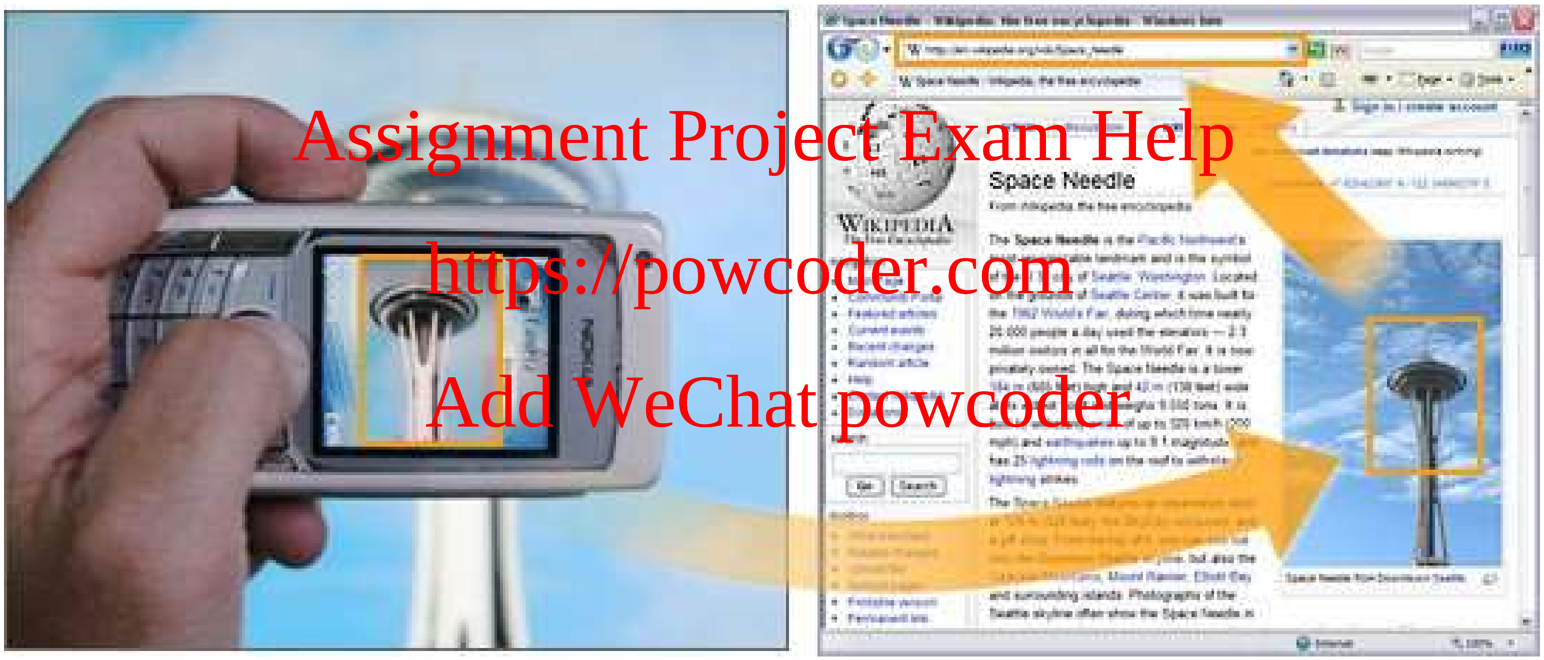
Converting scanned documents or number plates to processable text



Application: object recognition in supermarkets

[LaneHawk by Evolution Robotics Retail] provides a loss-prevention solution that detects bottom-of-basket (BOB) items in checkout lanes

Application: object recognition in phones



Application: autonomous vehicles

[Intel's Mobileye](#) makes cars safer and more autonomous

Application: space exploration

NASA's Mars Exploration Rover Spirit autonomously captured this picture in 2007

Vision systems used for panorama stitching, 3D terrain modeling, obstacle detection, position tracking

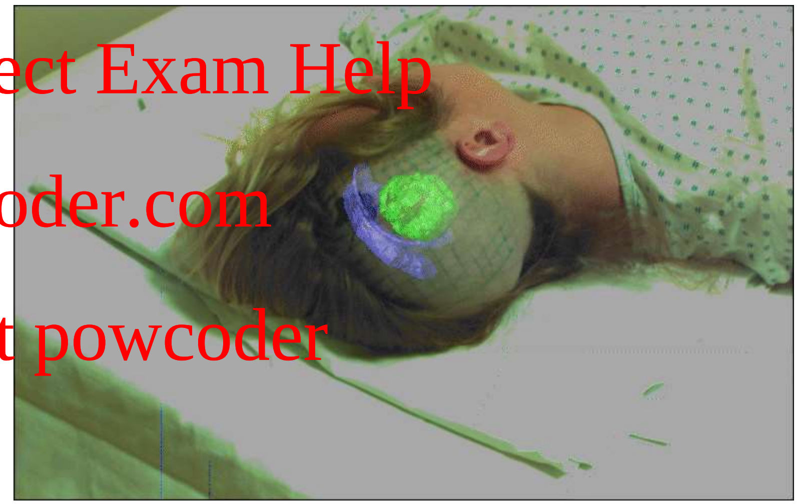
See [Computer Vision on Mars] for more information

Application: machine vision in robotics

NASA's Mars Spirit Rover

RoboCup

Application: medical imaging



Computer Aided Diagnosis

Image Guided Surgery

Application: video surveillance

- Traffic monitoring
- Person tracking
- Action recognition
- Speed estimation
- Object counting
- ...

Goals of Computer Vision

- Extract useful information from images: both metric and semantic
- Complexity of visual data is a challenge
- Recent progress due to higher processing power, memory, storage capacity
- Image→measurements→model→algorithms for learning and inference

Computer vision tasks

- Obtain simple inferences from individual pixel values
- Group pixels to separate object regions or infer shape information
- Recognise objects using geometric or statistical pixel information
- Combine information from multiple images into a coherent whole

Requires understanding of the physics of imaging and the use of mathematical and statistical models for information extraction

Critical issues in computer vision

- **Sensing**: how do sensors obtain images of the world?
- **Encoded Information**: how do images yield information of the scene, such as colour, texture, shape, motion…?
- **Representations**: what representations are appropriate to describe objects?
- **Algorithms**: what are appropriate algorithms to process image information and construct scene descriptions?

Low-level computer vision

This is almost entirely **digital image processing** (image in > image out)

- **Sensing**: image capture and digitisation
- **Preprocessing**: suppress noise and enhance object features
- **Segmentation**: separate objects from background and partition them
- **Description**: compute features that differentiate objects
- **Classification**: assign labels to image segments (regions)

High-level computer vision

This is about knowledge construction, representation, and inference

- **Recognition**: identify objects based on low-level information

- **Interpretation**: assign meaning to groups of recognized objects

- **Scene analysis**: complete understanding of the captured scene

Assumed knowledge

To do this course successfully you should:

- Be able to program well in Python or willing to learn it independently
- Be familiar with data structures and algorithms and basic statistics
- Be able/learn to use and integrate software packages (OpenCV, Scikit-Learn, Keras)
- Be familiar with vector calculus and linear algebra or willing to learn it independently

*Please self-assess **before** deciding to stay/enroll in the course*

Student learning outcomes

After completing this course you will be able to:

- Explain basic scientific, statistical, and engineering approaches to computer vision
- Implement and test computer vision algorithms using existing software platforms
- Build larger computer vision applications by integrating software modules
- Interpret and comment on articles in the computer vision literature

Course Changes in T1, 2021

Based on feedback received in 2020:

1. All course components, including the exam, are **entirely online**. Live online lectures will provide an opportunity to interact with the lecturer. Online labs will be interactive.
2. The **relative time allotted** to various topics will be adjusted to allow more time for new topics.
3. **Homework** has been introduced to help prepare for the new form of online exam. This is for practice only and will not be assessed; this would also reduce the proportion of marks allotted to writing style assessments.
4. The **labs** provide an opportunity **for coding and hands on work**, while lectures discuss the problems and solutions.

Weekly Class Structure

Week	Topic	Lecturer
1	Introduction, Image Formation, Image Processing	Professor Arcot Sowmya
2	Image Processing (continued)	Professor Arcot Sowmya
3	Feature Representation	Dr Yang Song
4	Pattern Recognition	Professor Arcot Sowmya
5	Image Segmentation	Professor Erik Meijering
6	Flexible Week (No Lectures, consultations will be held)	
7	Motion Tracking	Professor Erik Meijering
8	Applications	Professor Arcot Sowmya
9	Deep Learning, Applications	Dr Yang Song, Professor Arcot Sowmya
10	Project Demos	Professor Arcot Sowmya, Professor Erik Meijering, Dr Yang Song

Weekly Class Structure

Week	Lecture Time Lecturer	Topic
Week 1	Monday (12-2 pm) Prof Arcot Sowmya Thursday (2-4 pm) Prof Arcot Sowmya	Introduction, Image Formation Image Formation, Image Processing
Week 2	Monday (12-2 pm) Prof Arcot Sowmya Thursday (2-3 pm) Prof Arcot Sowmya	Image processing
Week 3	Monday (12-2 pm) Dr Yang Song Thursday (2-3 pm) Dr Yang Song	Feature representation
Week 4	Monday (12-2 pm) Prof Arcot Sowmya Thursday (2-3 pm) Prof Arcot Sowmya	 Pattern Recognition
Week 5	Monday (12-2 pm) Prof Erik Meijering Thursday (2-3 pm) Prof Erik Meijering	Image Segmentation
Week 6	FLEXIBLE WEEK	No lectures; consultations will be held
Week 7	Monday (12-2 pm) Prof Erik Meijering Thursday (2-3 pm) Prof Erik Meijering	Motion Tracking
Week 8	Monday PUBLIC HOLIDAY Thursday (2-3 pm) Prof Arcot Sowmya	 Applications
Week 9	Monday (12-2 pm) Dr Yang Song Thursday (2-3 pm) Prof Arcot Sowmya	Deep Learning of Computer Vision Applications/Exam preparation
Week 10	Monday (12-2 pm) Thursday (2-4 pm)	Project Demos

Weekly Class Structure

- Lectures: Mon and Thurs, see previous slide for details; note PUBLIC HOLIDAY on Monday 5th April
- Labs: 3-4 PM on Thursdays in weeks 2, 3, 4, 5
- Project consultations:
  - 3-4 PM Thursdays in weeks 6, 7, 8, 9
  - additional consultations require appointments with your assigned tutor
- Project demo:
  - On Monday and Thursday in week 10, during class hours
  - Detailed schedule will be announced on class web page

ALL changes will be announced on class web page on WebCMS3

Assessments

Late Submission Penalty

Unless you have received special dispensation from the Lecturer in Charge, work that is submitted after the deadline DURING THE TERM will incur a penalty of 10% per day, up to a maximum of 100%.

For the final examination, university exam rules will apply.

Assessment	Marks	Release	Due
Assignment	10%	Week 2	Week 4
Lab Work	10%	Weeks 2, 3, 4, 5	Weeks 3, 4, 5, 6
Project (multiple stages)			
- Individual component	15%	Week 5	Week 7
- Group component *	25%	Week 5	Week 10
Exam	40%	Exam Period	Exam Period

Communication- Modes and Etiquette

- **Online forum (Piazza) is your first port of call-** post query of wider interest on lectures, labs, assessments
- **Contact LIC** for late submission, absence, assessment deadlines, lab and assessment content
- **Contact Course admin** for issues with enrolment, file submission, group enrolment or any admin matter
- every effort will be made to respond quickly to queries- allow **maximum of 24 hours turnaround**
- Do observe **standards of equity and respect** in dealing with all students and staff- in person, emails, forum posts, all other communication
- Preferred language of **communication** is **English**

Special Consideration/Supp Policy

- If your work in this course is affected by unforeseen adverse circumstances, you should apply for **Special Consideration**

- UNSW handles special consideration requests centrally. **Do not just email the LIC** about special consideration.

- Special Consideration **requests must be accompanied by documentation**

- Mark calculated in the same way as other students who sat the original assessment
- If you are awarded a Supp and do not attend, then your exam mark will be zero.

More information on Course web page

Plagiarism Policy

READ the UNSW Policy and Procedure on this.

For the **purposes of COMP 9517**, plagiarism includes copying or obtaining all, or a substantial part, of the material for your assignment, whether written or graphical report material, or software code, **without written acknowledgement** in your assignment from:

- a location on the Internet
- a book, article or other written document (published or unpublished) whether electronic or on paper or other medium
- another student, whether in your class or another class
- someone else (e.g. from someone who writes assignments for money)

Plagiarism (ctd)

- If you copy material from another student or non-student **with acknowledgement**, you will not be penalised for plagiarism, but the marks you get for this will be at the marker's discretion, and will reflect the marker's perception of the amount of work you put into finding and/or adapting the code/text.
- If you use text found in a publication (on the Internet or otherwise) then the marks you get for this will be at the **marker's discretion**, and will reflect the marker's **perception of the amount of work** you put into finding and/or adapting the text.

The assessments provide opportunities for you to develop important skills. Use these opportunities!

Further information on WebCMS

Please be sure you are familiar with:

- Communication Etiquette
- Special Consideration
- Student Conduct
- Plagiarism Policy
- Academic Integrity

Further reading on lecture topics

In the lectures we will be referring to various online resources for further reading such as:

- Richard Szeliski, Computer Vision: Algorithms and Applications, Springer, 2021

- Dana H. Ballard and Christopher M. Brown, Computer Vision, Prentice Hall, 1982

- Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, MIT Press, 2016
- David A. Forsyth and Jean Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2011

- Simon J. D. Prince, Computer Vision: Models, Learning and Inference, Cambridge University Press, 2012

And other books, scientific articles, and other resources available online or via the UNSW Library

Further reading on today's topics

- Chapter 1 of Szeliski for a general introduction to computer vision
- Chapter 1, Shapiro and Stockman
- Appendix A and B of Szeliski for background on linear algebra, numerical techniques and statistics

Acknowledgements

- Some images on applications taken from Szeliski with original sources credited where possible
- Other images and videos credited where possible