

CSCM77: Computer Vision & Deep Learning

Prof. Xianghua Xie

x.xie@swansea.ac.uk

<http://csvision.swan.ac.uk>

Course Policies

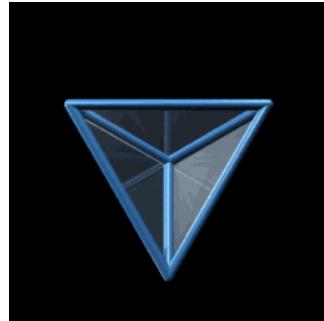
- Prerequisite: CSCM45/345 Big Data and Machine Learning
 - If you haven't taken this module, you can't enrol on this one.
- Lectures
 - 2 lectures per week,
 - See timetable on College Intranet for Time and Location
- Labs
 - 10 sessions, supervised
 - Lab works can only be signed off in allocated lab sessions (not outside these hours)
 - See timetable on College Intranet for Time and Location

Course Policies

- Assessment
 - Lab work: 20%
 - Coursework: 20%
 - viva
 - **Late submission: 0%**
 - Exam: 60%
- Recommended reading
 - D. Forsyth and J. Ponce, Computer Vision: a Modern Approach, 2003
 - C. Bishop, Pattern Recognition and Machine Learning, Springer 2006
- Tutor: all questions related to lab classes and programming
 - Mike Kenning (788486@swansea.ac.uk)

What is Computer Vision

- Computer Graphics
 - A layman definition: “Computer graphics are graphics **created** using computers”
 - Pictorial *synthesis* of real or imaginary objects from their computer-based models.



- Visualisation
 - “Visualization is any technique for creating images, diagrams, or animations to **communicate** a message”

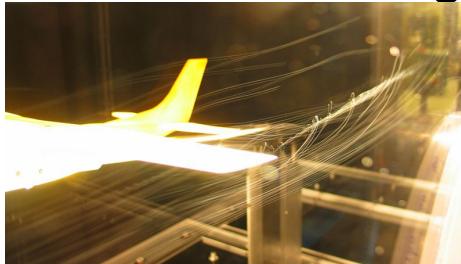
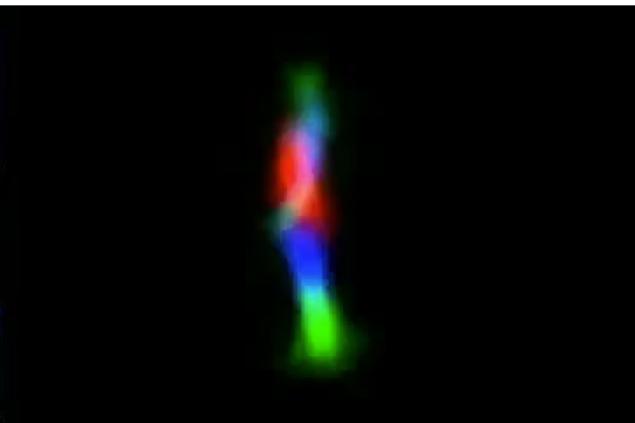
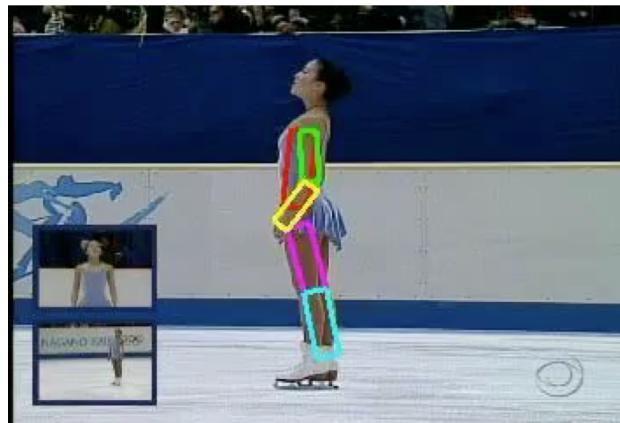


Image Source: Wikipedia & cs.swan.ac.uk
<http://www.ics.uci.edu/~dramanan>

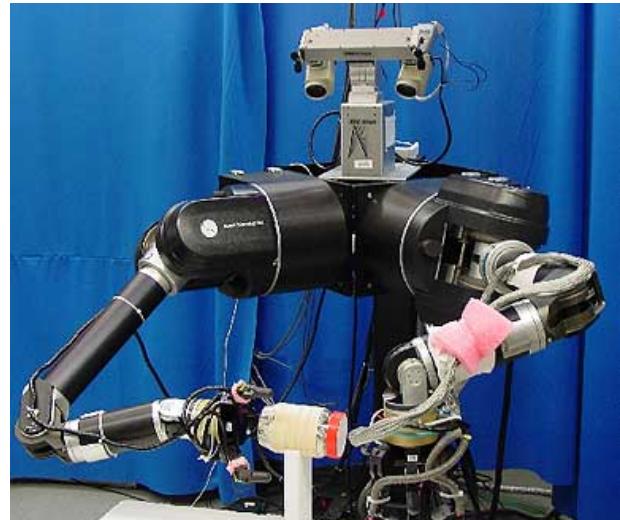
What is Computer Vision

- Computer Vision
 - A layman definition: “Computer vision is the science and technology of machines that **see**”
 - A broad term for the processing of image data, that commonly involves *feature extraction* and *machine learning*.
 - 1) underlying theories of optics, light and surfaces,
 - 2) underlying statistical, property and shape models,
 - 3) theory-based algorithms,
 - 4) issues related to what humans broadly relate to “understanding”.



What is Computer Vision

- “see” means more



“Concerned with the theory behind artificial systems that extract information from images”

<http://pubtheologian.files.wordpress.com/2011/06/read-bible.jpg>

http://www.palmerservices.net/images/Computer_on_desk.jpg

<http://www.cs.umass.edu/images/new.humanoid.jpg>

What is Computer Vision

- A brief history
 - Started in 60s, digital image processing;
 - UK largely began in 70s
 - Computer vision had a bad name in 80s (just like AI had its “fame” in 70s); however, lot of theoretical developments
 - 90s: great leap forward
 - Noughties: the most diverse, exciting; lots of practical applications
 - Last fifteen years: dramatic advances due to computing power and new machine learning methods
- Can you identify the object?
- No.
- You can't **recognise** it but you can **see** it?
- *That's right.*
- So we can **see** things that we can't **recognise**?

What is Computer Vision

- Imitating human vision and beyond
 - Recognise objects (people we know, things we own)
 - Localisation and tracking in space and time
 - Recognise actions and interactions
 - Understand intention (holy grail to many)
- Vision, however, is
 - Deceivingly easy
 - Deceptive
 - Demanding

What is Computer Vision

- Vision is deceptively easy
- Seeing seems to be effortless
 - Seeing seems simpler than “thinking”
 - Use ~70% of brain power for visual perception
 - Almost all “creatures” see, but don’t see alike



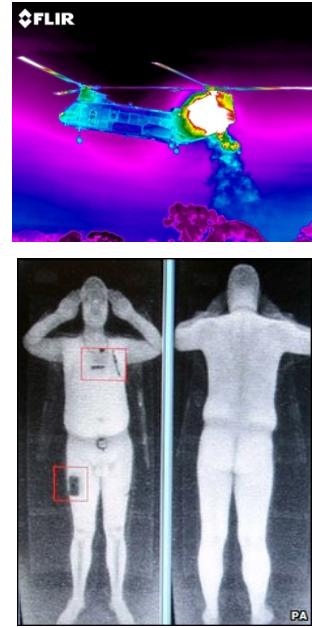
http://nikipedia87.files.wordpress.com/2008/11/cat_trig_002.jpg

<http://static.open.salon.com/files/cute-dog-yawning11233800905.jpg>

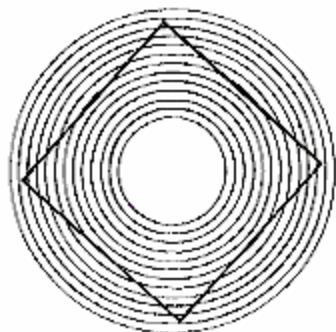
<http://craziestgadgets.com/wp-content/uploads/2009/03/dog-fanny-pack.jpg>

What is Computer Vision

- Vision is deceptive
- Vision is immediate,
- Exceptionally strong sensation,
- Performs reconstruction within our brain,
- Human vision is also limited (see <0.1% energy), passive (e.g. not like a bat)



Beyond human vision



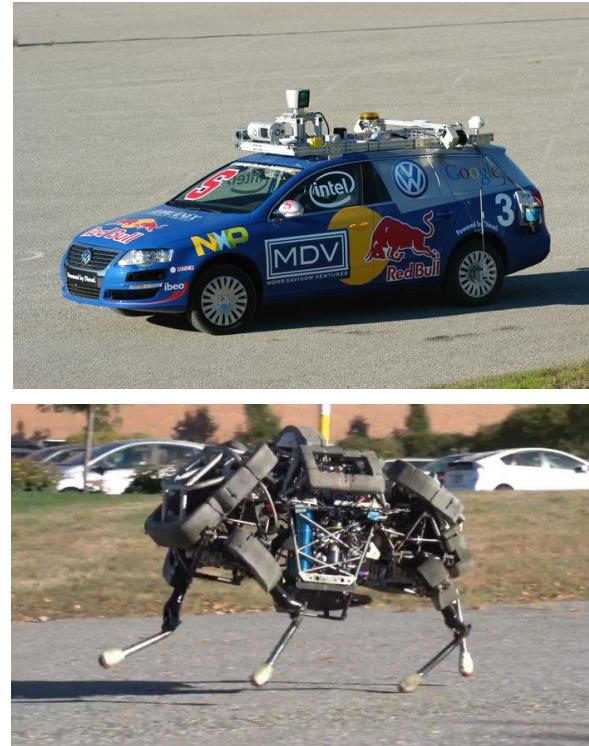
http://www.goinfrared.com/images/gallery/helicopter_5_ir.jpg

<http://news.bbc.co.uk>

CSCM77: CVDL

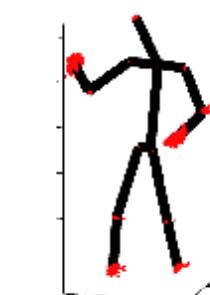
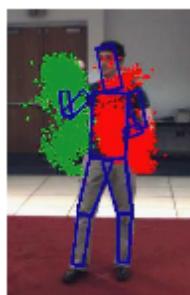
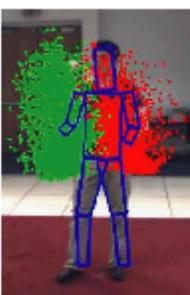
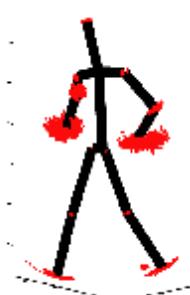
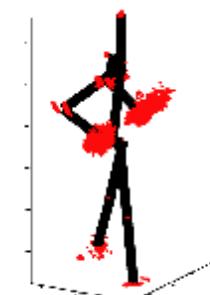
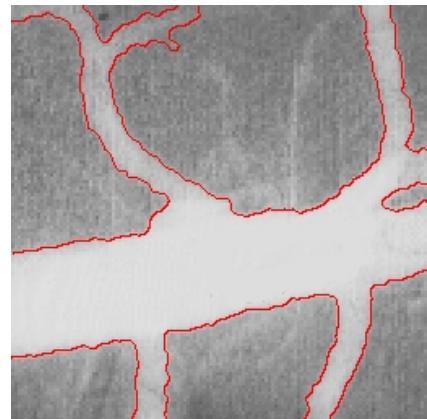
What is Computer Vision

- Generally very hard problems
- Often ill-posed problems
- Involves tools from various subject areas
 - Imaging
 - Signal processing
 - Mathematics
 - Linear systems, matrix decomposition, eigen analysis
 - Calculus (gradient, Taylor series), geometry (Fundamental matrix)
 - Probability, random variable ...
 - Machine learning & Artificial Intelligence
 - ...
 - Pattern recognition plays a fundamental role



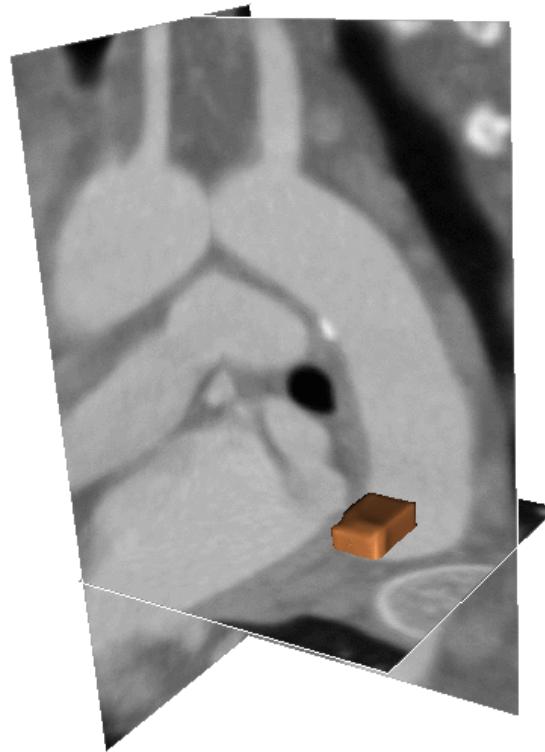
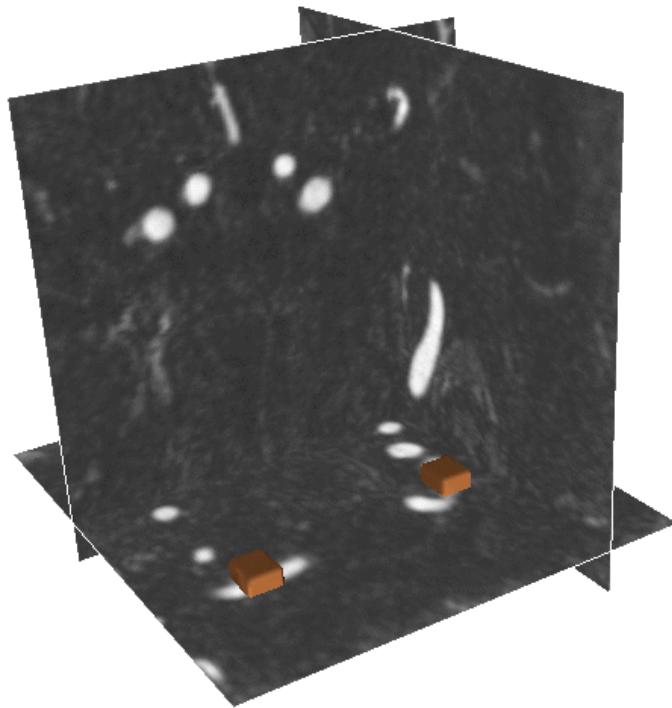
Computer Vision

- Processing single images



Computer Vision

- Processing multiple images
 - Volumetric data
 - images are spatially one after another



Computer Vision

- Processing multiple images
 - Video sequence: images are temporally one after another



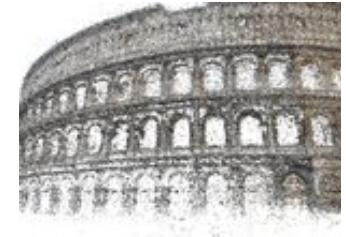
Computer Vision

- Processing multiple images
 - Thousands of individual images



Computer Vision

- Processing thousands of images



Computer Vision

- Processing thousands of images



Computer Vision

- Processing thousands of images



The old city of Dubrovnik, 4,619 images, 3,485,717 points

Building Rome in a Day, *Sameer Agarwal et al.*

Course Outline

- Part One: Kinect as an example vision system, Cameras, and Stereos
 - how Kinect works:
 - imaging and human pose detection,
 - concept of learning and random forests
 - camera model and calibration
- Part Two: Feature, Object, and Motion
 - Edges and Haar wavelet
 - HoG feature
 - Motion analysis
 - Object tracking

Course Outline

- Part Three: Deep Learning
 - Concept of deep learning
 - Deep Neural Network
 - Convolutional Neural Network
 - Region Proposal Network (Faster R-CNN)
 - Recurrent Neural Network

Deep learning is one of the most advanced and popular techniques in modern AI.



Is CVDL for me?

- Mathematics
 - Don't have to be very good at it, but willing to learn:
 - Vectors, matrices
 - Simple linear system
 - Basic probability
- Programming
 - Prior experience necessary
 - Python