Lecture 1: Introduction and Course Logistics
Course logistics

**Times:** Mon, Wed, Fri 12:00-1:00pm

**Instructor:** Jim Little

![Jim Little](image)

**E-mail:** little@cs.ubc.ca

**Course webpage:** [http://www.cs.ubc.ca/~little/425](http://www.cs.ubc.ca/~little/425)

**Discussion:** [piazza.com/ubc.ca/winterterm12019/cpsc425](piazza.com/ubc.ca/winterterm12019/cpsc425)
On-line Etiquette

Times: Mon, Wed, Fri 12-1pm
Locations: Online (Zoom)

Keep your **microphones muted**, unless you are asking a question.

**Raise your hand** (in zoom) if you want to ask a question, I will call on you (possibly not immediately), and then you can unmute and ask it, then mute again.

If you don’t have a microphone, you can ask a question in Chat, but chat is hard for me to monitor, so I would like this to be the option of “last resort”.

Course webpage: [https://www.cs.ubc.ca/~little/425](https://www.cs.ubc.ca/~little/425)
Discussion: [piazza.com/ubc.ca/winterterm22020/cpsc425201/home](piazza.com/ubc.ca/winterterm22020/cpsc425201/home)
**About me ...**

I have been working in *Computer Vision* for the last ~40 years.

<table>
<thead>
<tr>
<th>Position</th>
<th>Institution</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professor</td>
<td>University of British Columbia</td>
<td>1988 -</td>
</tr>
<tr>
<td>Research Scientist</td>
<td>Massachusetts Institute of Technology</td>
<td>1985-1988</td>
</tr>
<tr>
<td>PhD, MSc</td>
<td>University of British Columbia</td>
<td>1978 - 1985</td>
</tr>
<tr>
<td>Research Associate</td>
<td>Simon Fraser University</td>
<td>1975-1978</td>
</tr>
<tr>
<td>Research Analyst</td>
<td>Harvard University</td>
<td>1972 - 1975</td>
</tr>
</tbody>
</table>
Course logistics

**Times:** Mon, Wed, Fri 12:00-1:00pm

**Location:** Online (Zoom)

**Instructor:** Jim Little

**E-mail:** little@cs.ubc.ca

**Office:** ICICS 117

**TAs:**
- Rayat Hossain
  - rayat137@cs.ubc.ca
- Won Bae
- Kalli Leung
  - kalli.leung@alumni.ubc.ca
- Ariel Shann
  - shannari@cs.ubc.ca
- Gabriel Huang
  - whbae@cs.ubc.ca

**TA Contacts:**
- Won Bae: rayat137@cs.ubc.ca
- Kalli Leung: kalli.leung@alumni.ubc.ca
- Ariel Shann: shannari@cs.ubc.ca
- Gabriel Huang: whbae@cs.ubc.ca

**Course webpage:** http://www.cs.ubc.ca/~little/425.html

**Discussion:** piazza.com/ubc.ca/winterterm12019/cpsc425
Course **logistics**

**Times:** Mon, Wed, Fri 12-1pm

**Locations:** Online (Zoom)

Lectures will be on **Zoom**. Lectures will be recorded and made available on **Canvas**.

TA and Office hours: **Zoom**
Course **logistics**

**Times:** Mon, Wed, Fri 16:00-17:00pm

**Location:** DMP 310

Use **Piazza** for any questions related to material and assignments in the course

If you have a question, I can guarantee you that at least 10 students in the course have an identical question.

**Course webpage:** [http://www.cs.ubc.ca/~little/cpsc425](http://www.cs.ubc.ca/~little/cpsc425)

**Discussion:** [piazza.com/ubc.ca/winterterm12019/cpsc425](piazza.com/ubc.ca/winterterm12019/cpsc425)
Course logistics

**Times:** Mon, Wed, Fri 12-1pm

**Locations:** Online (Zoom)

I will use Canvas for assignment submission and grading.
I will use the **Course Webpage** for assignment and lecture slide distribution.

I will post slides before each lecture, so you can take notes over them if you wish.
Course logistic

**Times:** Mon, Wed, Fri 4-5pm

**Locations:** Online (Zoom)

Lectures (Live: Zoom; Recorded: Canvas; Slides: Canvas & Web Page)

Office and TA hours (Zoom)

Assignments (Instructions: Web Page & Canvas; Handin: Canvas)

Assigned Readings (Web Page)

Schedule (Web Page)

Questions & Assignment Support (Piazza)
Topics Covered

- Image Processing (Linear Filtering, Convolution)
- Filters as Templates
- Image Feature Detection (Edges & Corners)
- Texture & Colour
- Image Feature Description (SIFT)
- Model Fitting (RANSAC, The Hough Transform)
- Camera Models, Stereo Geometry
- Motion and Optical Flow
- Clustering and Image Segmentation
- Learning and Image Classification
- Deep Learning Introduction
CPSC 425 was originally developed by David Lowe, and then Bob Woodham, and has evolved over the years. Much of the material this year is adapted from material prepared by Bob, as well extensions developed by others who taught this course.

Previously taught by:

— 2020-2021 Term 1 by Leon Sigal
— 2019-2020 Term 2 by Leon Sigal
— 2019-2020 Term 1 by Jim Little
— 2018-2019 Term 2 by Leon Sigal
— 2018-2019 Term 1&2 by Leon Sigal
— 2017-2018 Term 2 by Leon Sigal
— 2016-2017 Term 2 by Jim Little
— 2015-2016 Term 2 by Fred Tung
— 2015-2015 Term 2 by Jim Little
Course Origins

The course is very **broad**, but relatively **shallow** introduction to a very diverse and complex field that draws material from geometry, statistics, AI, machine learning, computer graphics, psychology and many others.

— This means we will cover many topics and different algorithms

— I will give you as much background and connective tissue as I can

    ... but, there is no “linear” way to learn the material we will cover

    ... I will not be able to go into depth on some of the topics
How to do **Well** in the **Course**?

— It is easy to think that material is easy and course requires no studying

— Part of your job should be going over the slides and carefully **analyzing** not just what is on them, but the underlying assumptions, algorithmic steps and so on

— Don’t strive for “**template matching**” strive for true “**understanding**”
How to do **Well** in the **Course**?

— It is easy to think that material is easy and course requires no studying

— Part of your job should be going over the slides and carefully **analyzing** not just what is on them, but the underlying assumptions, algorithmic steps and so on

— Don’t strive for “**template matching**” strive for true “**understanding**”

— Some topics we will cover are **theoretic** and **fundamental** (e.g., geometry)

— Others are **algorithmic** (i.e., you make certain assumptions about the world, these assumptions may not always hold, but will be useful in building algorithms that ultimately perform well on a prescribed task)

— Computer vision is more of an **experimental** science - ultimately we are looking at performance to access if algorithmic choices are successful
Grading Criteria

Online Quizzes: 10%

Programming Assignments: 45%

6 graded and 1 ungraded (optional) assignment

Midterm Exam (TBD): 15%

Final Exam (TBD): 30%
Grading Criteria

You do NOT need to pass the final to pass the course

Online Quizzes: 10%

Programming Assignments: 45%

6 graded and 1 ungraded (optional) assignment

Midterm Exam (TBD): 15%

Final Exam (TBD): 30%
Quizzes

Will be made available on Canvas for a 24 hour window.

Number of quizzes has not been determined and each quiz may have a different number of questions / points.

Quizzes are designed to get you to think more deeply about what we are covering and to keep you on track with the material.
Assignments

There will be 7 assignments in total (6 marked)

— Approximately 1 every 2 weeks (two are 1.5 weeks)
— You will hand these in by 11:59pm on the due date (read hand in instructions and late policy on course webpage)

You will use the Python, with the following libraries: Python Imaging Library (PIL), NumPy, Matplotlib, SciPy, Scikit-Learn

— Assignment 0 (which is ungraded) will introduce you to this.

Assignments contribute 45% to your final score (equally distributed)
Midterm Exam

[Tentatively] on TBD

— In class during the lecture period
— Closed book, no notes allowed

Multiple choice, true / false and short answer questions

— Aimed to test your “understanding” of the content of the course

The Midterm exam will contribute 15% to your final score
The Final exam is held during the regular examination period and is scheduled by the Registrar’s Office.

Similar to the midterm but longer and with more extensive short/medium answer questions.

The Final exam will contribute 30% to your final score.
The course uses the following textbook, which is recommended (but not required):

Can be freely downloaded as a PDF from SpringerLink, through UBC Library Website (must login using CWL).

Computer Vision: Algorithms and Applications
By: R. Szeliski
Publisher: Springer
Pub. Date: 2010

Computer Vision: A Modern Approach (2nd edition)
By: D. Forsyth & J. Ponce
Publisher: Pearson
Pub. Date: 2012
Readings

You will be assigned *readings*.

— Sometimes you will be assigned readings from other sources

Do the reading **after coming** to the lecture

— Reading assignments will be posted on course webpage
— They will also be mentioned in class
How important is **Vision**?

To answer this question, we need to go back to about

.... **543 million years, B.C.**
How important is **Vision**?

To answer this question, we need to go back to about

.... 543 million years, B.C.

**Vision** is really fundamental to life and evolution
What is **Computer Vision**?

What is **Computer Vision**?

Computer vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.

What do you see?
What we would like **computer to infer?**

*Slide Credit: Jitendra Malik (UC Berkeley)*
What we would like **computer to infer**?

Will person B put some money into person C’s cup?
What is **Computer Vision**?

Computer vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.

Image (or video)  \[\rightarrow\]  **Sensing** Device  \[\rightarrow\]  **Interpreting** Device  \[\rightarrow\]  Interpretation

- blue sky,
- trees,
- fountains,
- UBC, ...

What is **Computer Vision**?

Compute vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.
“spend the summer linking a camera to a computer and getting the computer to describe what it saw”

- Marvin Minsky (1966), MIT
  Turing Award (1969)

... >50 years later
“You’ll notice that Sussman never worked in vision again!” – Berthold Horn
Can computers **match (or beat)** human vision?

• We’ve been at it for 50 years
Can computers **match (or beat)** human vision?

- How good is human vision?
Can computers **match (or beat)** human vision?
Can computers **match (or beat)** human vision?

Edward H. Adelson
Can computers **match (or beat)** human vision?

- **Yes and No** (mostly NO)
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*)
1. Vision for Measurement

Real-time stereo  

Structure from motion  

Tracking  

Wang et al.  

Snavely et al.  

Demirdjian et al.
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*)

**Ill-posed problem:** real world is much more complex than what we can measure in images: 3D -> 2D

It is (literally) impossible to invert the image formation process
Computer Vision Problems

1. Computing properties of the 3D world from visual data (measurement)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (perception and interpretation)
2. Vision for **Perception and Interpretation**
2. Vision for **Perception and Interpretation**

Objects
- The Wicked Twister
- Ferris wheel
- Lake Erie
- Deck
- Tree
- Tree
- Tree
- Bench
- Umbrellas
- Pedestrians
- People waiting in line
- People sitting on ride
- Writing
- Maxair

Activities
- Ride
- Ride
- Ride

Scenes
- Amusement park
- Cedar Point
- 12 E

Locations
- Sky
- Water
- Tree

Text / Writing
- Slide Credit: Kristen Grauman (UT Austin)

Faces

Gestures

Motions

Emotions

...
Computer Vision Problems

1. Computing properties of the 3D world from visual data (measurement)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (perception and interpretation)

It is computationally intensive / expensive
2. Vision for **Perception and Interpretation**

~ 55% of **cerebral cortex** in humans (13 billion neurons) are devoted to vision

more human brain devoted to vision than anything else
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*)

It is computationally intensive / expensive

We do not (fully) understand the processing mechanisms involved
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*)

3. Algorithms to mine, search, and interact with visual data (*search and organization*)
3. **Search and Organization**

![Diagram showing search and organization process](image)

- **Query**
- **Image or video archives**
- **Relevant content**

*Slide Credit: Kristen Grauman (UT Austin)*
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*).

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*).

3. Algorithms to mine, search, and interact with visual data (*search and organization*).

Scale is enormous, explosion of visual content.
3. Search and Organization

Snapchat
31.7 Million / hour

WhatsApp
29.2 Million / hour

Instagram
2.9 Million / hour

Flickr
18K hours / hour

Facebook
14.6 Million / hour

*based on article by Kimberlee Morrison in Social Times (2015)
3. **Search and Organization**

> 85% of all web content is multimedia content of visual form

*from iStock by GettyImages

*based on article by Kimberlee Morrison in Social Times (2015)
Computer **Vision Problems**

1. Computing properties of the 3D world from visual data (*measurement*)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*)

3. Algorithms to mine, search, and interact with visual data (*search and organization*)

4. Algorithms for manipulation or creation of image or video content (*visual imagination*)

*Slide Credit: Kristen Grauman (UT Austin)*
4. Visual **Imagination**

- Identity = ◆ | ▲ | ■
- Action = ○ walking | ● running | ◆ skipping | ▲ jumping jack | ▼ side step

He et al. ECCV 2018

Zhao et al. ECCV 2018
4. Visual **Imagination**

<table>
<thead>
<tr>
<th>Input</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>![Input Image]</td>
<td>![Black Hair]</td>
<td>![Blond Hair]</td>
</tr>
</tbody>
</table>

---

---

---

---
Computer Vision Problems

1. Computing properties of the 3D world from visual data (*measurement*)

2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*)

3. Algorithms to mine, search, and interact with visual data (*search and organization*)

4. Algorithms for manipulation or creation of image or video content (*visual imagination*)

*Slide Credit:* Kristen Grauman (UT Austin)
Can computers **match (or beat)** human vision?

- **Yes and No** (mostly NO)

- Let’s see some examples of state-of-the-art and where it is used
Optical Character Recognition (OCR)

Technology to convert **scanned documents to text**
(comes with any scanner now days)

Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition
Face Detection

Technology available in any digital camera now  
(one of the first big commercial successes of vision algorithms)
Smile Detection
Face Recognition

Facebook

Apple’s iPhoto

http://www.apple.com/ilife/iphoto/

Slide Credit: Devi Parikh (GA Tech) and Fei-Fei Li (Stanford)
Vision for **Biometrics**
Vision for **Biometrics**

“How the Afghan Girl was Identified by Her Iris Patterns”  Read the [story wikipedia](#)
Vision for Biometrics

Fingerprint scanners on many new laptops, other devices

Image Credit: James Hays (GA Tech)

iPhone X Face ID

Face recognition systems are not part of widely used technologies

How it works and how to fool it:

https://www.youtube.com/watch?v=FlbMLmsCax0
Object Recognition (in supermarkets)

https://www.youtube.com/watch?v=NrmMk1Myrxc
Object Recognition (in mobile devices)

https://www.youtube.com/watch?v=8SdwVCUJ0QE

Nikia’s Point & Find

https://en.wikipedia.org/wiki/Nokia_Point_%26_Find
3D Urban Modeling and Virtual Tourism

[ Agarwal, Furukawa, Snavely, Curless, Seitz, Szeliski, 2010 ]
Visual Special Effects (VFX): Shape and Motion Capture

Slide Credit: Stephen Seitz (University of Washington)
Vision in **Sports**

Sportvision first down line

Nice [explanation](http://www.howstuffworks.com) on [www.howstuffworks.com](http://www.howstuffworks.com)


**Slide Credit:** Stephen Seitz (University of Washington)
Automotive Safety and Smart Cars

Tesla’s Autopilot

Google Self-driving Cars

Mobileye
Interactive Games: **Kinect**

Sensor Components:
- Depth Sensor
- Power Light
- IR Emitters
- RGB Camera
- Microphone Array

Image: [Link](http://www.rat.com)
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.
Vision for **Medical Imaging**

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Slide Credit: James Hays (GA Tech)
Captioning and Visual Question Answering

Demo: [Vinyals et al., 2015]

Demo: [Seo et al., NIPS 2017]

Q1: What color is the bowl?
GT answer: White
Predicted answer: White
Rank of GT: 1

Q2: Do you see any people?
GT answer: No
Predicted answer: No, just the cat
Rank of GT: 2

Q3: What color is the cat?
GT answer: Grey, white, and black
Predicted answer: Grey, black and white
Rank of GT: 6
Related Disciplines

Artificial Intelligence (AI)

Computer Vision

Scope of CPSC 425

Image Processing

Geometric Reasoning

Recognition

Slide Credit: James Hays (GA Tech)
Prepare for the **Next Lecture**

**Readings:**

— Next Lecture: Forsyth & Ponce (2nd ed.) 1.1.1 — 1.1.3

**Reminders:**

— Start working on **Assignment 0** (ungraded) due **January 18**

— [optional] Watch TED talk by Prof. Fei-Fei Li
  https://www.youtube.com/watch?v=40riCqvRoMs
Related Disciplines: Vision and Graphics

Inverse problems: analysis and synthesis

(it is sometimes useful to think about computer vision as inverse graphics)
Why Study Computer Vision?

It is one of the **most exciting areas of research** in computer science.

Among the **fastest growing technologies** in the industry today.
Wired’s 100 **Most Influential People** in the World

63. Yann LeCun

*Director of AI research, Facebook, Menlo Park*

LeCun is a leading expert in deep learning and heads up what, for Facebook, could be a hugely significant source of revenue: understanding its user’s intentions.

62. Richard Branson

*Founder, Virgin Group, London*

Branson saw his personal fortune grow £550 million when Alaska Air bought Virgin America for $2.6 billion in April. He is pressing on with civilian space travel with Virgin Galactic.

61. Taylor Swift

*Entertainer, Los Angeles*
CVPR Attendance


Attendance Counts: 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 1000, 6000