Lecture 1: Introduction and Course Logistics

Image Credit: Devi Parikh
Course logistic

**Times:** Tues, Thurs 12:30-2:00pm  
**Locations:** West Mall Swing Space, Room 121

**Instructor:** Leonid Sigal

**E-mail:** lsigal@cs.ubc.ca  
**Office:** ICICS 119

**Course webpage:** [https://www.cs.ubc.ca/~lsigal/teaching19_Term2.html](https://www.cs.ubc.ca/~lsigal/teaching19_Term2.html)  
**Discussion:** [piazza.com/ubc.ca/winterterm22020/cpsc425201/home](piazza.com/ubc.ca/winterterm22020/cpsc425201/home)
About me ...
About me …

Software Engineer
1999 - 2001
About me …

Software Engineer
1999 - 2001
About me ...

PhD, MSc
2001 - 2008

Software Engineer
1999 - 2001
About me ...

Postdoctoral Researcher
2007 - 2009

PhD, MSc
2001 - 2008

Software Engineer
1999 - 2001
About me ...

Senior Research Scientist
2009 - 2017

Postdoctoral Researcher
2007 - 2009

PhD, MSc
2001 - 2008

Software Engineer
1999 - 2001
About me ...

Associate Professor
2017 -

Senior Research Scientist
2009 - 2017

Postdoctoral Researcher
2007 - 2009

PhD, MSc
2001 - 2008

Software Engineer
1999 - 2001
## About me …

I have been working in **Computer Vision** for the last ~20 years.

<table>
<thead>
<tr>
<th>Position</th>
<th>Institution</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associate Professor</td>
<td>UBC</td>
<td>2017 -</td>
</tr>
<tr>
<td>Senior Research Scientist</td>
<td>Disney Research</td>
<td>2009 - 2017</td>
</tr>
<tr>
<td>Postdoctoral Researcher</td>
<td>University of Toronto</td>
<td>2007 - 2009</td>
</tr>
<tr>
<td>PhD, MSc</td>
<td>Brown, Boston</td>
<td>2001 - 2008</td>
</tr>
<tr>
<td>Software Engineer</td>
<td>Cognex</td>
<td>1999 - 2001</td>
</tr>
</tbody>
</table>
Course logistic

**Times:** Tues, Thurs 12:30-2:00pm

**Locations:** West Mall Swing Space, Room 121

**Instructor:** Leonid Sigal

E-mail: lsigal@cs.ubc.ca

Office: ICICS 119

**TAs:**
- Mona Fadaviardakani
  - mfadavi@cs.ubc.ca
- Hafsa Zahid
  - hafsa.zahid@alumni.ubc.ca
- Tzu-Yun (Ariel) Shann
  - shannari@cs.ubc.ca
- Tianxin Tao
  - taotianx@cs.ubc.ca
- Suhail Mohammed
  - suhail33@cs.ubc.ca

**E-mail:**
- shannari@cs.ubc.ca
- suhail33@cs.ubc.ca

**Locations:**
- West Mall Swing Space, Room 121

**Times:**
- Tues, Thurs 12:30-2:00pm

**Course webpage:** [https://www.cs.ubc.ca/~lsigal/teaching19_Term2.html](https://www.cs.ubc.ca/~lsigal/teaching19_Term2.html)

**Discussion:** [piazza.com/ubc.ca/winterterm2020/cpsc425201/home](piazza.com/ubc.ca/winterterm2020/cpsc425201/home)
Course logistic

**Times:** Tues, Thurs 12:30-2:00pm  
**Locations:** West Mall Swing Space, Room 121

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 logistic

**Times:** Tues, Thurs 12:30-2:00pm  
**Locations:** West Mall Swing Space, Room 121

I will (only) use **Canvas** for assignment submission and grading.
I will use **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.
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.

*Image Credit: https://www.deviantart.com/infinitecreations/art/BioMech-Eye-168367549*
What do you see?

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

Will person B put some money into person C’s cup?

- Person A walking away carrying 3 bags
- Person B looking at C
- Person C playing Accord D
- Bag F
- Bench E with 3D model

**Slide Credit**: Jitendra Malik (UC Berkeley)
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.

![Diagram](Image (or video))

**Sensing Device** => **Interpreting Device** => **Interpretation**

- Image (or video)
- 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.


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.

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.

This implies a very strong connection between **Computer Vision** and **(Machine) Learning**.
Computer vision ... the beginning ...

The Summer Vision Project

"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
Computer vision … the beginning …

“You’ll notice that Sussman never worked in vision again!” – Berthold Horn

Slide Credit: Devi Parikh (GA Tech)
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?
Can computers **match (or beat)** human vision?

- How good is human vision?
  
  As a measuring device not very good, as a functioning device really good
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**

Wang et al.

**Structure from motion**

Snavely et al.

**Tracking**

Demirdjian et al.

*Slide Credit: Kristen Grauman (UT Austin)*
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**

*Slide Credit: Kristen Grauman (UT Austin)*
2. Vision for **Perception and Interpretation**

- **Objects**: Cedar Point, Lake Erie, The Wicked Twister, Ferris wheel, amusement park, 12 E, Deck, bench, umbrella, maxair, tree, people, pedestrians, people waiting in line, people sitting on ride.
- **Activities**: Riding rides, sitting, standing.
- **Scenes**: Amusement park scene, lake view, tree-filled area.
- **Locations**: Cedar Point, Lake Erie, 12 E.
- **Text / writing**: The Wicked Twister 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)

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

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

*from iStock by GettyImages
3. Search and Organization

Snapchat: 31.7 Million / hour
WhatsApp: 29.2 Million / hour
Facebook: 14.6 Million / hour
Instagram: 2.9 Million / hour
Flickr: 18K hours / 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

*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*)
4. Visual *Imagination*

- **Identity** = ♦ | ▲ | ■
- **Action** = walking | running | skipping | jumping jack | side step

He et al. ECCV 2018

**Demo:** https://layout2im.cs.ubc.ca/layout/

Zhao et al. ECCV 2018
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)*
Challenges: Viewpoint invariance

Michelangelo 1475-1564

*slide credit Fei-Fei, Fergus & Torralba
Challenges: Lighting

*image credit J. Koenderink*
Challenges: Scale

*slide credit Fei-Fei, Fergus & Torralba
Challenges: Deformation
Challenges: Occlusions

Rene Magritte 1965
Challenges: Background clutter

Kilmeny Niland 1995
Challenges: Local ambiguity and context

*image credit Fergus & Torralba*
Challenges: Local ambiguity and context
Challenges: Motion

*Image credit Peter Meer*
Challenges: Object inter-class variation
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
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](https://en.wikipedia.org/wiki/Afghan_Girl) wikipedia

*Slide Credit:* James Hays (GA Tech)
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=FhbMLmsCax0
Object Recognition (in supermarkets)

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

Google Goggles
Use pictures to search the web. Watch a video.

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

Nikia’s Point & Find

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

[ Agarwal, Furukawa, Snavely, Curless, Seitz, Szeliski, 2010 ]
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

Slide Credit: Amnon Shashua
Interactive Games: Kinect

Sensor Components

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

Image: [Link to Kinect image](http://www.tsl.com)
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: http://vqa.cloudcv.org
Demo: http://demo.visualdialog.org

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

[ Vinyals et al., 2015 ]
[ Seo et al., NIPS 2017 ]
Computer Vision

Scope of CPSC 425

Image Processing
Geometric Reasoning
Recognition

Deep Learning

Machine Learning

Robotics

Human Computer Interaction

Medical Imaging

Optics

Neuroscience

Graphics

Computational Photography

Slide Credit: James Hays (GA Tech)
Related Disciplines: Vision and Graphics
Related Disciplines: Vision and Graphics
Related Disciplines: Vision and Graphics

Model

Graphics

Slide Credit: Kristen Grauman (UT Austin)
Related Disciplines: Vision and Graphics

Images

Model

Graphics

Slide Credit: Kristen Grauman (UT Austin)
Related Disciplines: Vision and Graphics

Slide Credit: Kristen Grauman (UT Austin)
Related Disciplines: Vision and Graphics

**Inverse problems:** analysis and synthesis
Related Disciplines: Vision and Graphics

**Inverse problems:** analysis and synthesis

(it is sometimes useful to think about computer **vision as inverse graphics**)

Slide Credit: Kristen Grauman (UT Austin)
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.
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 £580 million when Alaska Air bought Virgin America for $1.6 billion in April. He is pressing on with civilian space travel with Virgin Galactic.

61. Taylor Swift

Entertainer, Los Angeles
CVPR 2019
Long Beach, CA
June 16th - June 20th
CVPR Attendance

![Attendance Chart]

- 2006: 1000
- 2007: 1000
- 2008: 1000
- 2009: 1000
- 2010: 2000
- 2011: 2000
- 2012: 2000
- 2013: 2000
- 2014: 2000
- 2015: 3000
- 2016: 4000
- 2017: 5000
- 2018: 6000
- 2019: 9000
Course logistic

**Instructor:** Leonid Sigal

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

**Office:** ICICS 119

**Times:** Tues, Thurs 12:30-2:00pm

**Locations:** West Mall Swing Space, Room 121

**TAs:**
- Mona Fadaviardakani
  - mfadavi@cs.ubc.ca
- Hafsa Zahid
  - hafsa.zahid@alumni.ubc.ca

**Tzu-Yun (Ariel) Shann**

**Discussion:**
- piazza.com/ubc.ca/winterterm22018/cpsc425

**TAs:**
- shannari@cs.ubc.ca
- Tianxin Tao
  - taotianx@cs.ubc.ca

**Suhail Mohammed**

**E-mail:** suhail33@cs.ubc.ca

**Office:** ICICS 119

**Times:** Tues, Thurs 12:30-2:00pm

**Locations:** West Mall Swing Space, Room 121

**Course webpage:** https://www.cs.ubc.ca/~lsigal/teaching19_Term2.html

**Discussion:** piazza.com/ubc.ca/winterterm2020/cpsc425201/home
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
Course Origins

CPSC 425 was originally developed by 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:

- 2019-2020 Term 2 by Leonid Sigal
- 2019-2020 Term 1 by Jim Little
- 2018-2019 Term 1 & 2 by Leonid Sigal
- 2016-2017 Term 2 by Jim Little
- 2015-2016 Term 2 by Fred Tung
- 2015-2015 Term 2 by Jim Little

Note: This is my 3rd time teaching CPSC 425
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 connection 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**”
Grading Criteria

In-class **clicker questions**: 10%

**Programming Assignments**: 25%

6 graded and 1 ungraded (optional) assignment

**Midterm** Exam (February 14th): 25%

**Final** Exam (TBD): 40%
Clicker Questions

Bring your i>Clickers to class

Register your remote: https://canvas.ubc.ca/ before the next class (we will test them during next lecture)

There will be clicker questions (not in every lecture):
— 1/2 point for participation
— 1/2 point for correct answer
*not all clicker quizzes are worth the same # of points, depends on # of questions.

The clicker questions contribute 10% to your total grade

Missing Quiz Policy: If you miss a quiz for a legitimate and documented reason, that quiz will be dropped (legitimate reasons: illness, conference travel, etc.) You are required to contact instructor and provide proof within 1 week of missed quiz.
Assignments

There will be 7 assignments in total (6 marked)

– Approximately 1 every 2 weeks (last one 1 week)
– 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 25% to your final score (each graded assignment is 5% of your grade)
Scheduled for **February 25th** (Tuesday after the break)

- Here 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 **25%** to your final score
The Final exam is held during the regular examination period, **April 14 — April 29, 2020**, 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 **40%** to your final score.
The course uses the following textbook, which is recommended (but not required):

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

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

Can be freely downloaded as a PDF from SpringerLink, through UBC Library Website (must login using CWL).
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
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 Tuesday, **January 14**

- [optional] Watch TED talk by Prof. Fei-Fei Li
  
  [https://www.youtube.com/watch?v=40riCqvRoMs](https://www.youtube.com/watch?v=40riCqvRoMs)