CMPT 412
Computer Vision

Yasutaka Furukawa
http://www.cs.sfu.ca/~furukawa/cmpt412
Today

• My introduction

• What is computer vision?

• Course fast-forward and logistics
My affiliations in the past
3D Computer Vision
Yasutaka Furukawa

Associate Professor
GrUM Lab, VML Lab
Simon Fraser University

Email: furukawa-atmark-sfu.ca
Office: TASC I 8023
Phone: 778-782-4619
Fax: 778-782-3045
Mail address: click here

Bio (CV, Google scholar)
I am an associate professor of Computing Science at Simon Fraser University. Prior to SFU, I was an assistant professor at Washington University in St. Louis. Before WUSTL, I was a software engineer at Google. Before Google, I was a post-doctoral research associate at University of Washington. I worked with Prof. Seitz and Prof. Curless at University of Washington, and Rick Szeliski at Facebook (was at Microsoft Research). I completed my Ph.D. under the supervision of Prof. Ponce at Computer Science Department of University of Illinois at Urbana-Champaign in May 2008.
What is computer vision?
What is Computer Vision?
For 5 years old...
What is Computer Vision?

“Make computers do what our eyes do”

but this is not really correct...
What is Computer Vision?

visual system (eyes and brain)

“Make computers do what our eyes do”

but this is not really correct...
What is the shape of this thing?
What is the shape of this thing?
What is Computer Vision?

“Make computers do what our eyes do”

but this is not really correct...
What is Computer Vision?

For certain tasks, Computer Vision can do much more...

“Make computers do what our eyes do”

but this is not really correct...
Can you create this w/o Computer Vision?

[ Apple Maps ]
What is Computer Vision?

Achieve (super) human-level perception
Is computer vision hard?
What a person sees
What a computer sees
Why are we able to interpret this image?
[ A. Kitaoka, 2003 ]
Important note:

In general, computer vision does not work (except in certain situations/conditions)
Important note:

In general, computer vision does not work
(except in certain situations/conditions)

5 years ago…
Important note:
In many important tasks, computer vision works

Now
Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Justus Thies¹, Michael Zollhöfer², Marc Stamminger¹, Christian Theobalt², Matthias Nießner³

¹University of Erlangen-Nuremberg
²Max-Planck-Institute for Informatics
³Stanford University

CVPR 2016 (Oral)
Real-Time Facial Motion Capture [Faceshift]
Deep Face
Google driving to be driverless

Google’s modified Toyota Prius uses an array of sensors to navigate public roads without a human driver. Other components, not shown, include a GPS receiver and an inertial motion sensor.

Laser-guided mapping
A rotating sensor with lasers called a LIDAR on the roof scans more than 200 feet in all directions to generate a precise three-dimensional map of the car’s surroundings.

Position estimator
A sensor mounted on the left rear wheel measures small movements made by the car and helps to accurately locate its position on the map.

Video camera
A camera mounted near the rear-view mirror detects traffic lights and helps the car’s onboard computers recognize moving obstacles—such as pedestrians and bicyclists.

Radar
Four standard automotive radar sensors, three in front and one in the rear, help determine the positions of distant objects.

Source: Google

NEW YORK TIMES; PHOTOGRAPHS BY RAMIY RAIMIAN FOR THE NEW YORK TIMES
Google Self-Driving Cars
Microsoft Kinect 2.0
Computer Vision for VR
Image stitching
Tango
Virtual Fitting
Effects magnification
It’s a good time to do computer vision
Industry aggressively hiring CV graduates, or even students!

(strong dominant industrial presence at conferences for recruitment)
Course logistics
Topics to be covered

Image processing:

- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.
Topics to be covered

Feature detection and correspondences:

• Corner detection.

• SIFT et al.

• Feature descriptors.

• RANSAC.
Topics to be covered

Transformations and geometry:

- Homographies and image alignment.
- Camera models.
- Fundamental matrix.
- Epipolar geometry and stereo.
- Structure from motion.
Topics to be covered

Physics-based vision:

• (Reflectance and image formation.)

• (Radiometry.)

• Shape from shading.

• Photometric stereo.

• (Color.)
Topics to be covered

Image understanding:

- Object recognition
- Image segmentation
- Face detection
- (Optical flow)
Topics to be covered

Deep neural network:

• Neural networks
• Convolutional neural networks
• Usages in Computer Vision
Slide credits

Most of these slides were adapted directly from:

- Ioannia Gkioulekas (16-385, Spring 2018)
- Kris Kitani (15-463, Fall 2016).

Inspiration and some examples also came from:

- Fredo Durand (Digital and Computational Photography, MIT).
- Kayvon Fatahalian (15-769, Fall 2016).
Grading

• 6 coding assignments (96% = 16% x 6)

• 4 % is a bonus

Assignments:
• a lot of programming in Matlab.
• hours and hours of programming.
• days and days of debugging.

Please:
• Be around for lectures.
• Post on Piazza discussions.
• Ask and answer questions.
Proposal: Assignments

Computer Vision - CMPT 412, Fall 2019

Project 1: Image Filtering and Hough Transform

Project 2: Augmented Reality with Planar Homographies

Project 3: 3D Reconstruction

Project 4: Scene Recognition with Bag of Words

Project 5: Digit Recognition with Convolutional Neural Networks

Project 6: Deep Learning with TensorFlow

Final Project (choice instead of projects 5 & 6)
**Computer Vision - CMPT 412, Fall 2019**

**Course Schedule**

<table>
<thead>
<tr>
<th>Dates</th>
<th>Tuesdays (14:30pm - 16:20pm)</th>
<th>Thursdays (14:30pm - 15:20pm)</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/3 &amp; 9/5</td>
<td>Introduction</td>
<td>Image filtering</td>
<td></td>
</tr>
<tr>
<td>9/10 &amp; 9/12</td>
<td>Image pyramids and Fourier transform</td>
<td>Hough transform</td>
<td></td>
</tr>
<tr>
<td>9/17 &amp; 9/19</td>
<td>Feature and corner detection</td>
<td>Feature detection and matching</td>
<td>Assign. 3 dot</td>
</tr>
<tr>
<td>9/24 &amp; 9/26</td>
<td>3D transformations</td>
<td>Image homographies</td>
<td></td>
</tr>
<tr>
<td>10/1 &amp; 10/2</td>
<td>3D models</td>
<td>3D reconstruction</td>
<td>Assign. 2 dot</td>
</tr>
<tr>
<td>10/8 &amp; 10/10</td>
<td>Stereo</td>
<td>Stereo</td>
<td></td>
</tr>
<tr>
<td>10/15 &amp; 10/17</td>
<td>Structure from motion</td>
<td>Structure from motion</td>
<td>Assign. 3 dot</td>
</tr>
<tr>
<td>10/22 &amp; 10/24</td>
<td>3D reconstruction to recognition</td>
<td>Shape matching</td>
<td></td>
</tr>
<tr>
<td>10/29 &amp; 10/31</td>
<td>Tracking</td>
<td>Segmentation</td>
<td>Assign. 4 dot</td>
</tr>
<tr>
<td>11/5 &amp; 11/7</td>
<td>Neural networks</td>
<td>Convolutional neural networks 1</td>
<td></td>
</tr>
<tr>
<td>11/12 &amp; 11/14</td>
<td>Training CNNs</td>
<td>Detection CNNs</td>
<td>Assign. 5 dot</td>
</tr>
<tr>
<td>11/19 &amp; 11/22</td>
<td>Segmentation CNNs</td>
<td>State-of-the-art CNN techniques</td>
<td></td>
</tr>
<tr>
<td>11/26 &amp; 11/28</td>
<td>Photometric stereo and shape from shading</td>
<td>Face detection</td>
<td>Assign. 6 dot</td>
</tr>
</tbody>
</table>

- Short break in the middle for Tuesdays
- Lecture contents for some weeks may change depending on the progress.
Late Policy and Mercy Rule

• For a late submission, penalty of 20% score reduction per 12 hours is applied. If one submits an assignment within 12 hours after the deadline, the score becomes 80%. Within 24 hours, 60%, and so on.

• If one is struggling but needs to obtain a passing grade, one has to complete at least “half” of the task for at least 4 assignments and must choose regular coding assignments instead of the final project.
Is this late or not?

- Deadline is by the end of 9/21 for example...

- Is this late?
  - 9/20 23:59:35
  - 9/21 24:04:02
  - 9/21 24:01:03
  - 9/21 24:00:00
Is this late or not?

- Deadline is by the end of 9/21 for example…

- Is this late?
  - 9/20 23:59:35  Safe
  - 9/21 24:04:02  Late
  - 9/21 24:01:03  Late
  - 9/21 24:00:00  Late
Recommended Book

Computer Vision
Algorithms and Applications
Richard Szeliski
Springer

PDF online
http://szeliski.org/Book/
Prerequisites

• Formally, MATH 152, and 9 units in Computing upper division courses or permission of the instructor. Highly recommend CMPT 307.

• Warning: Computer Vision is sexy but requires very solid math and extensive coding.
Office hours TBD
Final note

• We expect that you see all the posts on Piazza.

• Ask questions on Piazza to be shared by everybody.

• Do not send personal emails to me or TA, which can be posted to Piazza.

• Other things (e.g., academic integrity), please look at the course homepage: http://www.cs.sfu.ca/~furukawa/cmpt412
Questions?