Introduction
Overview of today’s lecture

• Teaching staff introductions
• What is computer vision?
• Course fast-forward and logistics
Teaching staff introductions
Instructor: Ioannis (Yannis) Gkioulekas

I won’t hold it against you if you mispronounce my last name

Originally from Greece

National Technical University of Athens (2004-2009)

Harvard University (2009-2017)

Carnegie Mellon University (2017-now)

My website: http://www.cs.cmu.edu/~igkioule

See also: http://imaging.cs.cmu.edu/
Looking around corners

visible wall

scan point

source & sensor

occluder

NLOS object

what a regular camera sees

what we can reconstruct

http://imaging.cs.cmu.edu/
Looking inside deep scattering objects

camera  thick smoke cloud  simulated camera measurements  reconstructed cloud volume  slice through the cloud
Seeing light in flight

camera for capturing video at $10^{15}$ frames per second

http://imaging.cs.cmu.edu/
Rendering wave effects

What real laser images look like

Speckle: noise-like pattern

What real laser videos look like

Match wave equation solvers, $10^5 \times$ faster

Reproduce physical effects like memory effect

http://imaging.cs.cmu.edu/
Differentiable rendering

visible surface

occluder

source and sensor

NLOS scene

reconstruction evolution

http://imaging.cs.cmu.edu/
Master of Science in Computer Vision (MSCV)

Research Interests:
Simultaneous Localization and Mapping (SLAM),
Deep Learning, Object Detection/Tracking

Current Area of Research:
Long term mapping for SLAM for dynamic environments

Past Research:
Video Stabilization, Object Tracking

TA: Anand Bhoraskar
TA: Prakhar Kulshreshtha (PK)

- **Master of Science in Computer Vision (MSCV)**

- **Research Interests:**
  - Deep Learning for Detection and Instance Segmentation
  - SLAM
  - Online Learning in Computer Vision

- **Current area of research:**
  - Long-term SLAM for Dynamic environments (under Prof. Michael Kaess)

- **Past Research:**
  - Instance segmentation for quality estimation of food grains on a smartphone
  - Text Intelligence in smartphone keyboard apps
  - Online Face Clustering
What is computer vision?
What a person sees
What a computer sees
Why are we able to interpret this image?
The goal of computer vision is to give computers (super) human-level perception.
typical perception pipeline

representation

‘fancy math’

output
typical perception pipeline

representation

'fancy math'

output

what should we look at?
(image features)

what can we understand?
(semantic segmentation)
typical perception pipeline

representation

‘fancy math’

output

easy to get lost in the techniques

what should we look at?
(image features)

what can we understand?
(semantic segmentation)
The parts that we are most interested in

representation

'fancy math'

output

what should we look at?
(image features)

what can we understand?
(semantic segmentation)
Important note:

In general, computer vision does not work
Important note:

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

Automated visual inspection
Object Recognition

Toshiba Tech IS-910T 2013

DataLogic LaneHawk LH4000 2012
Face detection

Sony Cyber-shot

Age recognition

Smile recognition
Face makeovers
First-down line
BMW night vision

BMW 5 series
Infinity EX

“Around view” camera
The system converts image data taken by 4 super-wide angle cameras, to display a virtual image of the vehicle from above.
Vision in Cars

2015
Image stitching
Photosynth
Tango
Virtual Fitting
Computer Vision for VR
Deep Face
Deep Dream
Facebook video style transfer 2016
Face2Face: Real-time Face Capture and Reenactment of RGB Videos

Justus Thies\textsuperscript{1}, Michael Zollhöfer\textsuperscript{2}, Marc Stamminger\textsuperscript{1}, Christian Theobalt\textsuperscript{2}, Matthias Nießner\textsuperscript{3}

\textsuperscript{1}University of Erlangen-Nuremberg
\textsuperscript{2}Max-Planck-Institute for Informatics
\textsuperscript{3}Stanford University

CVPR 2016 (Oral)
It’s a good time to do computer vision
Industry aggressively hiring CV faculty from universities
Industry aggressively hiring CV graduates, or even students!

(strong dominant industrial presence at conferences for recruitment)
CVPR GROWTH

Number of papers at CVPR

Original slide courtesy of CVPR 2016
CVPR GROWTH
Number of attendees at CVPR

Original slide courtesy of CVPR 2016
Computer vision at CMU

Dedicated courses for each subject we cover in this class:

- Physics-based Methods in Vision
- Geometry-based Methods in Computer Vision
- Computational Photography
- Visual Learning and Recognition
- Statistical Techniques in Robotics
- Sensors and sensing

... plus an entire department’s worth of ML courses.

CVPR 2019: CMU was the second most common academic affiliation among authors

(can you guess the first?)
Master in Computer Vision at CMU
Course logistics
Website

http://www.cs.cmu.edu/~16385/
(includes links to Canvas and Piazza)

Assignments

https://canvas.cmu.edu/courses/14118

Discussion&notes

https://piazza.com/class/k53x5h48my264d
(you should sign up here on your own)
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

Objects, faces, and learning:

- Basics of probability.
- K-means, KNN, PCA, SVM.
- Bag of words.
- Viola-Jones face detection.
- Perceptron, backpropagation.
- Convolutional neural networks.
Topics to be covered

Dealing with motion:

• Optical flow (LK, HS).

• Image registration.

• Kalman Filtering.

• Tracking (KLT, Mean-Shift).
Topics to be covered

Special topics:

• Computational photography.

• ???
Grading

- Seven two-week programming assignments: 70%
- Twelve weekly take-home quizzes: 27%
- Class and Piazza participation: 3%

Take-home quizzes:
- New this year.
- Two-three theory questions.
- Replace mid-term and final.

Participation:
- Be around for lectures.
- Post on Piazza discussions.
- Ask and answer questions.
Programming assignments

Assignment 1 Hough Transform
Assignment 2 Homography
Assignment 3 Stereo
Assignment 4 Photometric Stereo
Assignment 5 Bag of Words
Assignment 6 Convolutional Neural Nets
Assignment 7 Lucas-Kanade Tracking

• a lot of programming in Matlab and Python.
• hours and hours of programming.
• days and days of debugging.
• generous grading policy (like grad school)
• take advantage of extra credit
Programming assignments

Assignment 1 Hough Transform
Assignment 2 Homography
Assignment 3 Stereo
Assignment 4 Photometric Stereo
Assignment 5 Bag of Words
Assignment 6 Convolutional Neural Nets
Assignment 7 Lucas-Kanade Tracking

- a lot of programming in Matlab and Python.
- hours and hours of programming.
- days and days of debugging.
- generous grading policy (like grad school)
- take advantage of extra credit

seriously, a lot of programming, so start early!
Schedule

- Tentative schedule on course website.
- Likely to change.
- Always check course website and Piazza for updates!

<table>
<thead>
<tr>
<th>Date</th>
<th>Topics</th>
<th>Slides</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>M, Jan 13</td>
<td>Introduction</td>
<td></td>
<td>PA1 out</td>
</tr>
<tr>
<td>W, Jan 15</td>
<td>Image filtering</td>
<td></td>
<td>TQ1 out</td>
</tr>
<tr>
<td>M, Jan 20</td>
<td><strong>No class</strong> (Martin Luther King day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Jan 22</td>
<td>Image pyramids and Fourier transform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Jan 27</td>
<td>Hough transform</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Jan 28</td>
<td>Feature and corner detection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Feb 3</td>
<td>Feature descriptors and matching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Feb 5</td>
<td>2D transformations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Feb 10</td>
<td>Image homographies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Feb 12</td>
<td>Camera models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Feb 17</td>
<td>Two-view geometry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Feb 19</td>
<td>Stereo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Feb 24</td>
<td>Structure from motion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Feb 26</td>
<td>Radiometry and reflectance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Mar 2</td>
<td>Photometric stereo and shape from shading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Mar 4</td>
<td>Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Mar 9</td>
<td><strong>No class</strong> (spring break)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Mar 11</td>
<td><strong>No class</strong> (spring break)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Mar 16</td>
<td>Image processing pipeline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Mar 18</td>
<td>Introduction to recognition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Mar 23</td>
<td>Bag of works</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Mar 25</td>
<td>Neural networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Mar 30</td>
<td>Convolutional neural networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Apr 1</td>
<td>Optimization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Apr 6</td>
<td>Faces</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Apr 8</td>
<td>Optical flow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Apr 13</td>
<td>Alignment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Apr 15</td>
<td>Tracking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Apr 20</td>
<td>Temporal models and SLAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Apr 22</td>
<td>Graph-based methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M, Apr 27</td>
<td>Segmentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W, Apr 29</td>
<td>Wrap-up and discussion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PA1 out, PA2 out, PA3 out, PA4 out, PA5 out, PA6 out, PA7 out, TQ1 due, TQ2 due, TQ2 due, TQ3 out, TQ1 due, TQ2 due, TQ3 due, TQ4 due, TQ5 due, TQ5 due, TQ6 out, TQ6 due, TQ7 due, TQ8 due, TQ7 due, TQ8 due, TQ8 due, TQ9 out, TQ9 due, TQ10 out, TQ10 due, TQ11 due, TQ11 due, TQ12 out.
Leniency

Late days for programming assignments:

• 10% reduction of points per late day
• 6 free late days total
• use them wisely… save for later (harder) assignments!

Option to skip take-home quizzes:

• you only need to submit 9 out of 12 quizzes
• late quizzes will not be graded
Book

We will be posting readings after each lecture

PDF online

http://szeliski.org/Book/
Prerequisites

We assume familiarity with calculus, linear algebra, basic probability, and programming.

Formal prerequisites:

- "Mathematical Foundations of Electrical Engineering" (18-202) and "Principles of Imperative Computation" (15-122)

OR

- "Matrix Algebra with Applications" (21-240) and "Matrices and Linear Transformations" (21-241) and "Calculus in Three Dimensions" (21-259) and "Principles of Imperative Computation" (15-122)

If you are missing a prerequisite but still want to enroll, let me know and we’ll discuss it.
Contact information and office hours

• Feel free to email us about administrative questions.
  o please use [16385] in email title!

• Technical questions should be asked on Piazza.
  o we won’t answer technical questions through email.
  o you can post anonymously if you prefer.

• Office hours will be determined by poll.
  o feel free to email Yannis about additional office hours.
  o you can also just drop by Yannis’ office (Smith Hall (EDSH) Rm 225).

Yannis will announce office hours for this week.
Please take the course survey before the next lecture!

(also posted on Piazza)