Université Concordia University
FACULTY OF ENGINEERING AND COMPUTER SCIENCE
Department of Computer Science and Software Engineering

COMP498G/691G COMPUTER VISION

LECTURE 1: INTRODUCTION TO COMPUTER VISION

Today's Lecture

- Introductions
- What is Computer Vision?
- Computer Vision @ Concordia University
- Course administrative details
- Questions

Teaching Assistant/Grader

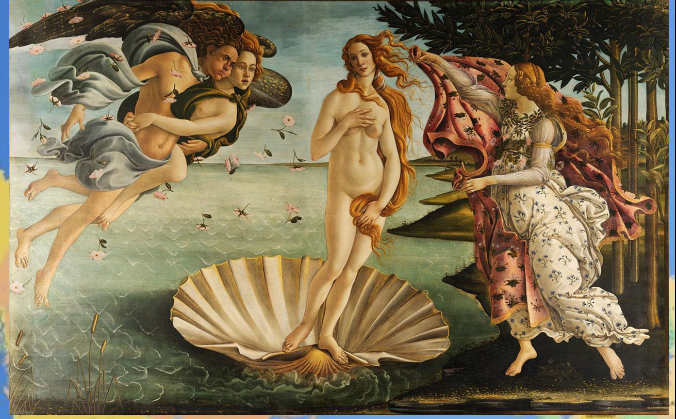
GTA: Xichen Zhou
Office: EV9.113
Office hours: During lab sessions or by appointment

Instructor

Dr. Charalambos [Charis] Poullis
Associate Professor
Office: EV3.183


GREENLAND (DENMARK)
ALASKA (USA)
CANADA
UNITED STATES OF AMERICA
MEXICO
RUSSIA
KAZAKHSTAN
MONGOLIA
CHINA
INDIA
SPAIN
ITALY
TURKEY
UKRAINE
BELARUS
POLAND
SWEDEN
FINLAND
NORWAY
FRANCE
GERMANY
IRAQ
IRAN
SAUDI ARABIA
ALGERIA
LIBYA
EGYPT
MALI
NIGER
CHAD
SUDAN
ETHIOPIA
KENYA
TANZANIA
ANGOLA
ZAMBIA
NAMIBIA
SOUTH AFRICA
BRAZIL
PERU
BOLIVIA
ARGENTINA
COLOMBIA
AUSTRALIA


CYPRUS
GREENLAND (DENMARK)
ALASKA (USA)
CANADA
UNITED STATES OF AMERICA
MEXICO
BRAZIL
PERU
BOLIVIA
ARGENTINA
COLOMBIA
SWEDEN
FINLAND
NORWAY
POLAND
BELARUS
UKRAINE
GERMANY
FRANCE
SPAIN
ITALY
GREECE
TURKEY
SYRIA
IRAQ
IRAN
KAZAKHSTAN
MONGOLIA
CHINA
INDIA
JAPAN
THAILAND
ALGERIA
LIBYA
EGYPT
SAUDI ARABIA
YEMEN
MALI
NIGER
CHAD
SUDAN
ETHIOPIA
NIGERIA
GUINEA
CENTRAL AFRICAN REPUBLIC
DEMOCRATIC REPUBLIC OF THE CONGO
UGANDA
KENYA
TANZANIA
ANGOLA
ZAMBIA
NAMIBIA
BOTSWANA
SOUTH AFRICA
AUSTRALIA



ALASKA (USA)
UNITED STATES OF AMERICA
MEXICO
SPAIN
ITALY
ALGERIA
LIBYA
MALI
NIGER
CHAD
SUDAN
NIGERIA
ETHIOPIA
KENYA
UGANDA
TANZANIA
ANGOLA
ZAMBIA
NAMIBIA
SOUTH AFRICA
COLOMBIA
PERU
BRAZIL
BOLIVIA
ARGENTINA
AUSTRALIA


UNIVERSITY OF MANCHESTER


Powell Library
UCLA
May 25, 2010
VerticalGeo
Connecting Our World
www.verticalgeo.com

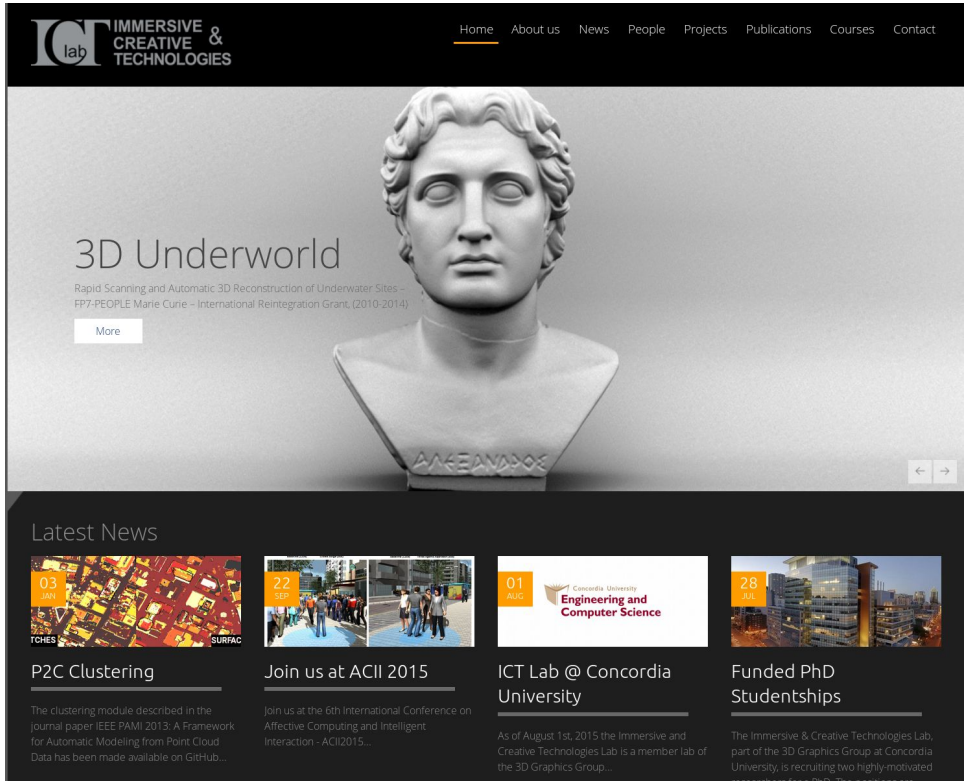

CYPRUS
ALASKA (USA)
GREENLAND (DENMARK)
CANADA
UNITED STATES OF AMERICA
ICELAND
SWEDEN
FINLAND
NORWAY
POLAND
BELARUS
UKRAINE
KAZAKHSTAN
MONGOLIA
SPAIN
ITALY
GREECE
TURKEY
SYRIA
IRAQ
IRAN
CHINA
ALGERIA
LIBYA
EGYPT
SAUDI ARABIA
INDIA
MALI
NIGER
CHAD
SUDAN
ETHIOPIA
KENYA
TANZANIA
ANGOLA
ZAMBIA
NAMIBIA
SOUTH AFRICA
AUSTRALIA


CYPRUS


ALASKA (USA)
CANADA
UNITED STATES OF AMERICA
MEXICO
COLOMBIA
PERU
BOLIVIA
ARGENTINA
Concordia
Flickr: Viola Ng

Research Interests

- Computer Vision
  - Computer Graphics
- Virtual Reality
- Augmented Reality



IMMERSIVE & CREATIVE TECHNOLOGIES

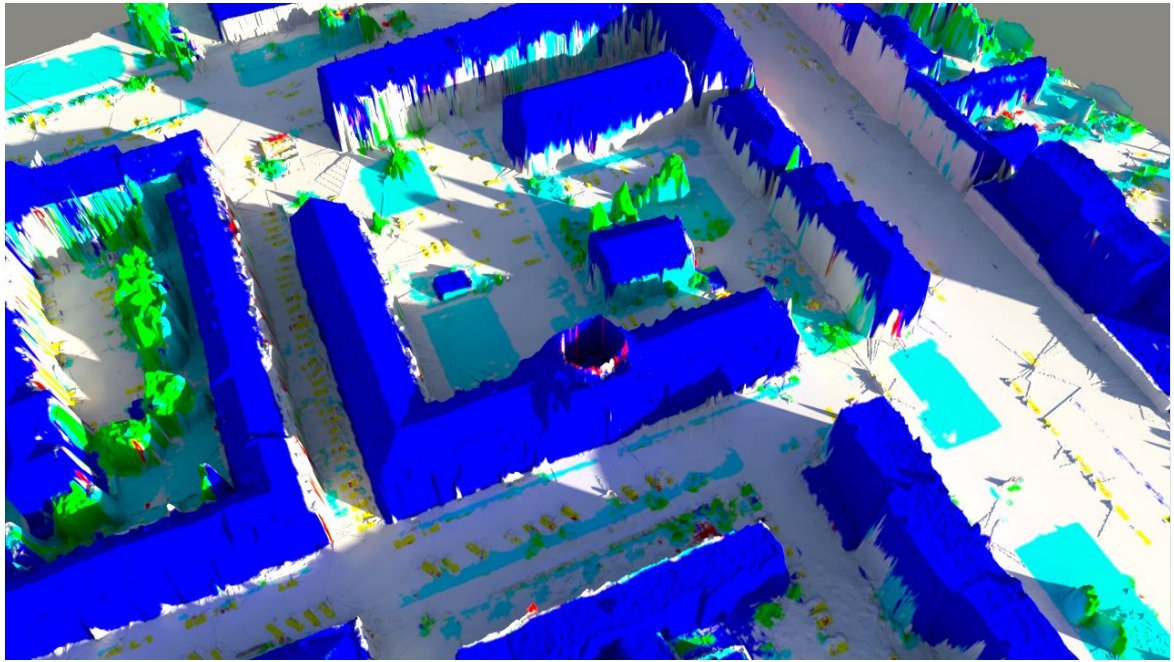
www.theICTlab.org

Recent work

Multi-label Pixelwise Classification for Reconstruction of Large-scale Urban Areas, Y. He, S. Mudur, C. Poullis

Recent work

Multi-label Pixelwise Classification for Reconstruction of Large-scale Urban Areas, Y. He, S. Mudur, C. Poullis



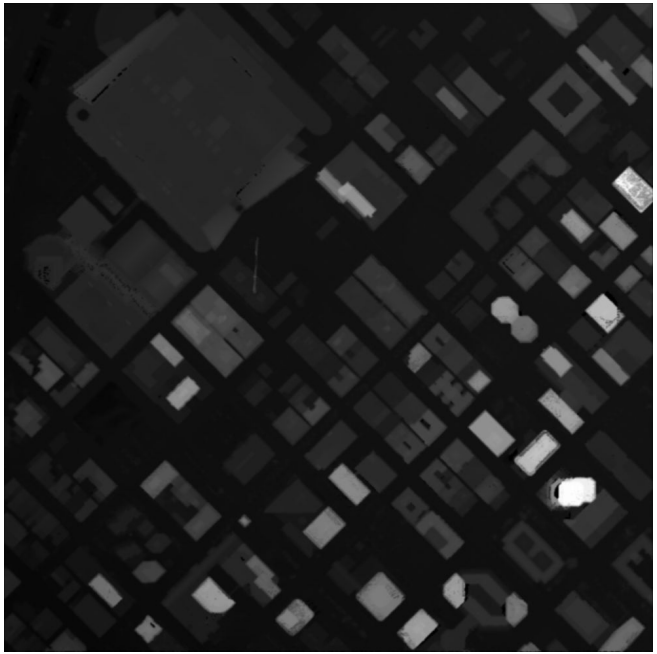
Recent work

Multi-label Pixelwise Classification for Reconstruction of Large-scale Urban Areas, Y. He, S. Mudur, C. Poullis

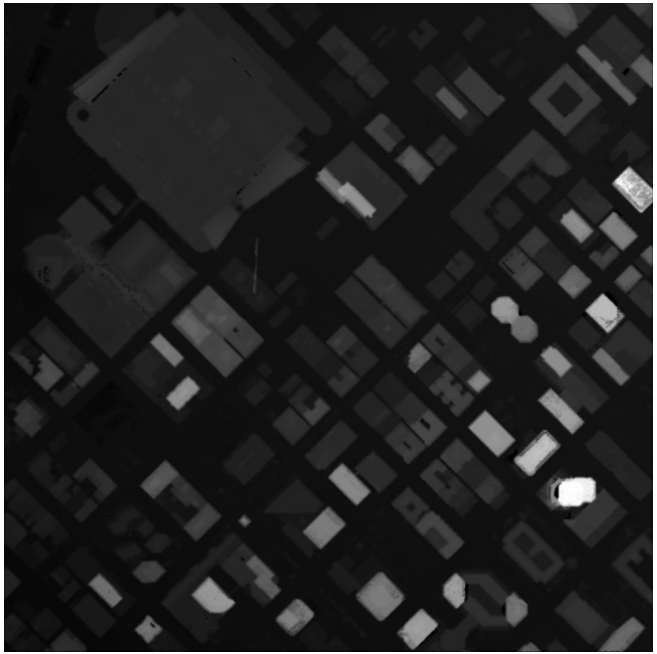
Recent work

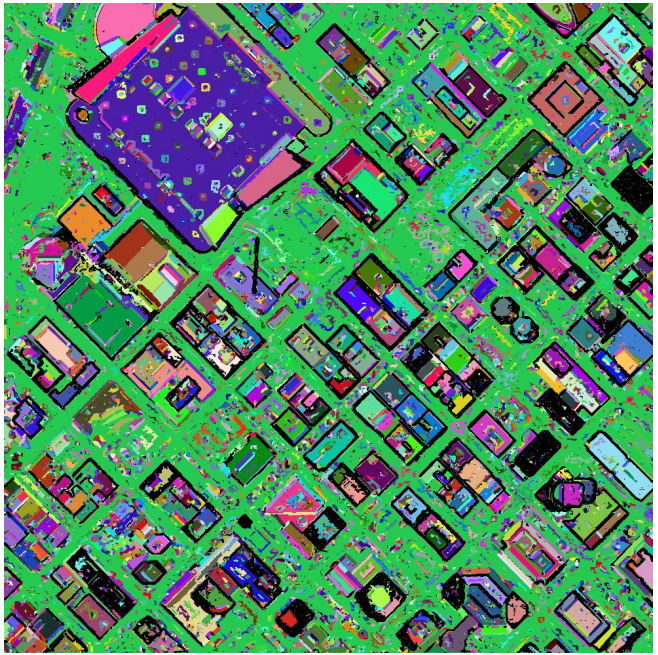
Multi-label Pixelwise Classification for Reconstruction of Large-scale Urban Areas, Y. He, S. Mudur, C. Poullis

Recent work



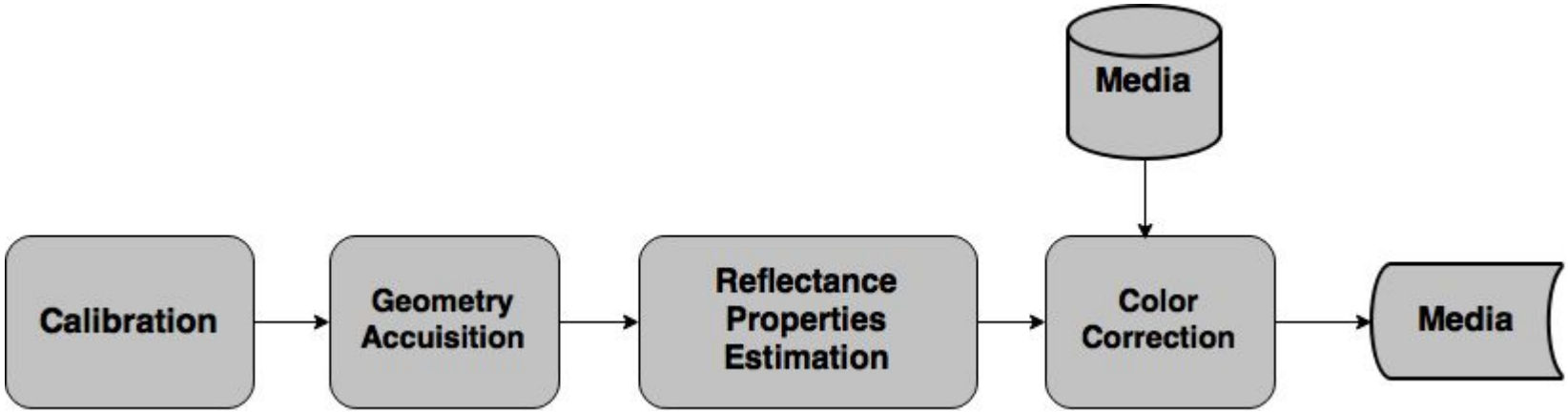
Recent work





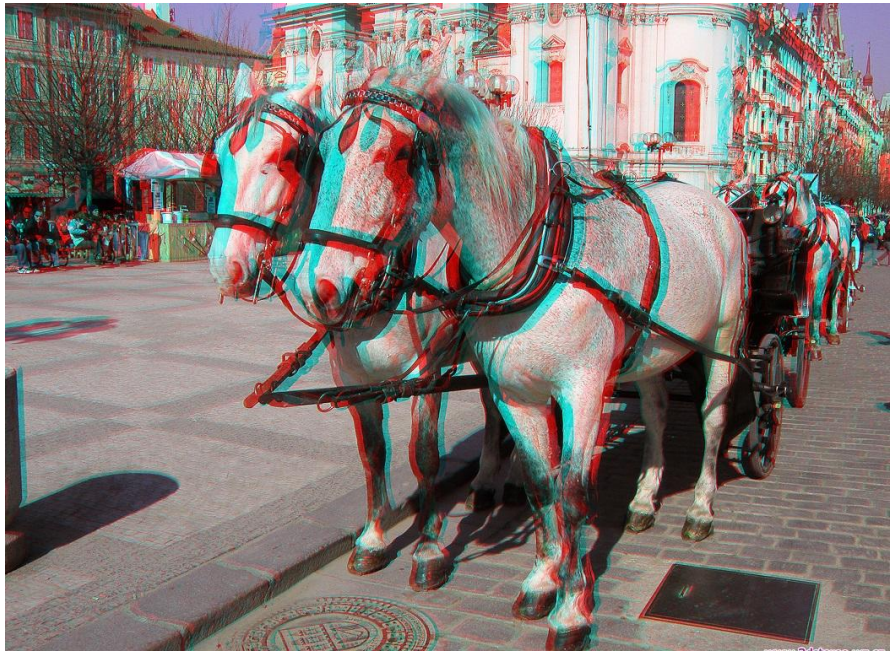
Recent work

A Long-Range Vision System for Projection Mapping of Stereoscopic Content in Outdoor Areas, B. Maneshgar, L. Sujir, S. Mudur, C. Poullis



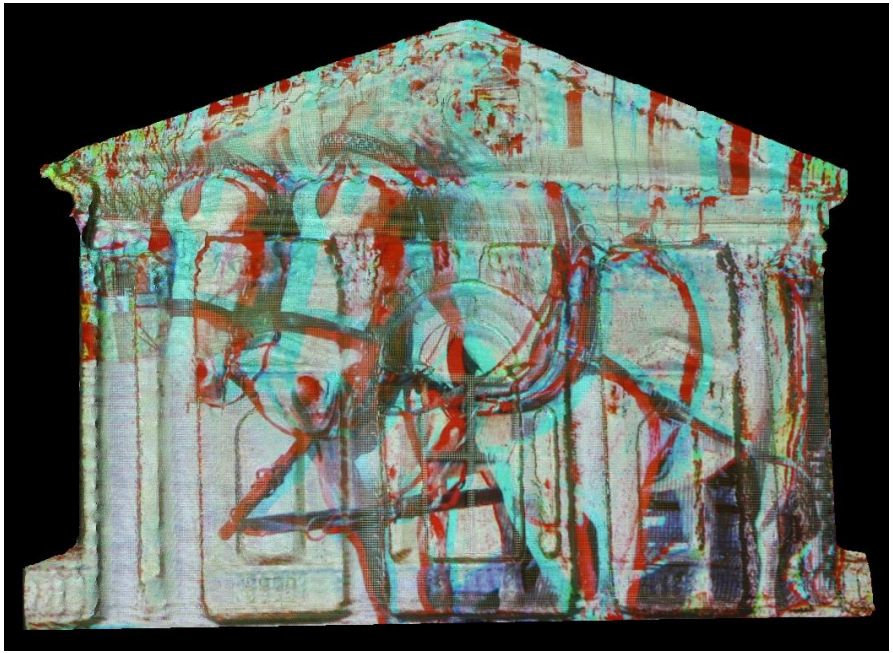
Recent work

A Long-Range Vision System for Projection Mapping of Stereoscopic Content in Outdoor Areas, B. Maneshgar, L. Sujir, S. Mudur, C. Poullis



Recent work

A Long-Range Vision System for Projection Mapping of Stereoscopic Content in Outdoor Areas, B. Maneshgar, L. Sujir, S. Mudur, C. Poullis



The goal of computer vision

- To extract “meaning” from pixels

What we see

Source: S. Narasimhan

The goal of computer vision

- To extract “meaning” from pixels

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

The goal of computer vision

- To extract “meaning” from pixels

Humans are remarkably good at this…

Source: “80 million tiny images” by Torralba et al.

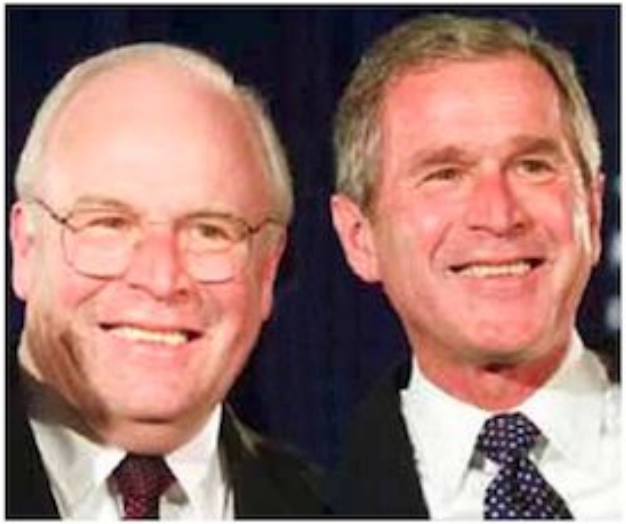
Can computers match (or beat) human vision?

Can computers match (or beat) human vision?

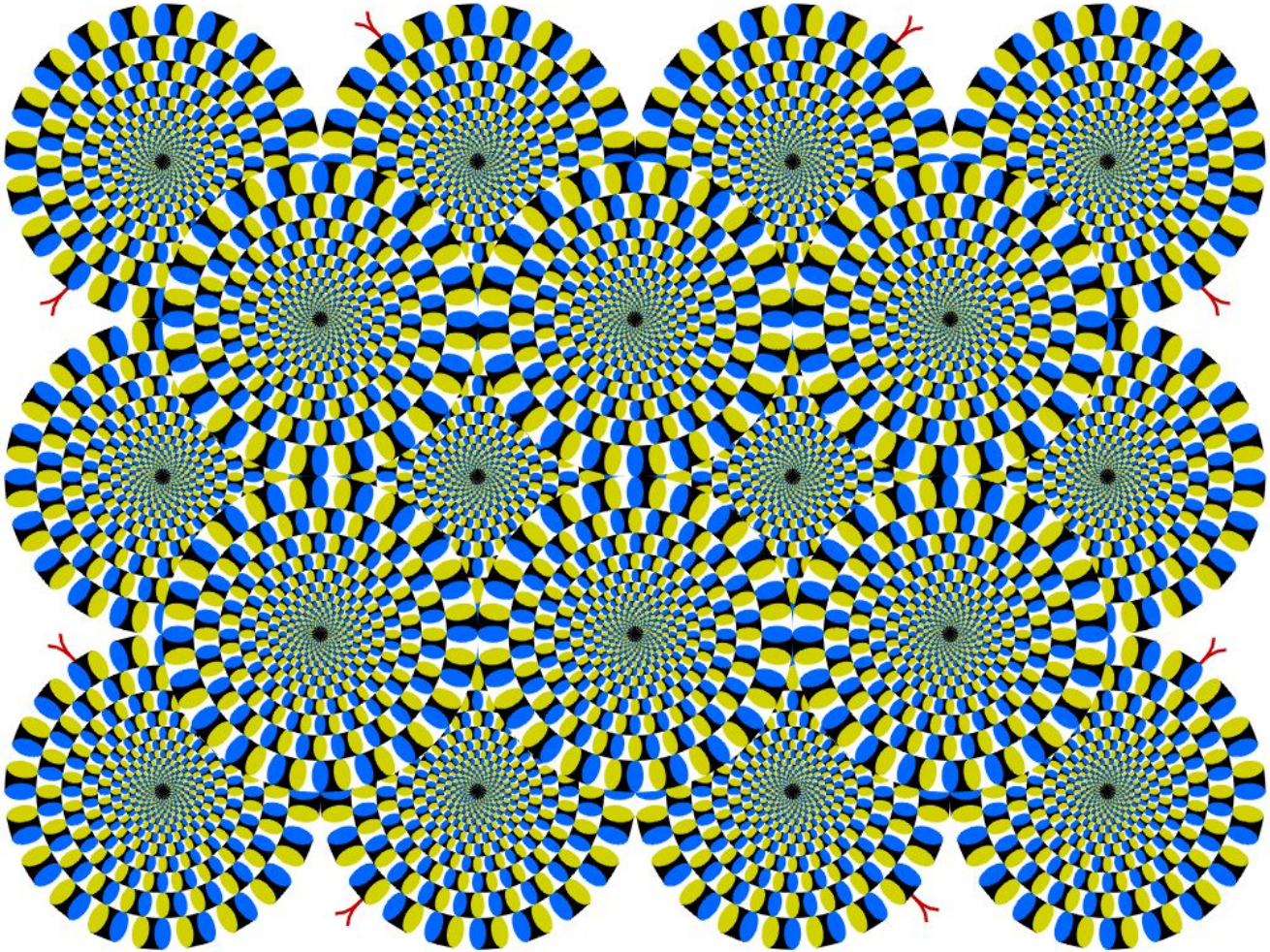
- Yes and no (but mostly no!)
  - humans are much better at “hard”
  - computers can be better at “easy”

Human perception has its shortcomings…

individuals. Unlike many computer vision systems and models of face processing that rely on fine details of the internal features and their precise spatial configuration, the encoding used by the human visual system appears to emphasize the whole head structure.



With the change of the administration in the White House, and the imminent publication of new editions of the books in which the illusion was reproduced, we have been urged by several colleagues to update the illusion. The new version ("Republican



What kind of information can be extracted from an image?

What kind of information can be extracted from an image?

Semantic information

What kind of information can be extracted from an image?

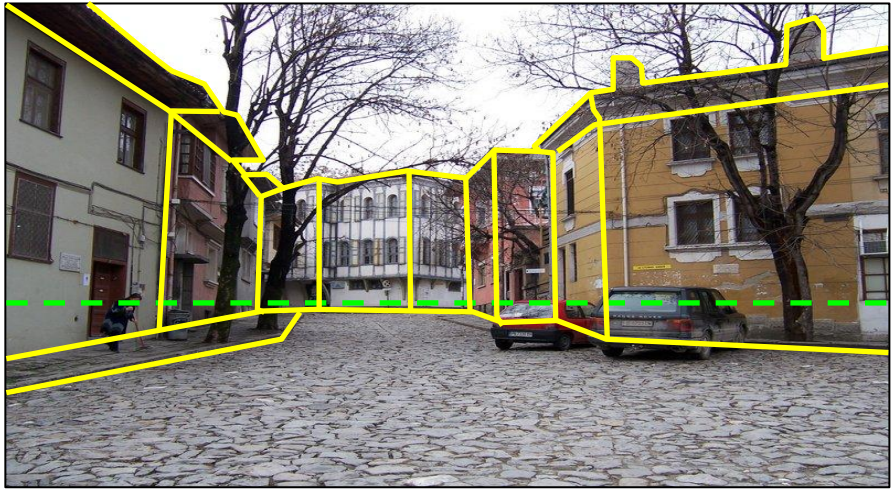
Semantic information

What kind of information can be extracted from an image?

Semantic information

What kind of information can be extracted from an image?

Semantic information



Geometric information

Why study computer vision?

Why study computer vision?

- Vision is useful

Why study computer vision?

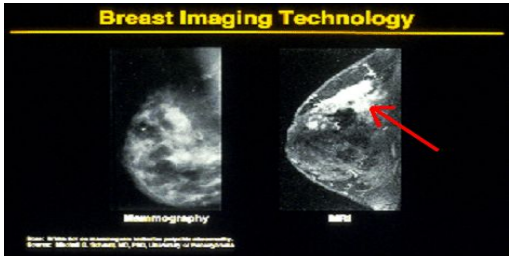
- Vision is useful
- Vision is interesting

Why study computer vision?

- Vision is useful
- Vision is interesting
- Vision is difficult
  - Half of primate cerebral cortex is devoted to visual processing
  - Achieving human-level visual perception is probably “AI-complete”

Why computer vision matters

Safety

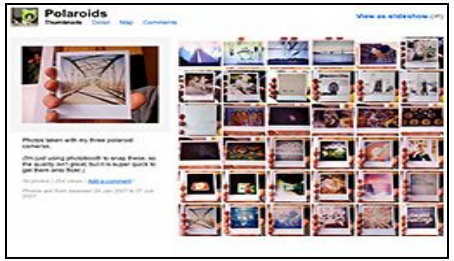


Health

Security

Comfort

Fun



Access

Ridiculously brief history of computer vision

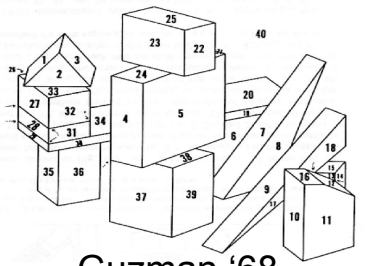
- 1966: Minsky assigns computer vision as an undergrad summer project

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"

Ridiculously brief history of computer vision

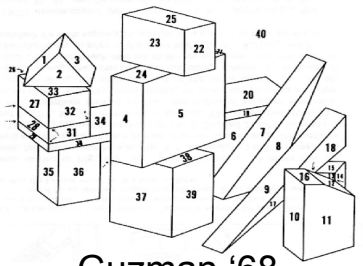
- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds



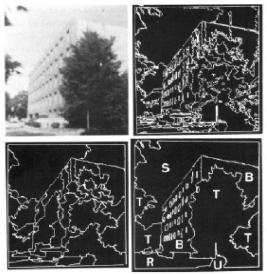
Guzman '68

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images



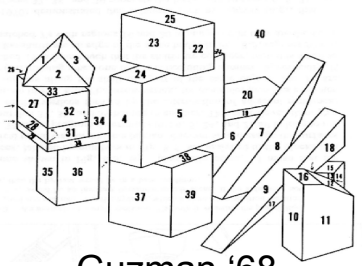
Guzman '68



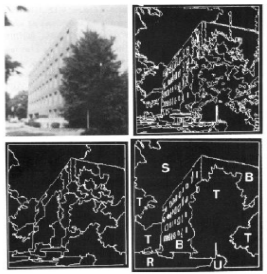
Ohta Kanade '78

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor



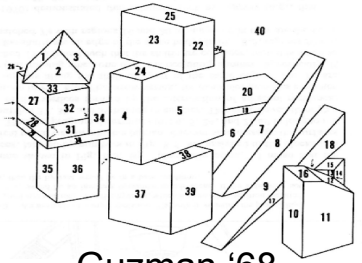
Guzman '68



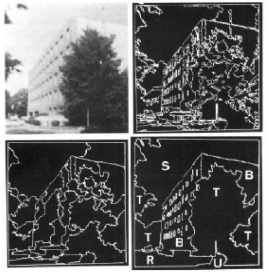
Ohta Kanade '78

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue



Guzman '68

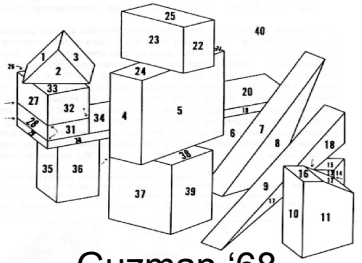


Ohta Kanade '78

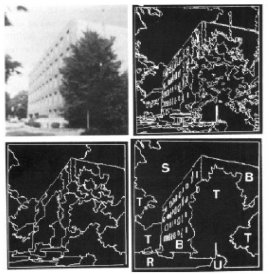
Turk and Pentland '91

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts



Guzman '68

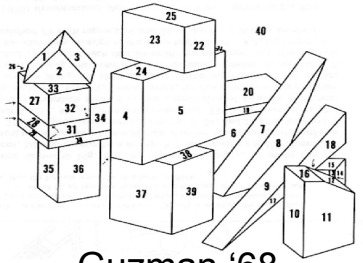


Ohta Kanade '78

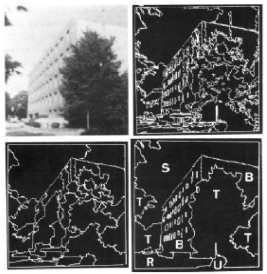
Turk and Pentland '91

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: Deep learning with ConvNets



Guzman '68

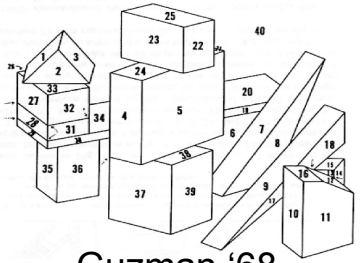


Ohta Kanade '78

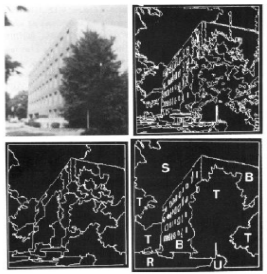
Turk and Pentland '91

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
  - "spend the summer linking a camera to a computer and getting the computer to describe what it saw"
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: Deep learning with ConvNets
- 2030's: robot uprising?



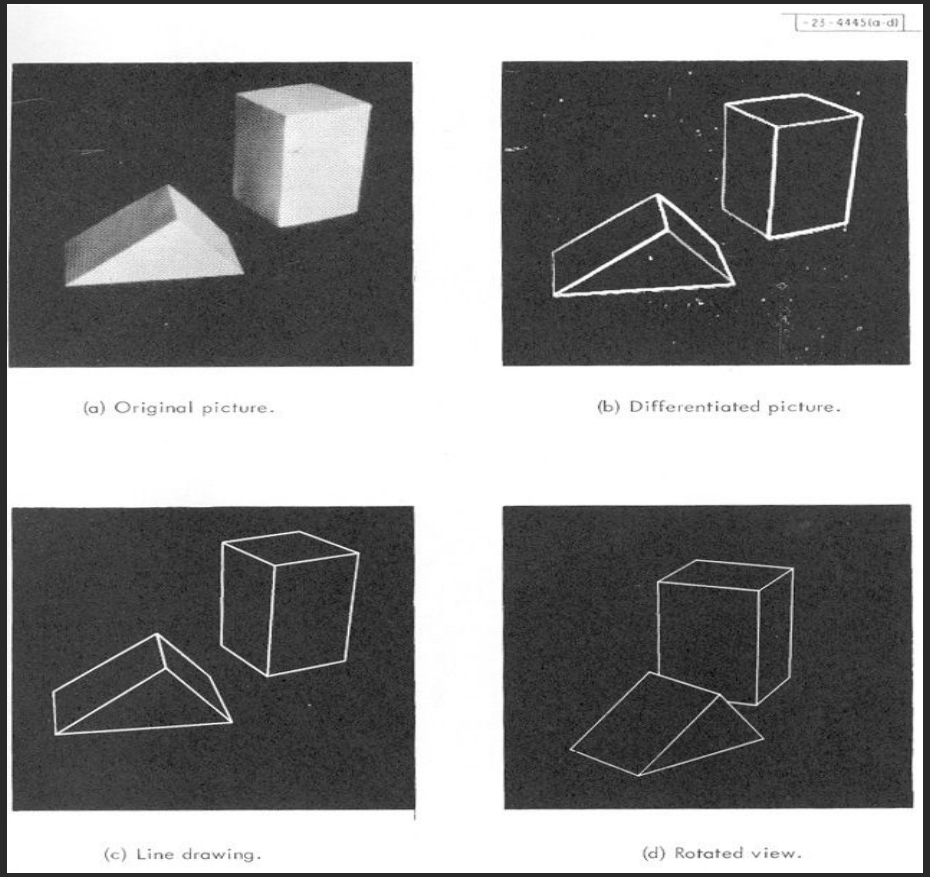
Guzman '68



Ohta Kanade '78

Turk and Pentland '91

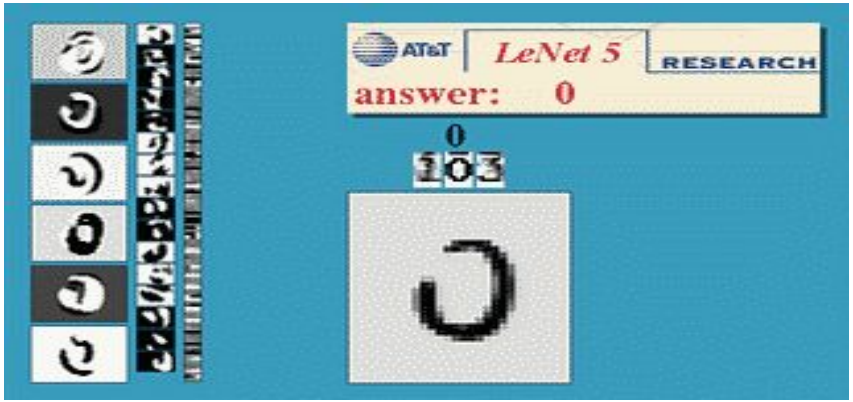
Origins of computer vision



L. G. Roberts, *Machine Perception of Three Dimensional Solids*, Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

Successes of computer vision to date

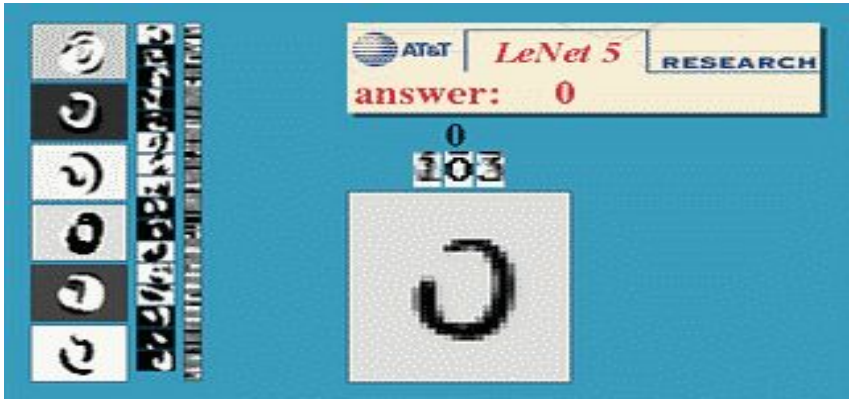
Optical character recognition (OCR)



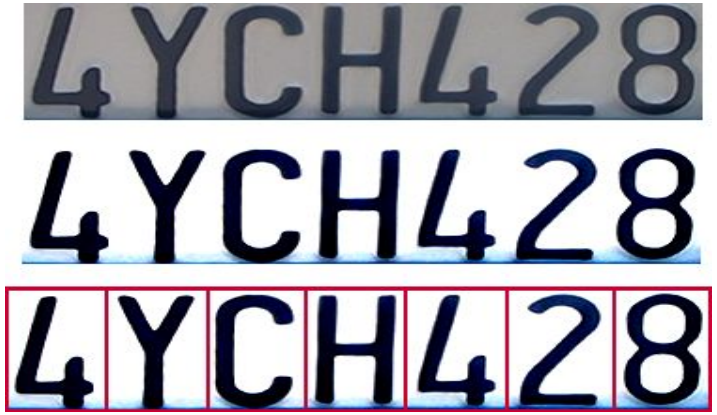
Digit recognition
yann.lecun.com

Source: S. Seitz, N. Snavely

Optical character recognition (OCR)



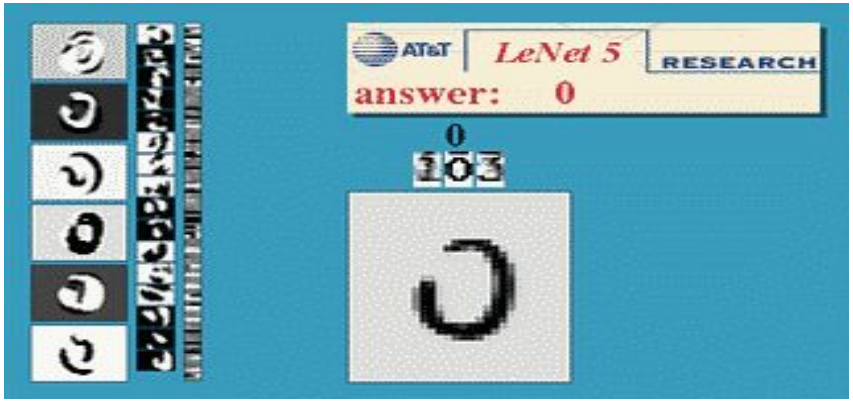
Digit recognition
yann.lecun.com



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

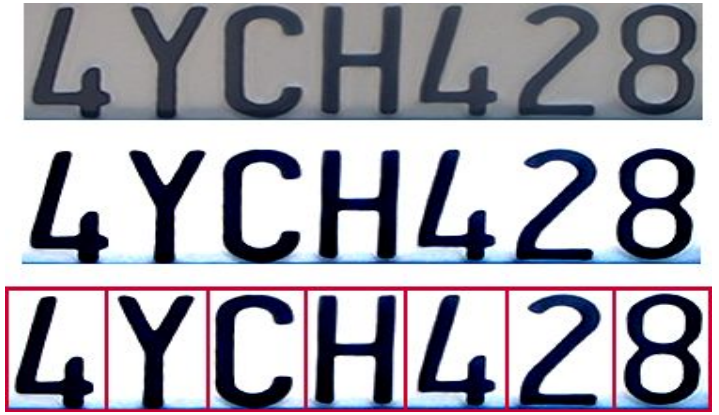
Source: S. Seitz, N. Snavely

Optical character recognition (OCR)



Digit recognition

yann.lecun.com



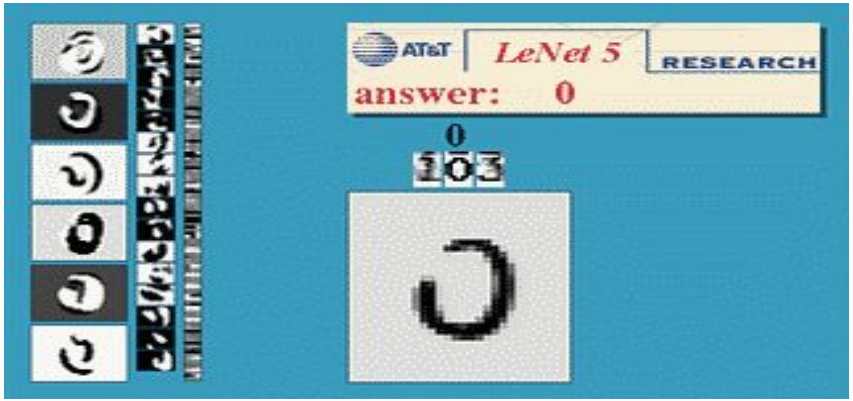
License plate readers

http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Automatic check processing

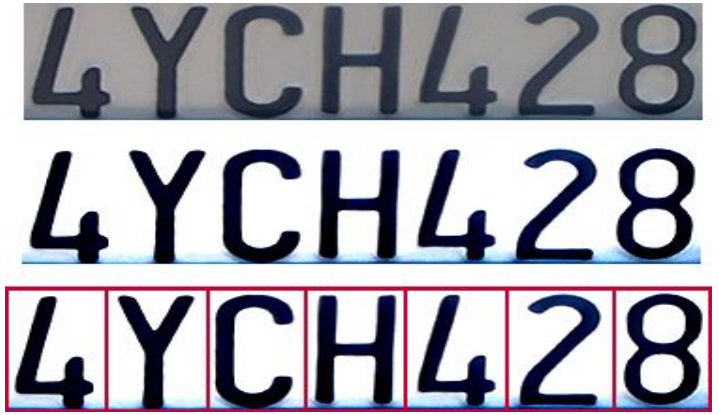
Source: S. Seitz, N. Snavely

Optical character recognition (OCR)



Digit recognition

yann.lecun.com



License plate readers

http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Automatic check processing

Sudoku grabber

http://sudokugrab.blogspot.com/

https://www.youtube.com/watch?v=olmMJ6p6mKE

Source: S. Seitz, N. Snavely

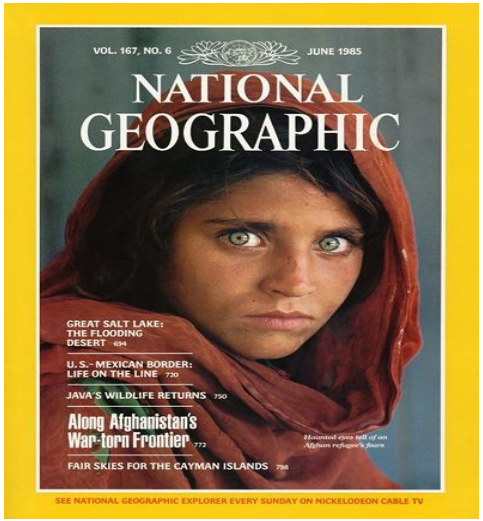
Biometrics

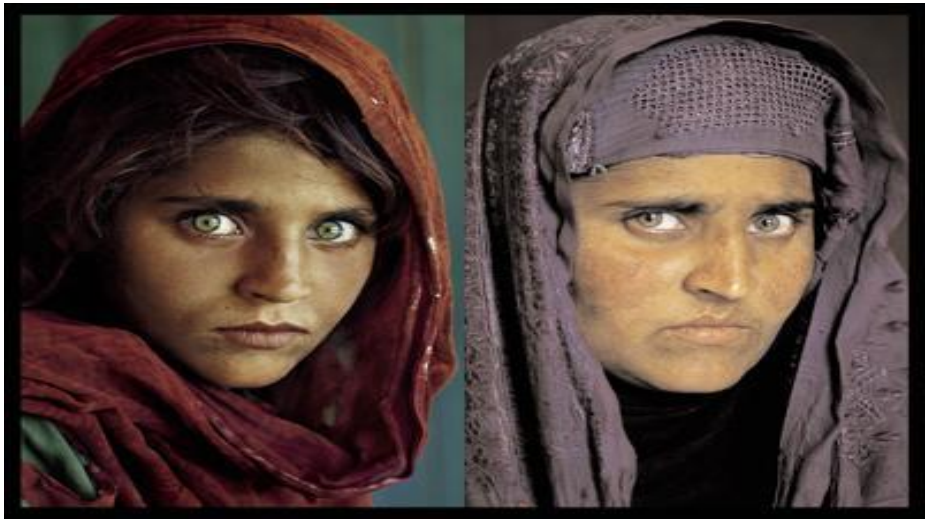
Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

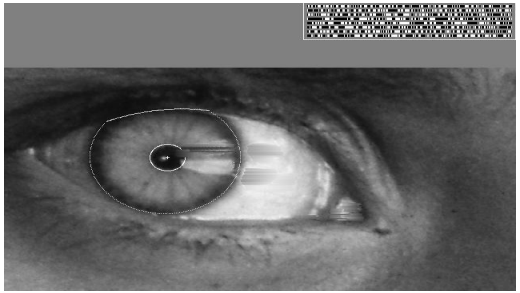
Source: S. Seitz

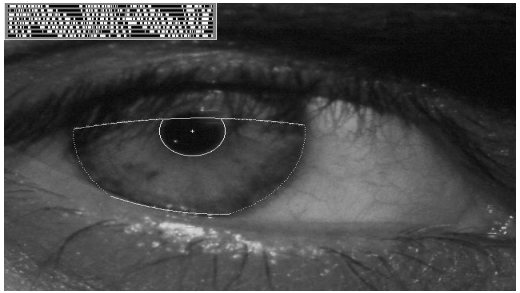
Biometrics





[How the Afghan Girl was Identified by Her Iris Patterns]





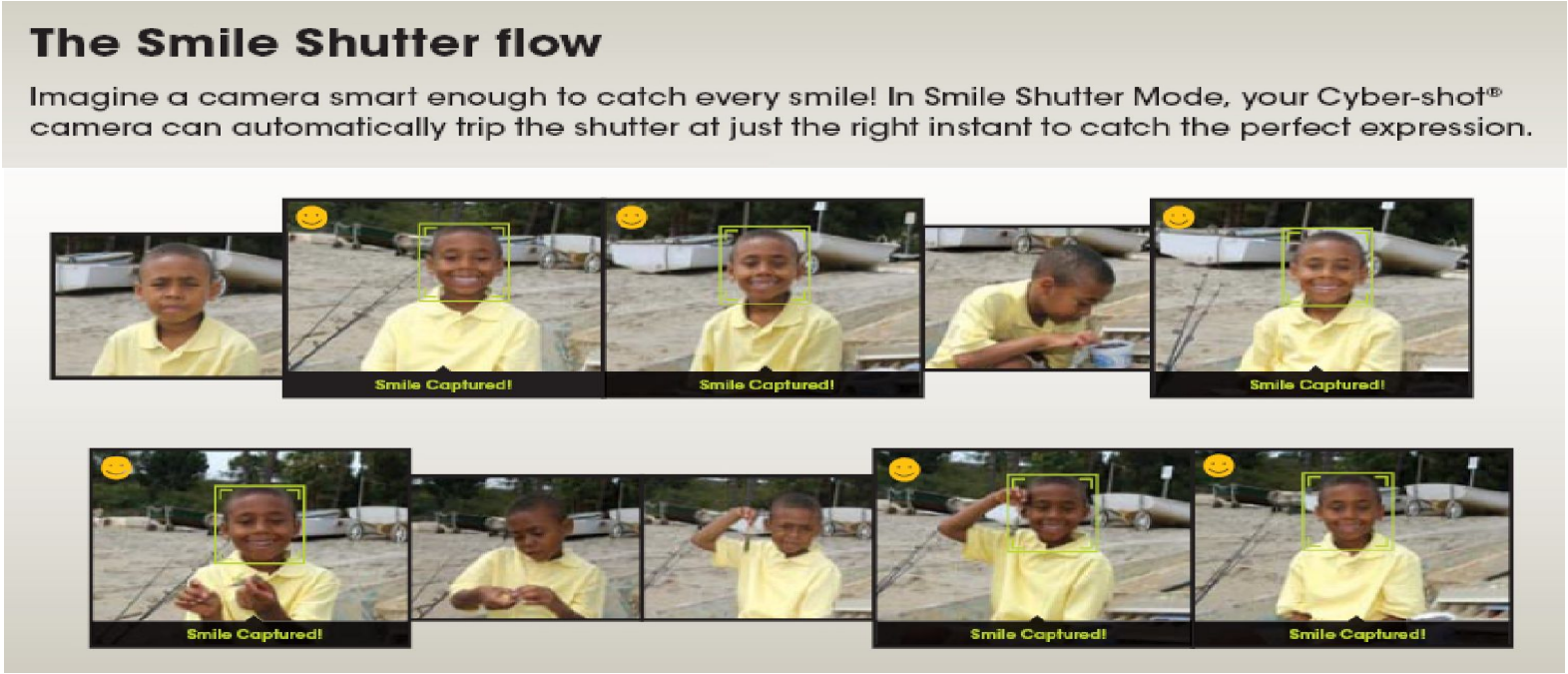
Source: S. Seitz

Face detection

Many consumer digital cameras now detect faces

Source: S. Seitz

Smile detection



Sony Cyber-shot® T70 Digital Still Camera

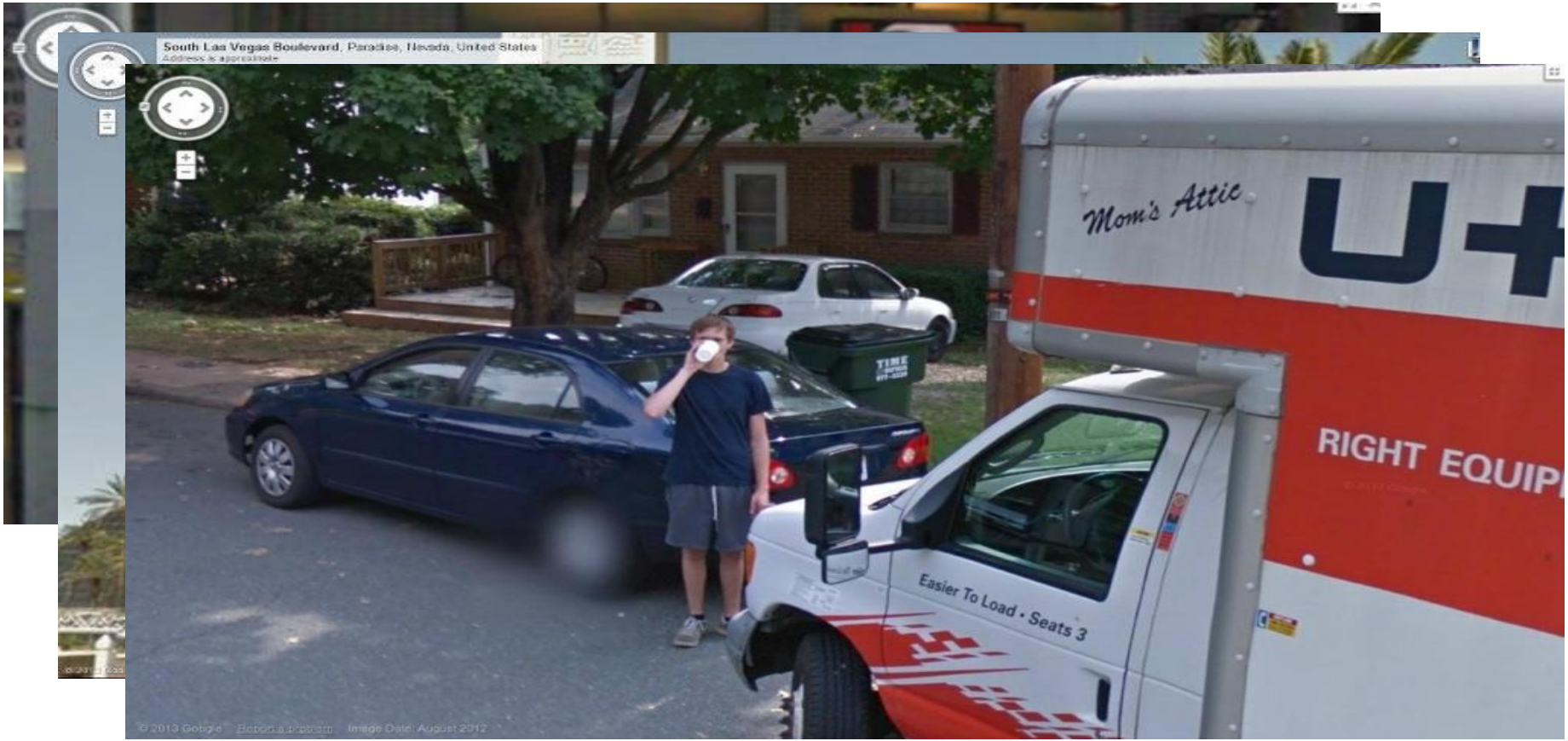
Source: S. Seitz

Face detection for privacy protection

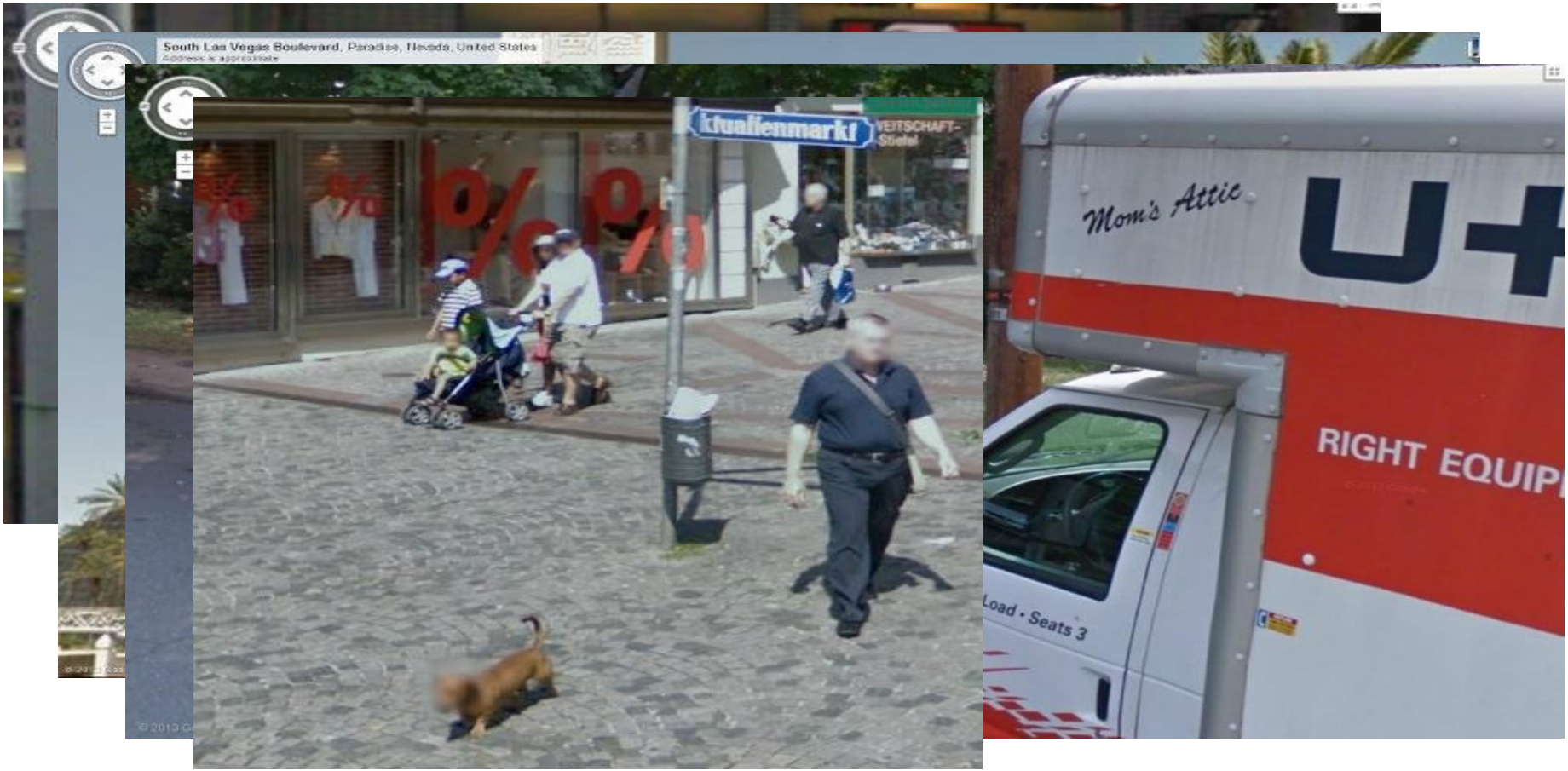
Technology gone wild…

Technology gone wild…

Technology gone wild…



Technology gone wild…

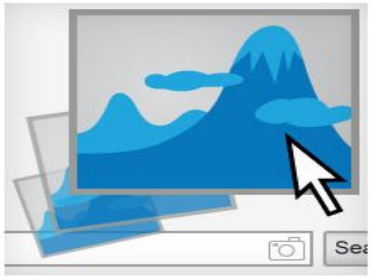


Face recognition: Apple iPhoto software

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

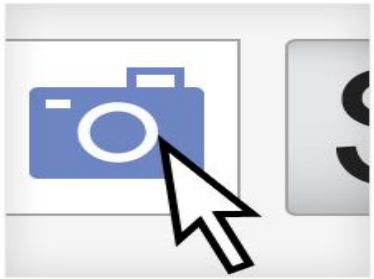
Visual search: Google search by image

Four ways to search by image



Drag and drop

Drag and drop an image from the web or your computer into the search box on images.google.com.



Upload an image

On images.google.com, click the camera icon, then select "Upload an image." Select the image you want to use to start your search.

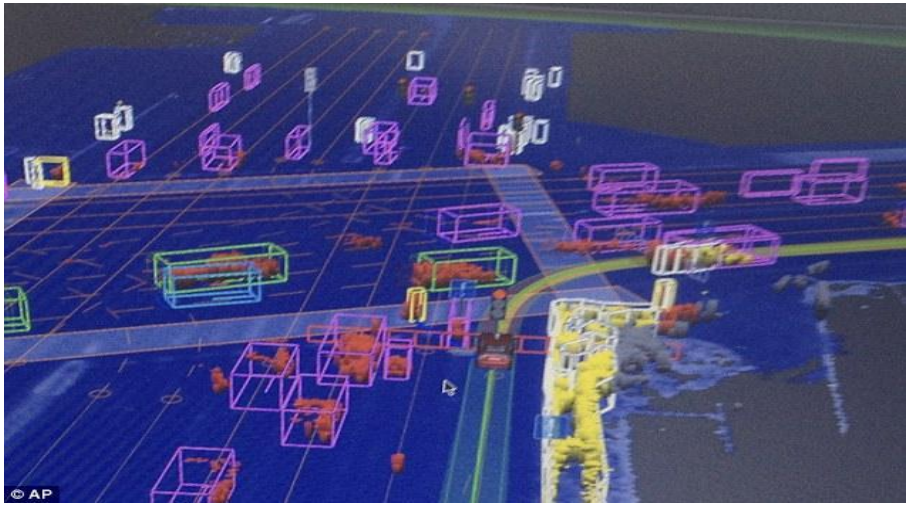
Copy and paste the URL for an image

Found an image on the web you're curious about? Right-click the image to copy the URL. On images.google.com, click the camera icon, and "Paste image URL".

Right-click an image on the web

To search by image even faster, download the Chrome extension or the Firefox extension. With the extension installed, simply right-click an image on the web to search Google with that image.

Google self-driving cars



https://www.youtube.com/watch?v=bDOnn0-4Nq8

- Google's self-driving car passes 300,000 miles (Forbes, 8/15/2012)
- Nissan pledges affordable self-driving car models by 2020 (CNET, 8/27/2013)

Automotive safety

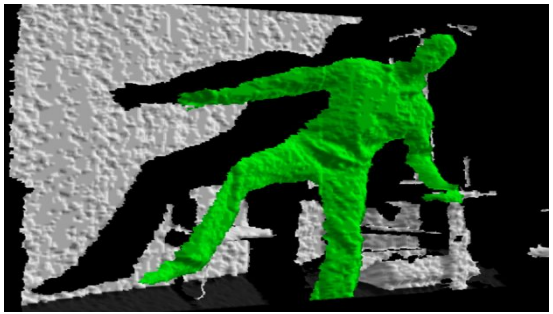
[Mobileye](): Vision systems in high-end BMW, GM, Volvo models

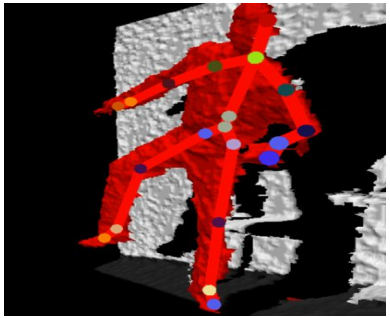
- Pedestrian collision warning
- Forward collision warning
- Lane departure warning
- Headway monitoring and warning

https://www.youtube.com/watch?v=HXpiyLUEOOY

Source: A. Shashua, S. Seitz

Vision-based interaction: Xbox Kinect





3D Reconstruction: Kinect Fusion

YouTube
Video

Occipital: 3D structure sensor for iPad

http://www.kickstarter.com/projects/occipital/structure-sensor-capture-the-world-in-3d

http://structure.io/#home-about-vid

3D reconstruction from photo collections

Colosseum, Rome, Italy

San Marco Square, Venice, Italy

Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. Seitz, The Visual Turing Test for Scene Reconstruction, 3DV 2013

YouTube Video

Object recognition

[Google Goggles]
[Bing Vision]

Special effects: shape and motion capture

Source: S. Seitz

Vision in space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

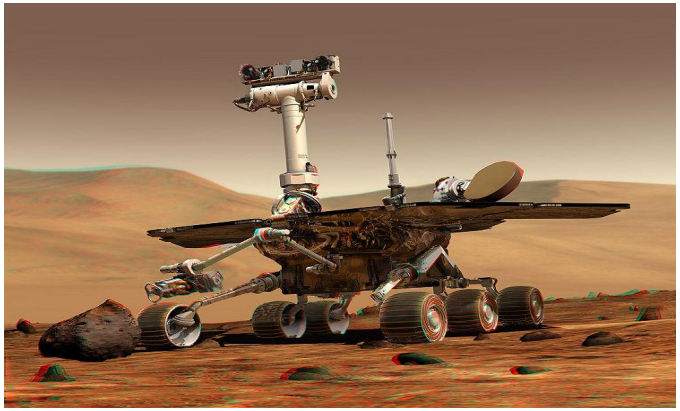
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

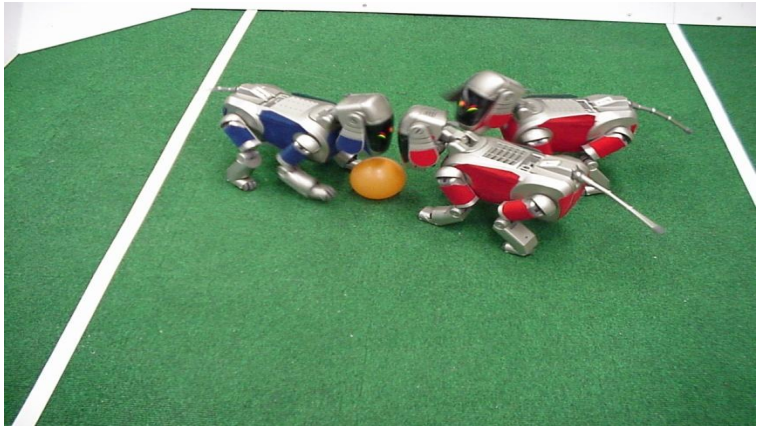
Industrial robots

[Vision-guided robots position nut runners on wheels]

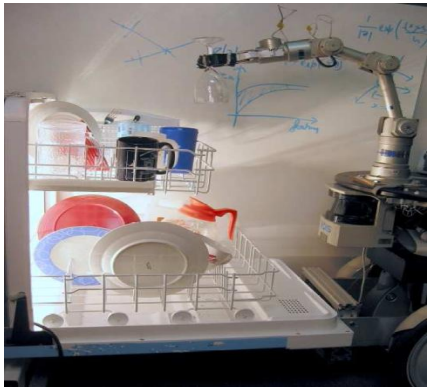
Mobile robots



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

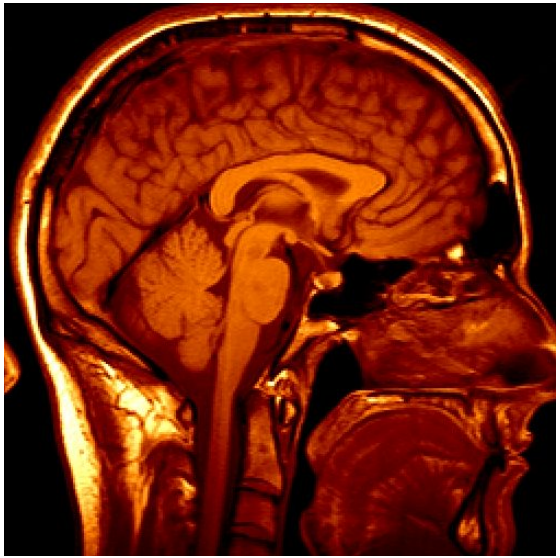


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

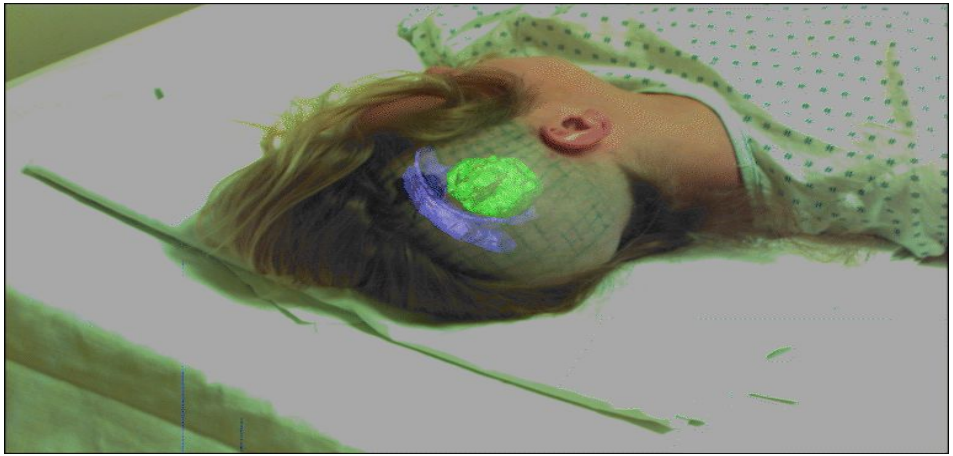
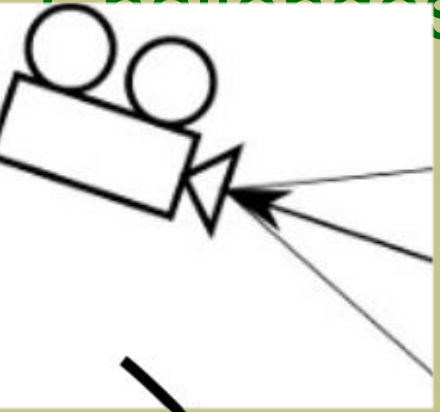
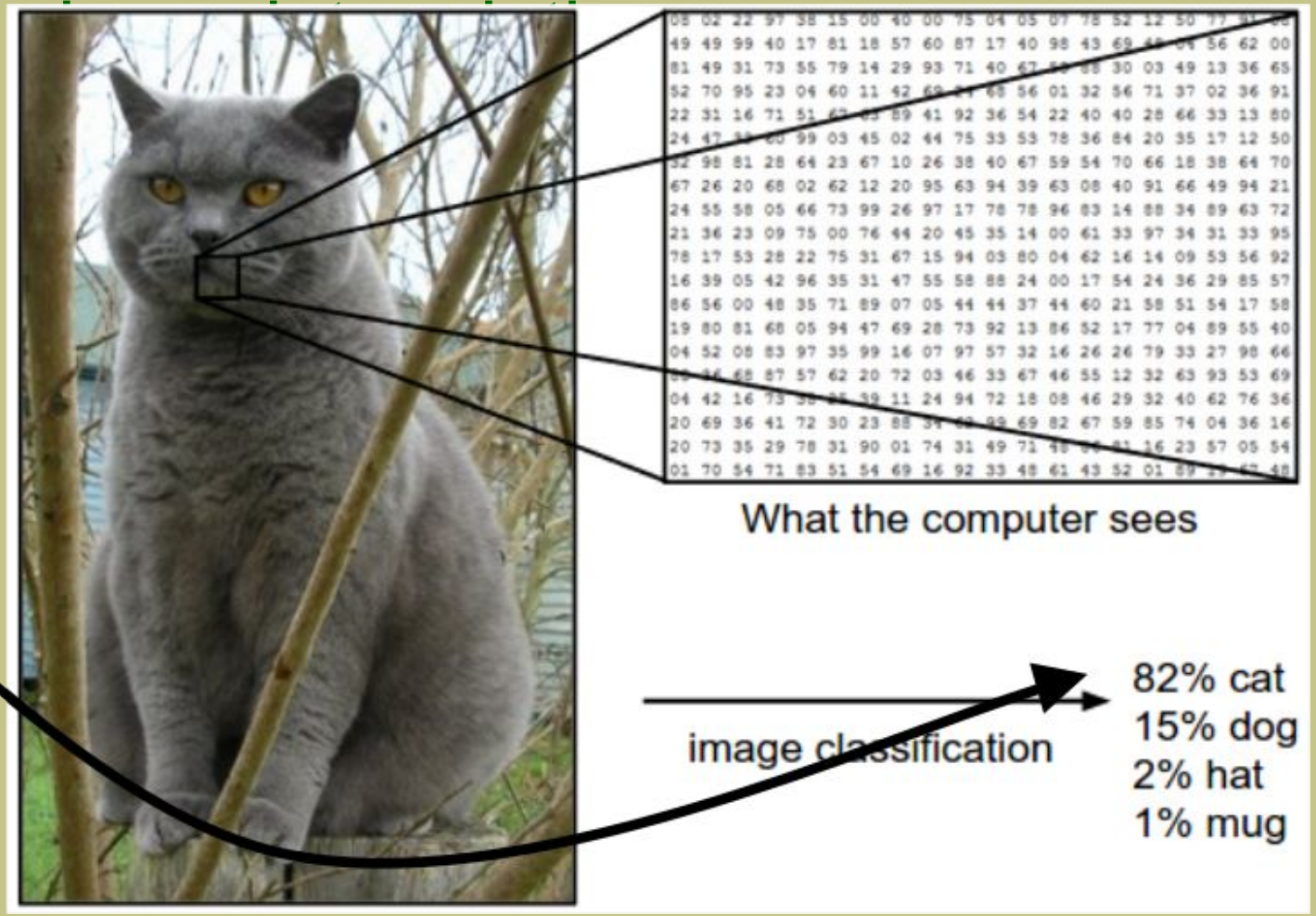


Image guided surgery
Grimson et al., MIT

Why is computer vision difficult?

Challenges:





slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges: illumination

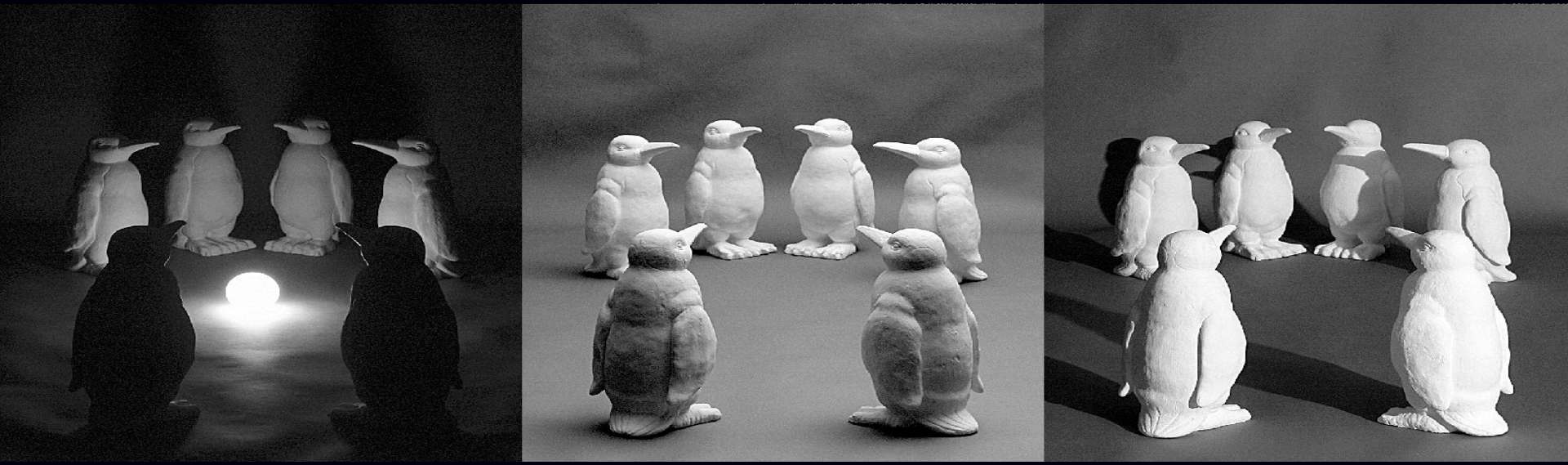
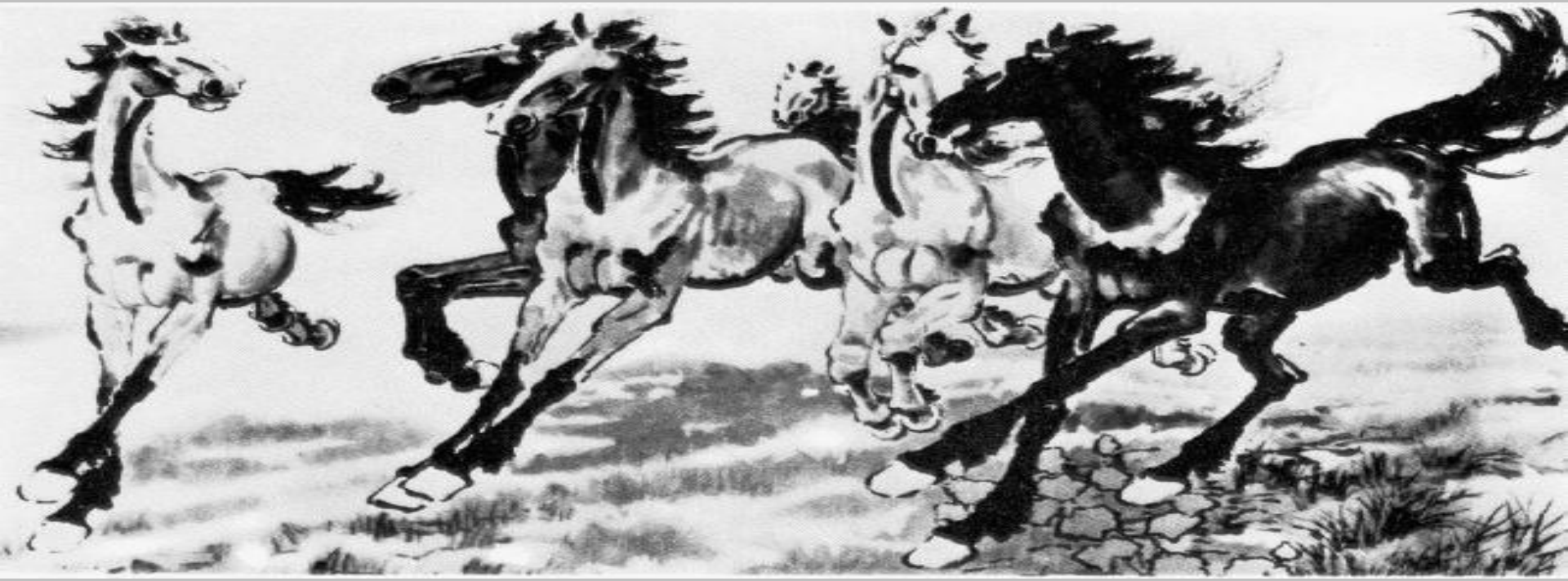


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation

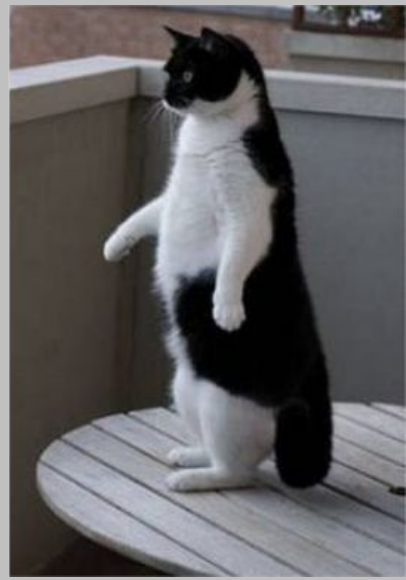


Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation



slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: object intra-class variation

slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges: occlusion

slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges:clutter

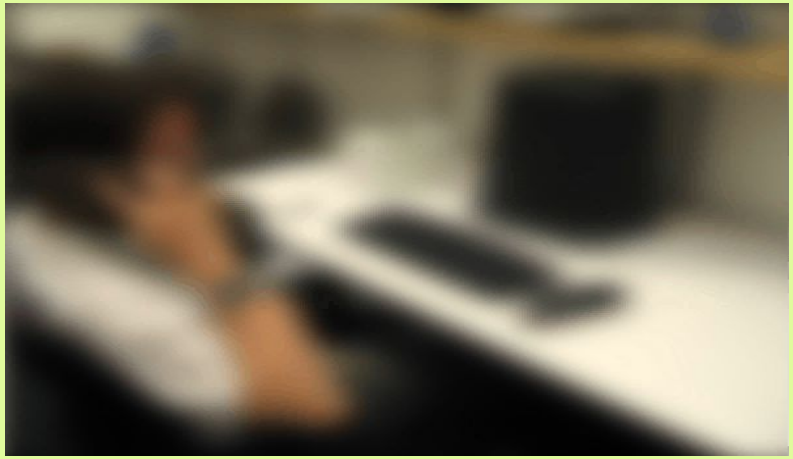
slide credit: Fei Fei Li, Andrej Karpathy, Justin Johnson

Challenges: occlusion, clutter

Image source: National Geographic

Challenges: Motion

Challenges: ambiguity

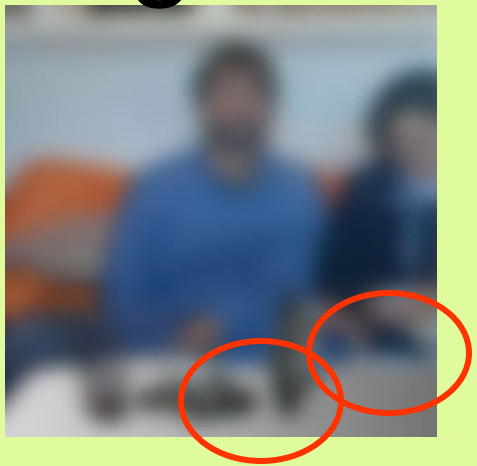


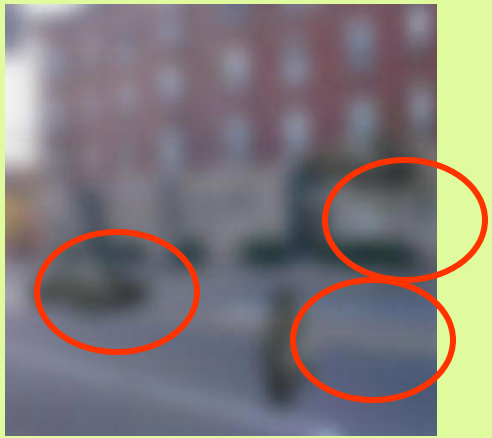
Source: Rob Fergus and Antonio Torralba

Challenges: ambiguity

Source: Rob Fergus and Antonio Torralba

Challenges: ambiguity





slide credit: Fei-Fei, Fergus & Torralba

Challenges: ambiguity

- Many different 3D scenes could have given rise to a particular 2D picture

Review: Intro to computer vision

- State-of-the-art applications
- Challenges of vision
  - Viewpoint and lighting variation
  - Intra-class variations: size, shape, deformation, etc.
  - Nuisances: motion, blur, noise, etc.
  - Intrinsic ambiguity

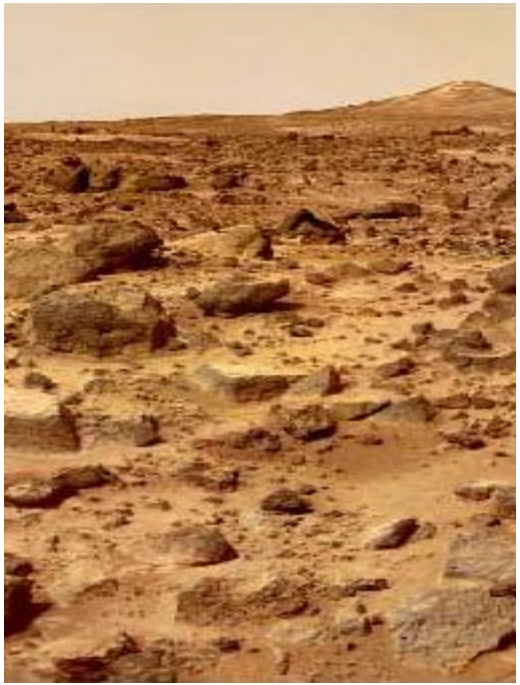
Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

Depth cues: Linear perspective

Depth cues: Parallax

Depth cues: Parallax



Shape cues: Texture gradient

Shape and lighting cues: Shading

Grouping cues: Similarity (color, texture, proximity)

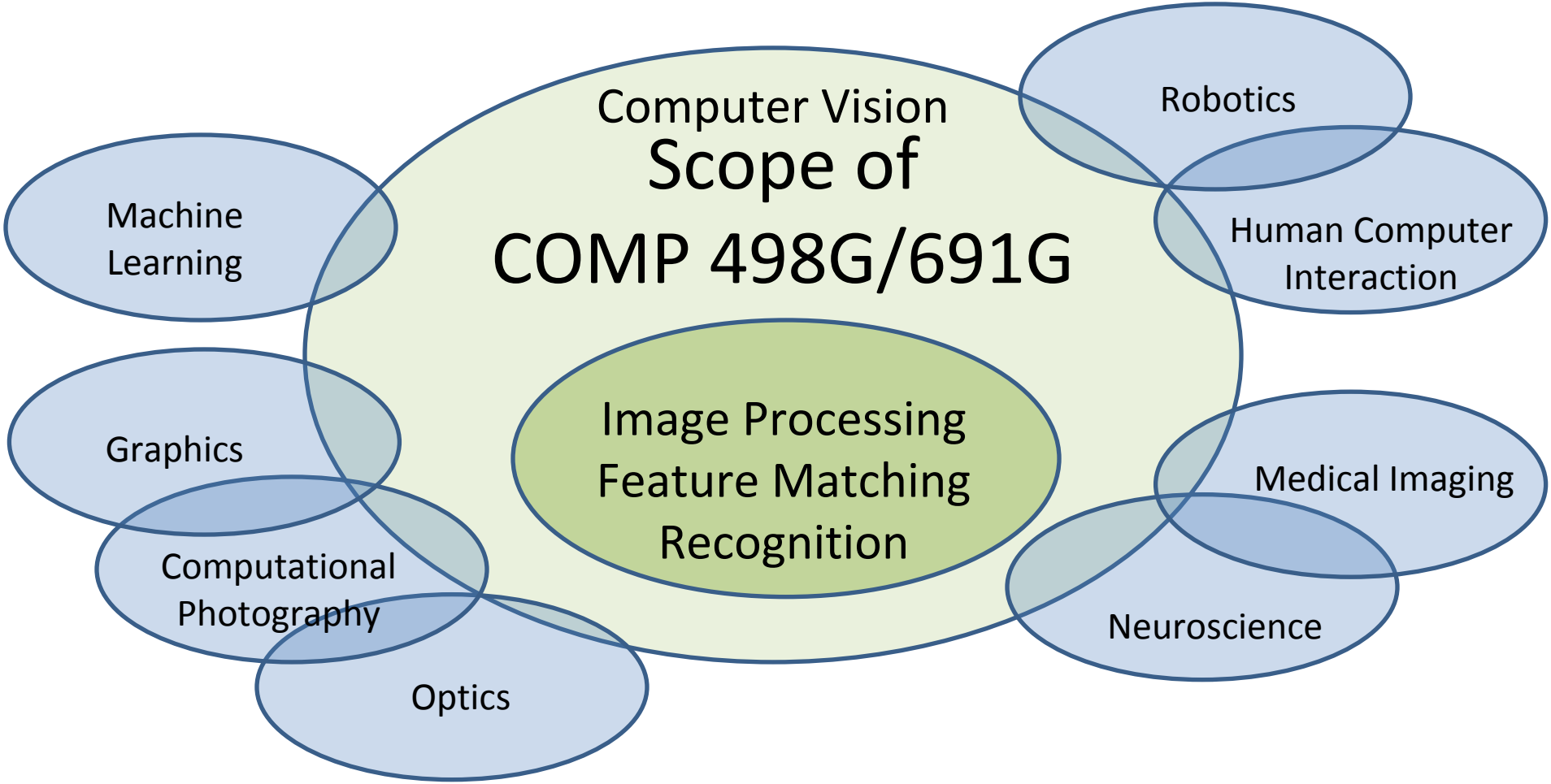
Grouping cues: “Common fate”

Image credit: Arthus-Bertrand (via F. Durand)

The computer vision industry

- A list of companies here:

  http://www.cs.ubc.ca/spider/lowe/vision.html


Computer Vision
Scope of
COMP 498G/691G
Robotics
Machine Learning
Human Computer Interaction
Image Processing
Feature Matching
Recognition
Graphics
Medical Imaging
Computational Photography
Neuroscience
Optics

Computer Vision and Nearby Fields

- Computer Graphics: Models to Images
- Comp. Photography: Images to Images
- Computer Vision: Images to Models

Course Topics

- Interpreting Intensities
  - What determines the brightness and color of a pixel?
  - How can we use image filters to extract meaningful information from the image?
- Correspondence and Alignment
  - How can we find corresponding points in objects or scenes?
  - How can we estimate the transformation between them?
- Grouping and Segmentation
  - How can we group pixels into meaningful regions?
- Categorization and Object Recognition
  - How can we represent images and categorize them?
  - How can we recognize categories of objects?
- Advanced Topics
  - Action recognition, 3D scenes and context, human-in-the-loop vision…

Course Website

http://poullis.org/courses/2017/Winter/COMP498G-691G/

- lecture notes are posted the day of the class
- bring to class and keep notes

- The lecture notes are password protected
  - user: comp_vision
  - pass: 2017_notes

Home | Course Syllabus | Course Schedule | Assignments | Exams | Links

Instructor: Charalambos Poullis

Programmer-on-Duty [POD]: TBA

Office: EV3.183

Email: charalambos [at] poullis [dot] org

Lectures: TBA @ xx:xx – xx:xx, FG-B055

POD office hours: TBA

Office hours: Wednesday xx:xx - xx:xx and by appointment

Other: Please see below

COURSE DESCRIPTION

Over the past few years computer vision has re-emerged as one of the most popular and challenging technical areas in computer science. Recent advances in computer graphics hardware as well as the development of novel algorithms for the fast and accurate extraction of salient features in images, have made it possible to significantly progress the state-of-the-art to a point where nowadays many commercial products incorporate some type of embedded computer vision system.

The goal of the proposed course is to introduce graduate students to the different aspects of computer vision, give them the ability to understand, apply, analyze and evaluate computer vision algorithms, and implement components that are fundamental to many modern vision systems.

Prerequisites: N/A

Suggested textbooks:
1. Computer Vision: A Modern Approach (2nd Edition) by David A. Forsyth, Jean Ponce. ISBN-13: 978-0136085928
2. Computer Vision: Algorithms and Applications Hardcover by Richard Szeliski. ISBN-13: 978-1848829343

Announcements

[01.02.2016] Welcome to the course!

Immersive and Creative Technologies Lab | Copyright © Charalambos Poullis | Concordia University

CHECK REGULARLY

Syllabus - Course Outline

Posted on the website

Pre-requisites

- Data structures
- A good working knowledge of C and C++ programming
- Linear algebra
- Vector Calculus

- No prior knowledge of image processing, computer graphics or computer vision is assumed

Grading

- Programming assignments (x3) (45%)
- Project (20%)
- Quiz 1 (15%)
- Quiz 2 (15%)
- Participation (5%)

- Late submission policy
  - Extra 7 days at the cost of 2 points [out of 10]
  - No extra credit will be awarded

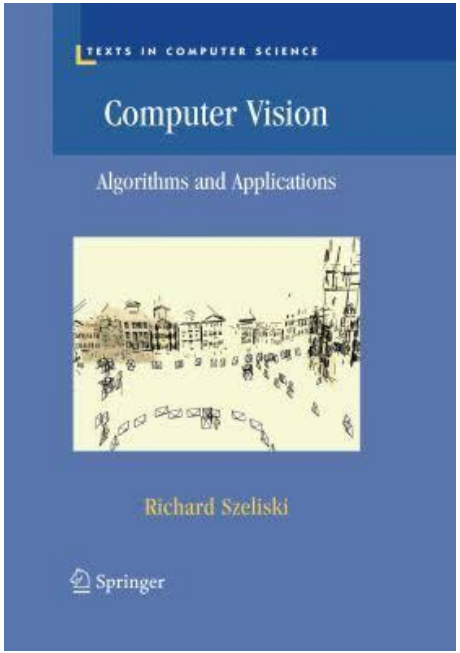
Academic Integrity Policy

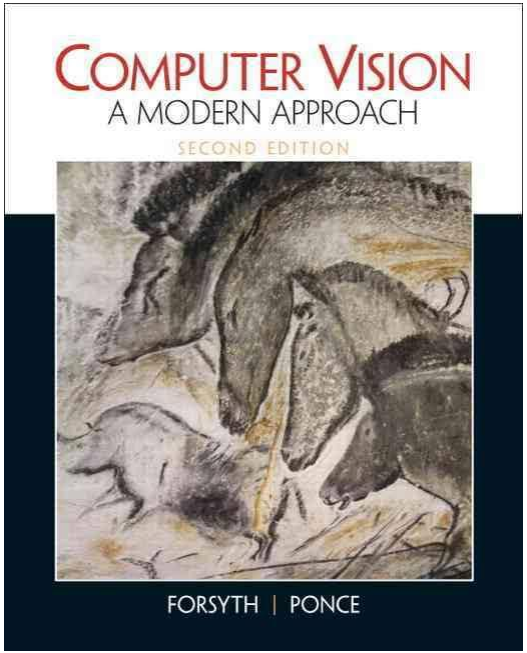
- Feel free to discuss assignments with each other, but coding must be done individually
- Feel free to incorporate code or tips you find on the Web, provided this doesn't make the assignment trivial and you explicitly acknowledge your sources
- Remember: we can Google too!

Submissions

- Electronic Assignment Submission
  - https://fis.encs.concordia.ca/eas/
- Late submissions must be sent to the TA via email

Suggested Textbooks





http://szeliski.org/Book/

Labs

- Start on Tuesday, 17th January 2017
- Tutorials in OpenCV
  - OpenCV API

Slides Acknowledgements

The slides are a combination of multiple resources and materials generously made publicly available by

S. Lazebnik,
D. Forsyth,
J. Ponce,
J. Koenderink,
S. Seitz,

R. Szeliski,
B. Freeman,
M. Pollefeys,
D. Lowe,
K. Grauman,

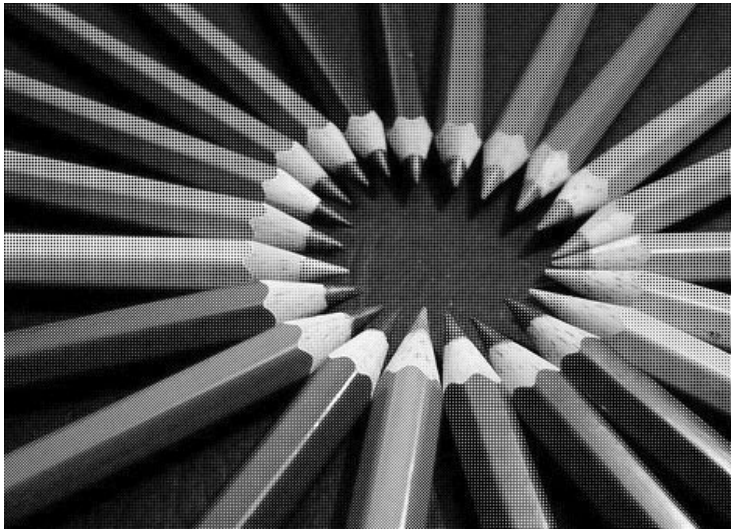
A. Efros,
F. Durand,
L. Fei-Fei,
A. Torralba,
R. Fergus,
J. Hays

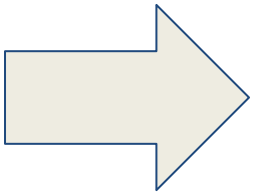
Assignments

- Demosaicing
- Hybrid Images
- Feature Detection and Matching
- Image Stitching
- [Tentative] 3D Reconstruction from Photometric Stereo

Demosaicing

- Implement the conversion of the Bayer pixel pattern to an RGB representation where each pixel has red, green and blue color channels





Hybrid Images

- Implement image filtering to separate high and low frequencies
- Combine high frequencies and low frequencies from different images to create an image with scale-dependent interpretation

Feature Detection and Matching

- Implement a feature detection and matching algorithm
- Feed feature matches to a structure-from-motion system

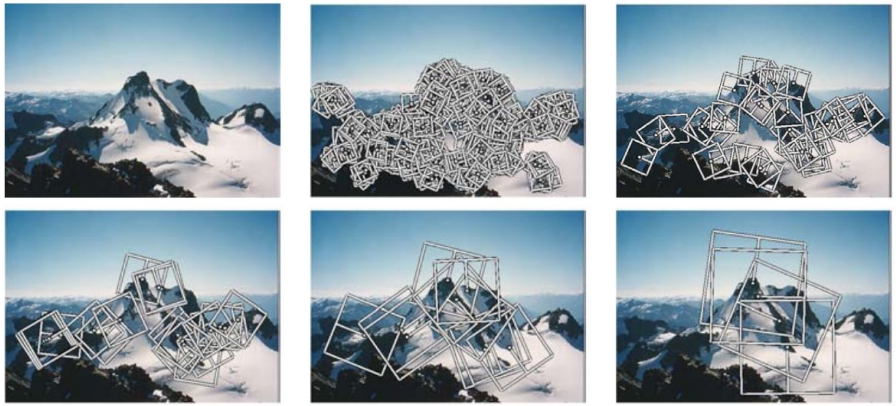


Figure 4.10 Multi-scale oriented patches (MOPS) extracted at five pyramid levels (Brown, Szeliski, and Winder 2005) © 2005 IEEE. The boxes show the feature orientation and the region from which the descriptor vectors are sampled.

Image credit: S. Szeliski

Image credit: J. Hays

Image Stitching

- Implement a feature-based alignment and image stitching algorithm

Image credit: S. Szeliski

3D Reconstruction from Photometric Stereo

- Implement an algorithm to construct a height field from a series of 2D images.

CONCORDIA.CA