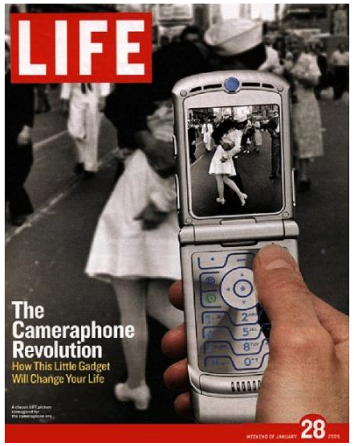
Introduction to Computer Vision

Introduction

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Welcome!


LIFE
The Cameraphone Revolution
How This Little Gadget Will Change Your Life
28

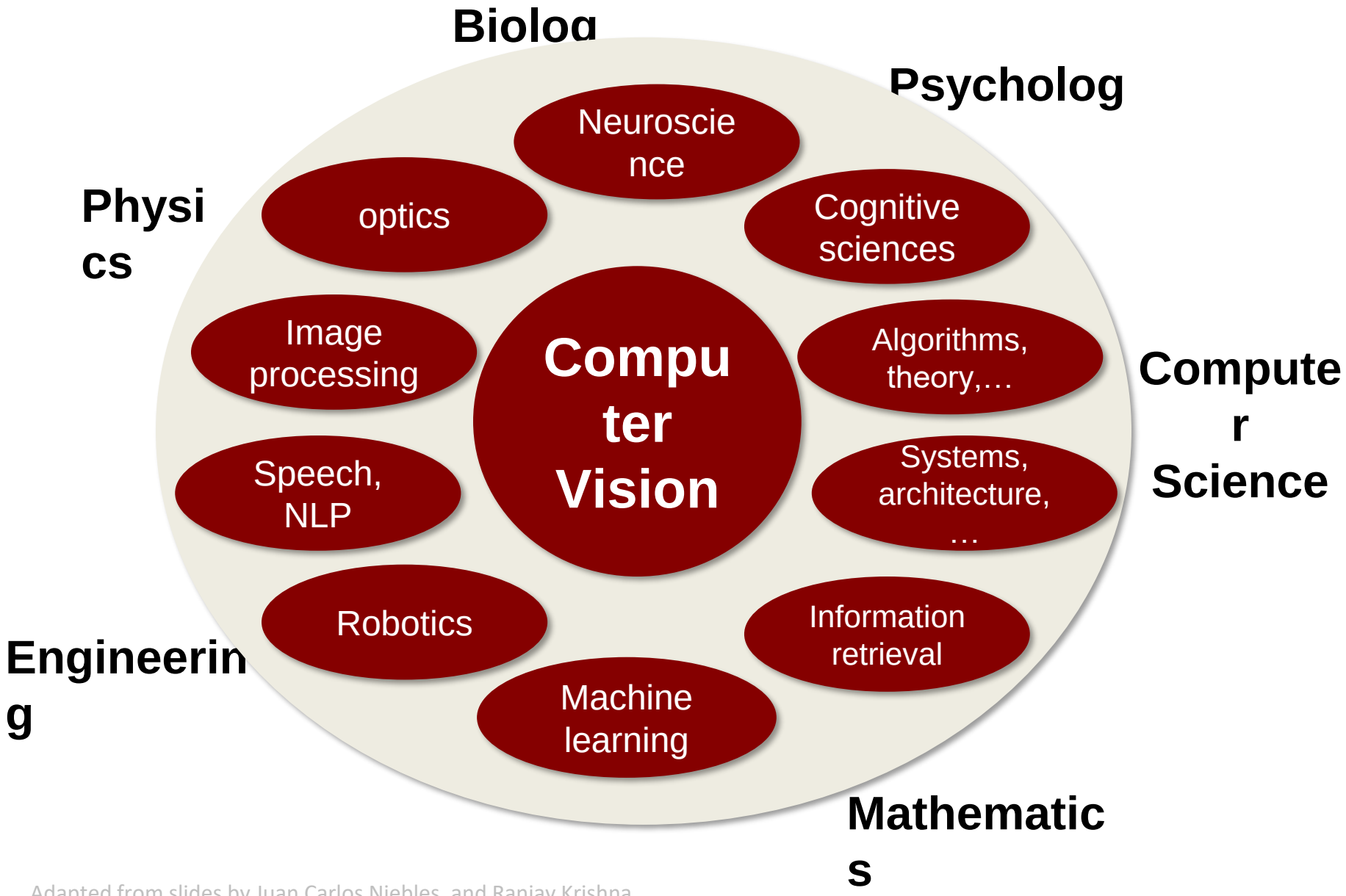

UNSATURATED OCEAN
BEFORE
SATURATED OCEAN
AFTER


Detector Camera

Introductory course for computer vision

- :
  - Enthusiastic undergrads
  - Want to get to know this exciting technology
  - Covers “DL for vision”, in a semi-applied way
- CENG 783: Deep Learning
- CENG 793: Advanced Deep Learning
- CENG 796: Deep Generative Models

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna



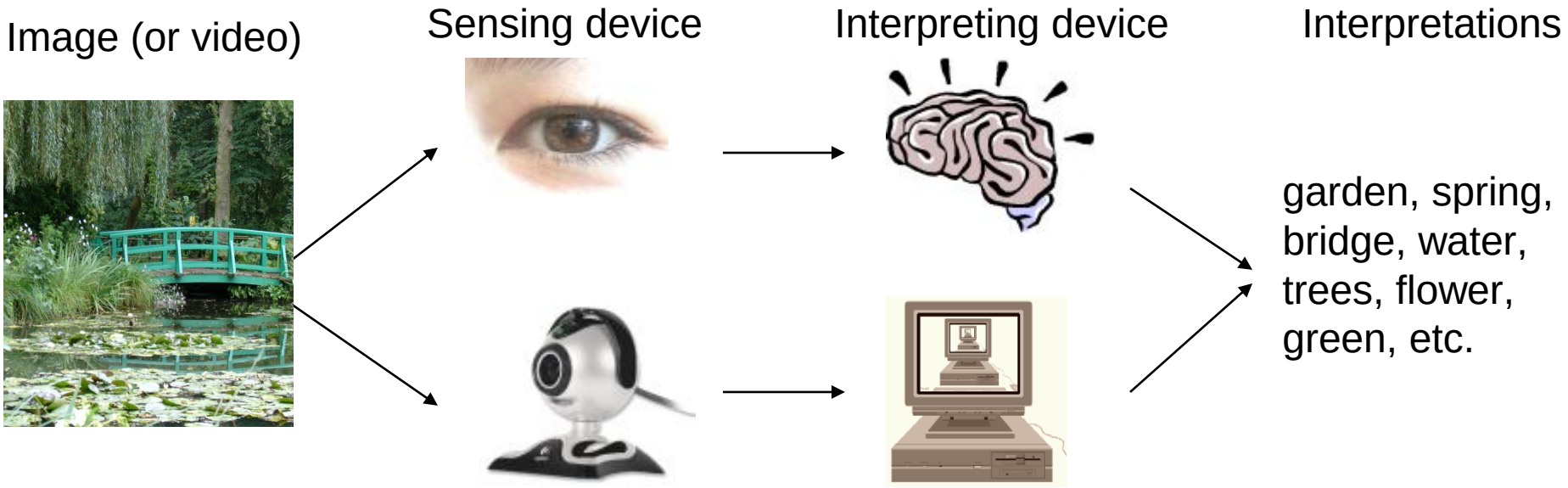
Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

This week's agenda

- Course overview
- Introduction to computer vision

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

What is (computer) vision?



Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

The goal of computer vision

- To bridge the gap between pixels and “meaning”

What we see

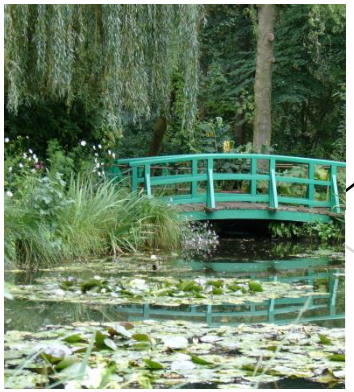
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

What is (computer) vision?

Image (or video)



Sensing device

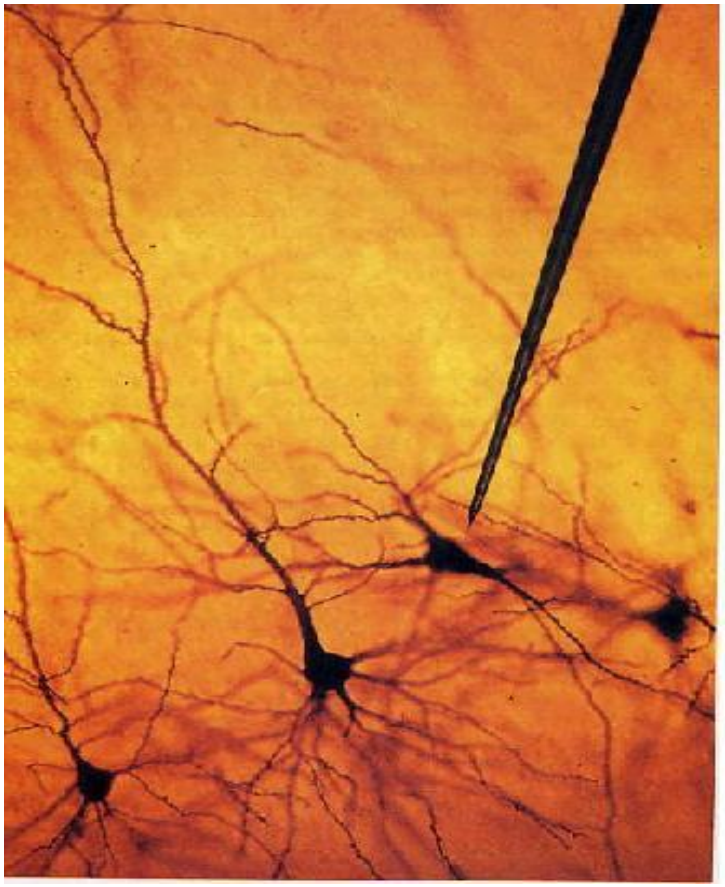
Interpreting device

Interpretations

garden, spring, bridge, water, trees, flower, green, etc.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

1981: Nobel Prize in medicine

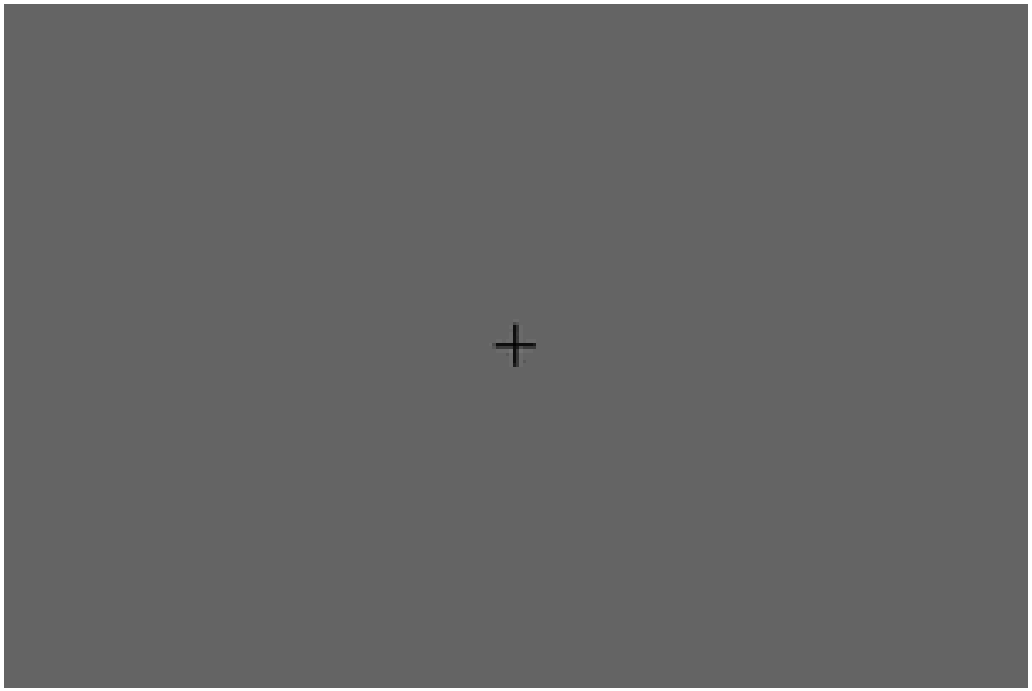


Hubel & Wiesel

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Human vision is superbly efficient

Can you detect a person?

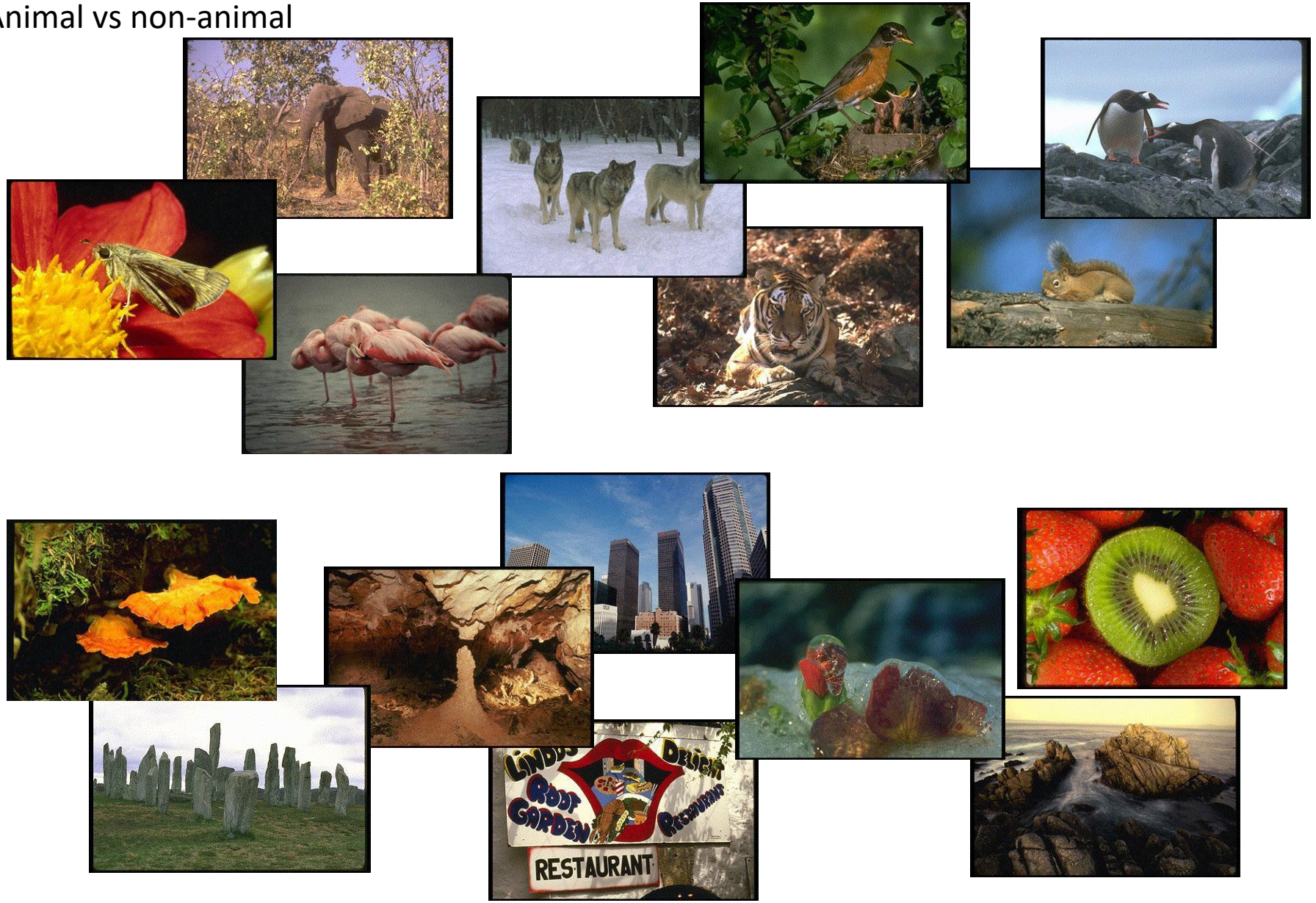


Potter, Biederman, etc. 1970s

100 ms per frame,
Never seen the picture, do not know the person

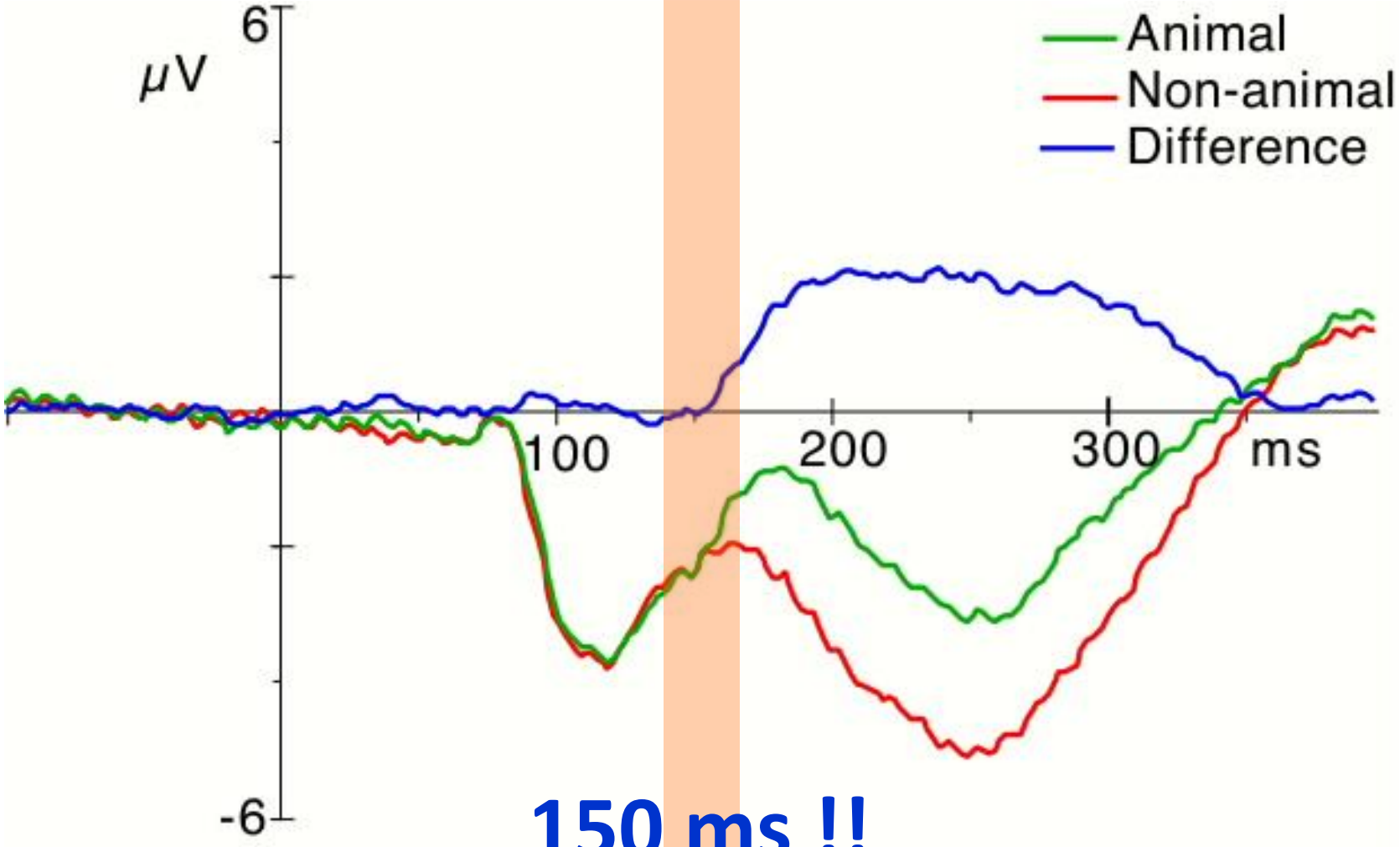
Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Animal vs non-animal



Thorpe, et al. *Nature,* 1996

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna



150 ms !!

Thorpe, et al. *Nature,* 1996

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Change blindness

Raise hand when you spot a difference

Rensink, O'regan, Simon, etc.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Change blindness

Rensink, O'regan, Simon, etc.

Change blindess

This is why spot-the-difference is a game ☺

Rensink, O'regan, Simon, etc.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

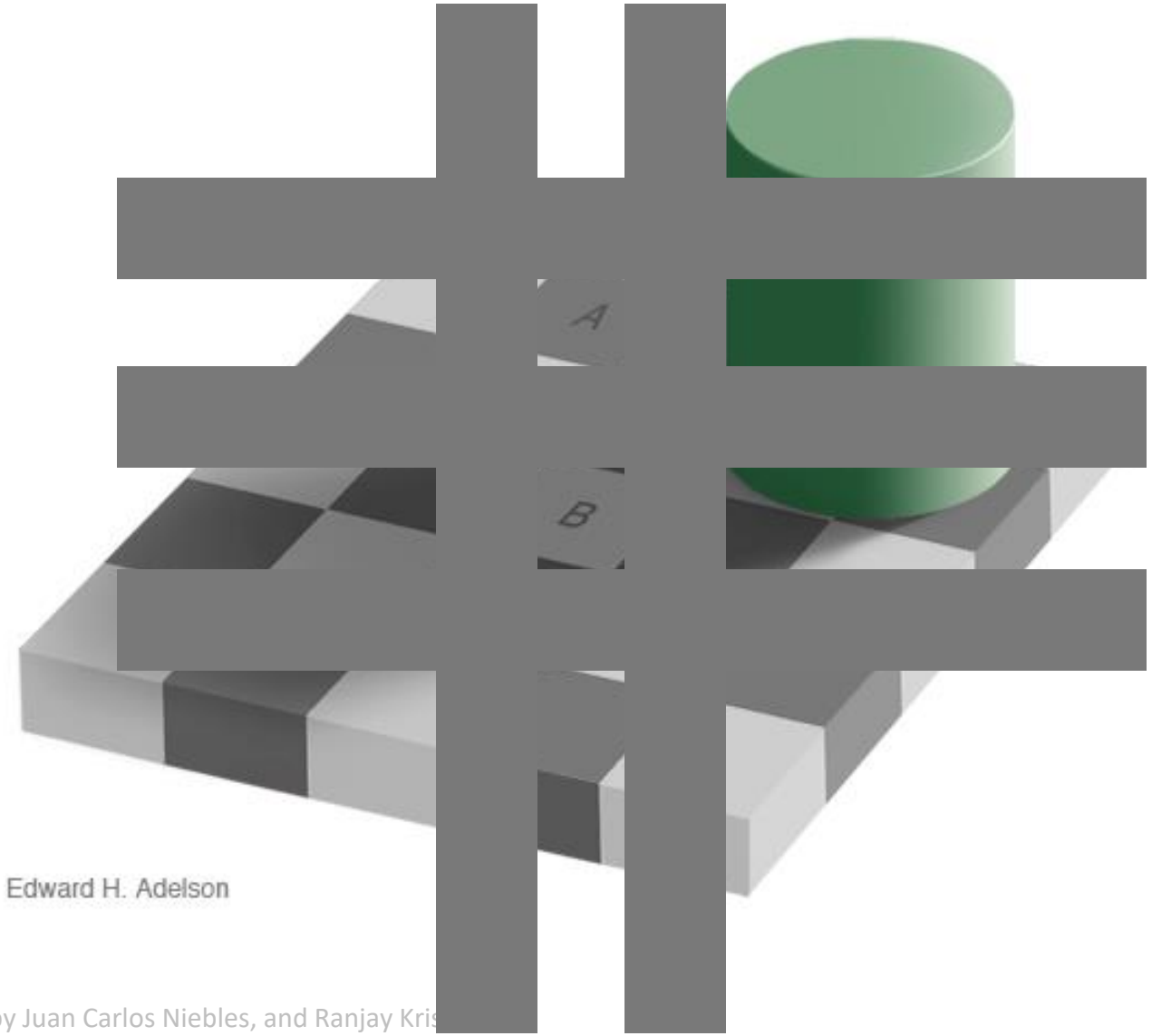
Segmentation

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Perception

Same face but the one behind looks like Al Gore due to feats of human face perception

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna



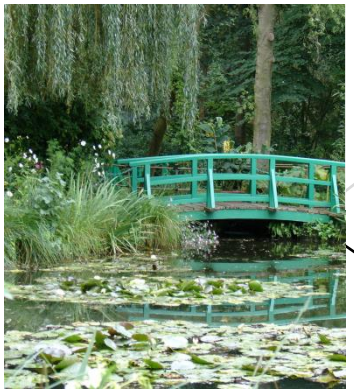
Edward H. Adelson

Human vision is GREAT, but its (mostly unknown) approach is not necessarily *the best* on for *computer* vision

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

What is (computer) vision?

Image (or video)



Sensing device

Interpreting device

Interpretations

garden, spring, bridge, water, trees, flower, green, etc.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

The goal of computer vision

- To bridge the gap between pixels and “meaning”

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

Course Logistics

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Contacting instructor and TAs

- Instructor: Gokberk Cinbis
  - Website: http://user.ceng.metu.edu.tr/~gcinbis
  - Primary Research Area: Data-efficient ML/DL, primarily on visual data
  - Office hours: please email to get appointment
- Teaching Assistant: İlker Işık

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Contacting instructor and TAs

- All announcements, Q&A will be made on ODTUCLASS
  - https://odtuclass.metu.edu.tr
- Please post your questions on ODTUCLASS
- Please email us if you encounter problems with forums

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Overall philosophy

- Breadth
  - Computer vision is a huge field, now tightly integrated with deep learning research
  - It can impact every aspect of life and society
  - It will drive the next information and AI revolution
  - Pixels are everywhere in our lives and cyber space
  - This is an introductory course, we will not cover all topics of CV
  - Lectures are mixture of details techniques and high level ideas
  - Speak our “language”

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Overall philosophy

- Depth
  - Master bread-and-butter techniques: eg. edges, corners, feature extraction, image classification
  - Programming assignments: be a good coder AND a good writer
  - Gather hands-on experience on using & developing basic deep learning approaches

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Lecture hours

- Lectures will start at **8:40** on both Tue & Fri.

- We'll effectively have 4 lectures per week (instead of 3), therefore, we skip some lecture days. This will help us to make faster progress in terms of theoretical discussions so that homeworks are not delayed in the semester, also makes it easy to make-up cancelled classes.

Office hours: after class / appointment by email.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Lecture recordings?

Probably not, for your own sake. Your primary resource for the classes are slides + lectures themselves.

Great lectures notes for a related course can be found here:
https://github.com/StanfordVL/CS131_notes

See syllabus for the textbook recommendations

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Zoom?

Again, probably not, for your own sake. Hybrid lectures, in my experience, greatly reduce the in-class discussion quality.

Grading policy

- Homeworks: 55% (Tentatively 3 HWs)
- Midterm: 20%
- Final: 25%

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Homework - late day policy

- Late day policy (important!)
  - Each student has 3 free calendar days to turn in late **in total (throughout the semester).** Use them wisely.
  - Example: if a student submits hw1 5 hours late, he/she will consume 1 full day, and will have 2 remaining late days.
  - Once late days are exhausted, 100% penalty will be applied to late submissions.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Homework policies

- Homeworks must be turned in on ODTUClass, unless otherwise specified.

Exception: if you encounter an issue when submitting the ODTUClass, please let us know. If deadline is about to pass, email TA your solution immediately! If the file is too large, email link to the dropbox/google drive/etc directory containing your solution. (We must see the solution date.)

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Collaboration policy

- Collaboration policy
  - Plagiarism will not be tolerated
  - Self-plagiarism will not be tolerated
  - METU Guidelines: http://oidb.metu.edu.tr/en/system/files/Academic%20Integrity%20Guide%20for%20Students.pdf
  - Ask ahead of time whenever you are in doubt.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Midterm/Final Exams

- Will contain written questions from the concept covered in class or any questions in the homeworks.
- Can require you to solve technical math problems.
- Midterm/Final maybe done in a lab environment.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

More details

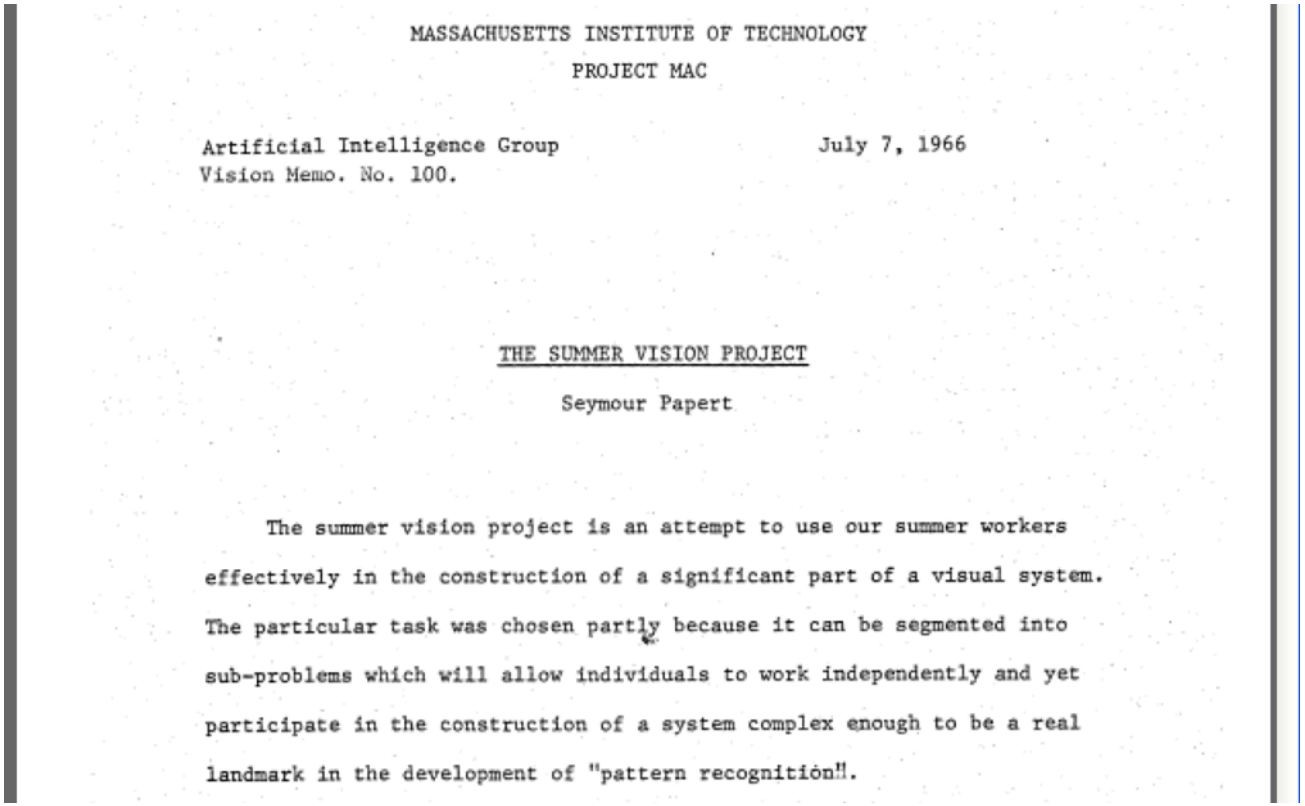
- Course webpage, syllabus:

http://user.ceng.metu.edu.tr/~gcinbis/courses/Fall23/CENG483/

Also, if interested in, about my research:
http://user.ceng.metu.edu.tr/~gcinbis

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Origins of computer vision: an MIT undergraduate summer project



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

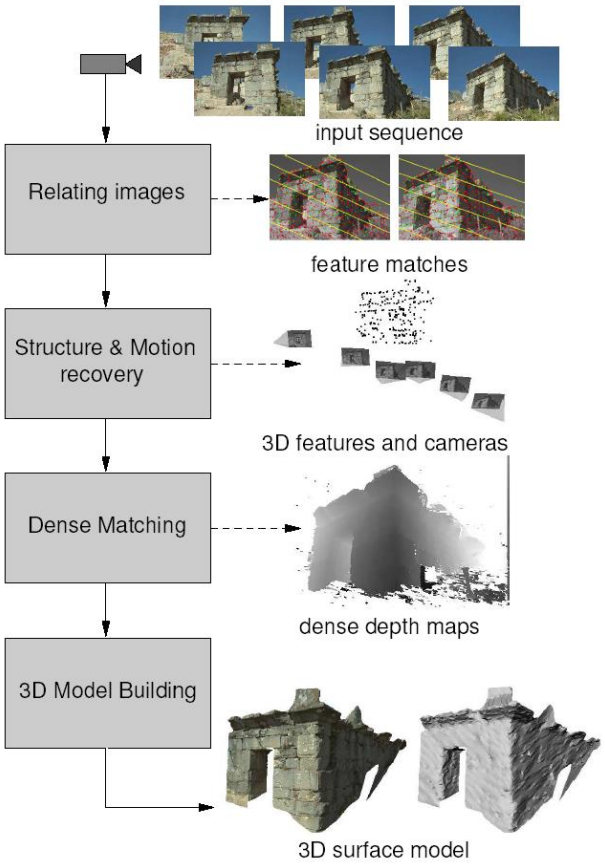
Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

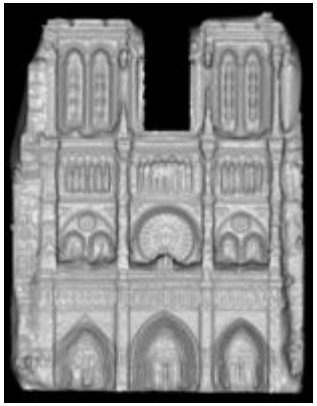
What kind of information can we extract from an image?

- (Geo)metric information
- Semantic information

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Vision as measurement device





Pollefeys et al.

Goesele et al.

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Vision as a source of semantic information

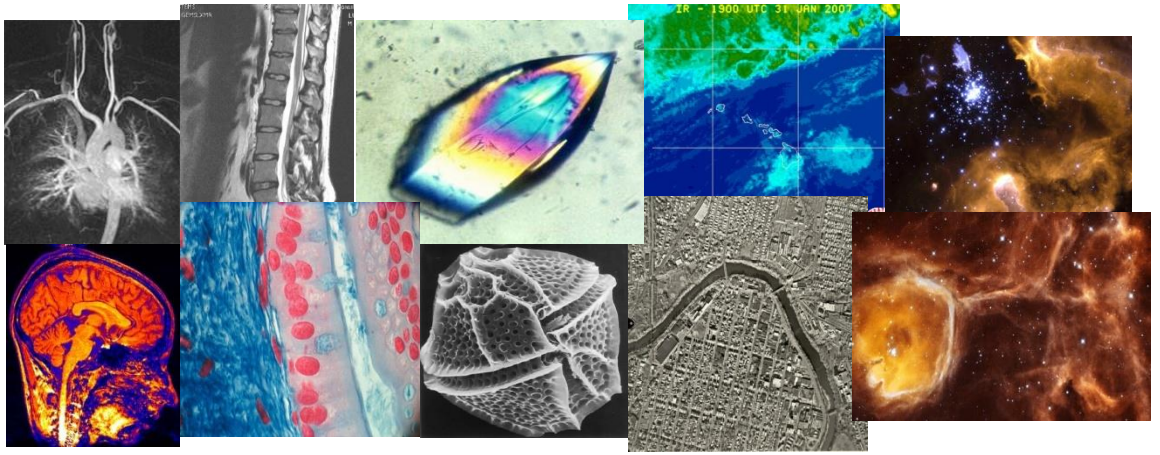
Slide credit: Kristen Grauman

Why study computer vision?

- Vision is useful: Images and video are everywhere!

Google Image Search Google Photos flickr webshots picsearch YouTube Broadcast Yourself

Surveillance and security



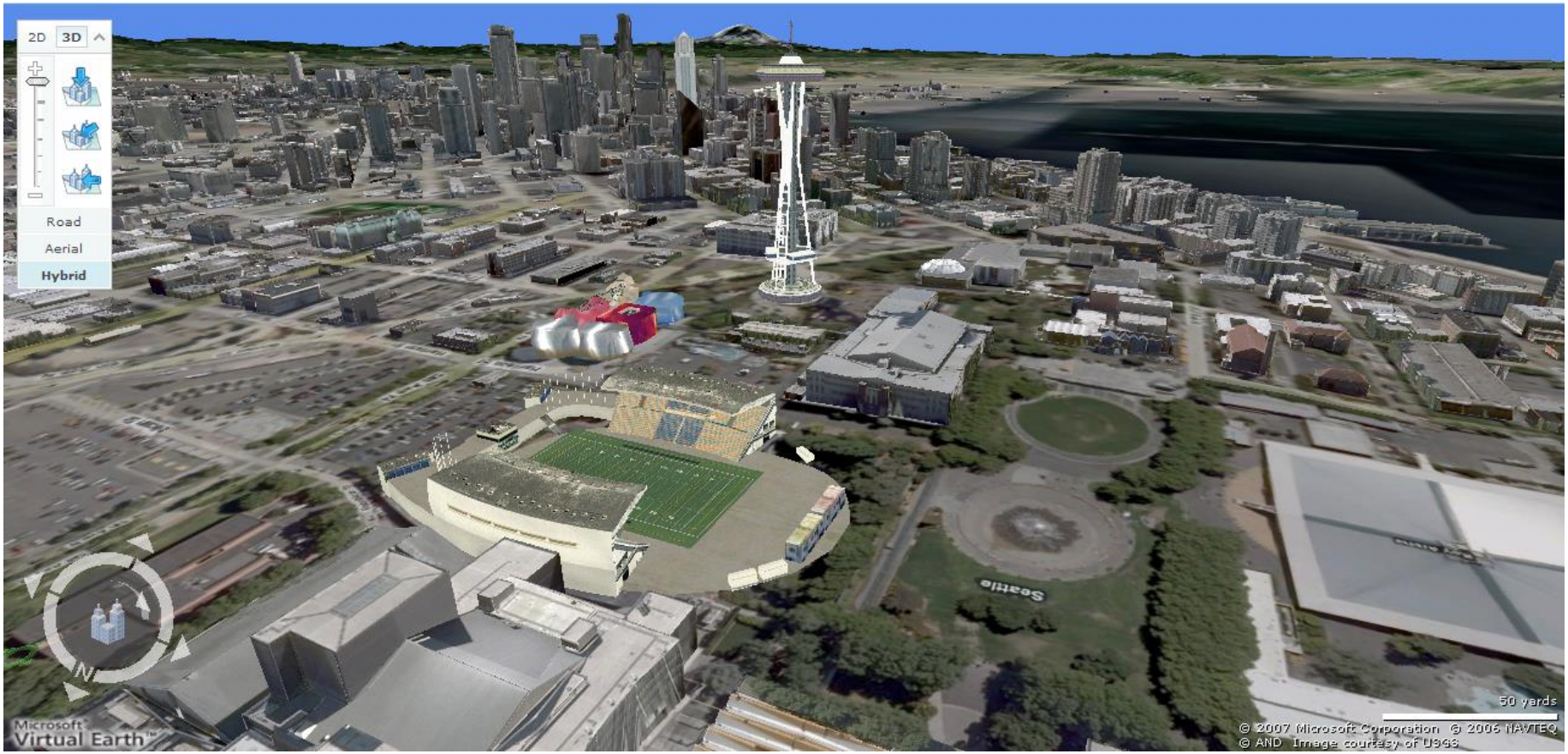
Medical and scientific images

Special effects: shape and motion capture

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

3D urban modeling



Bing maps, Google Streetview

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

3D urban modeling: Microsoft Photosynth

http://photosynth.net (defunct)

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

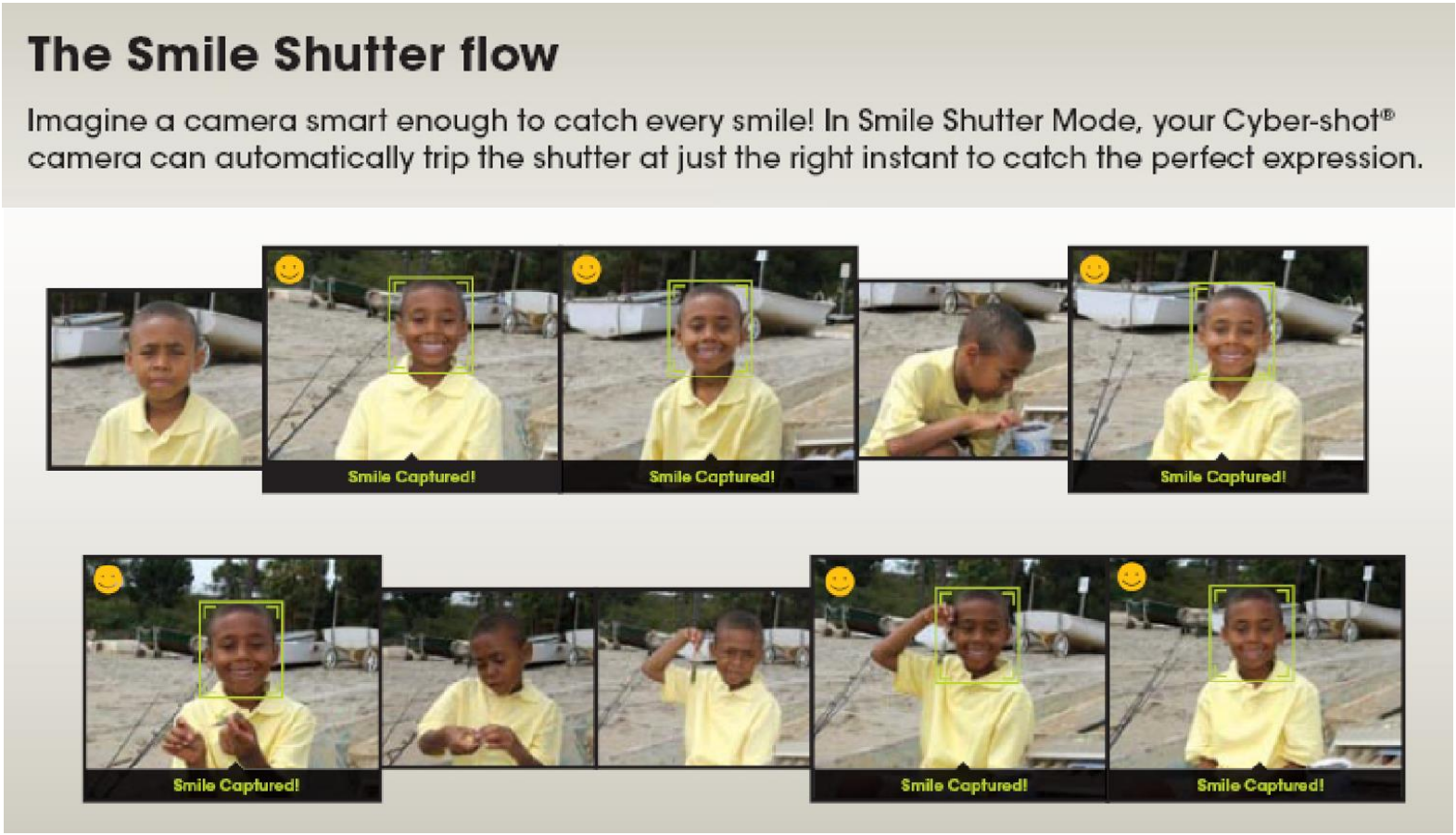
Face detection

- Many digital cameras now detect faces
  - Canon, Sony, Fuji, …

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Smile detection



Sony Cyber-shot® T70 Digital Still Camera

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Face recognition: Apple, Facebook, etc.

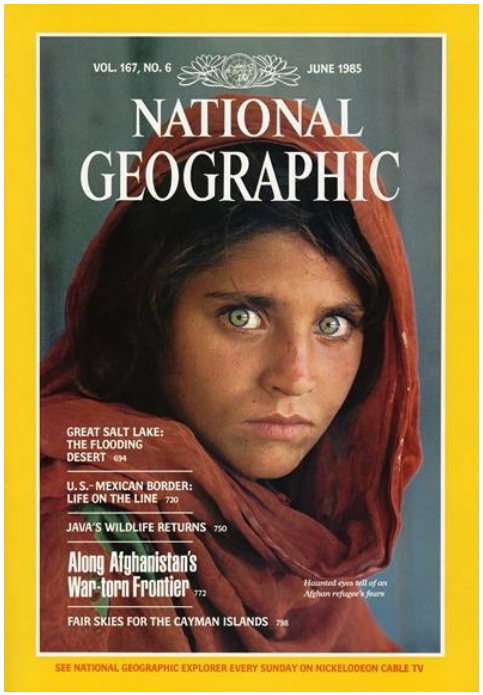
http://www.apple.com/ilife/iphoto/

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

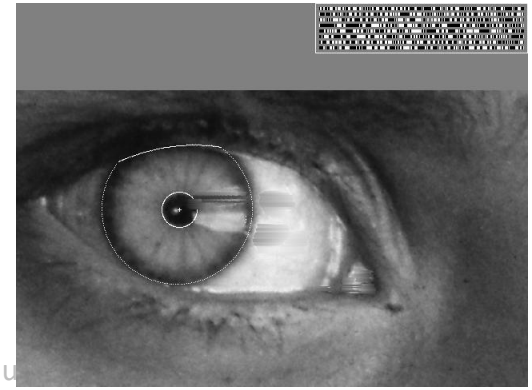
Portrait Focus

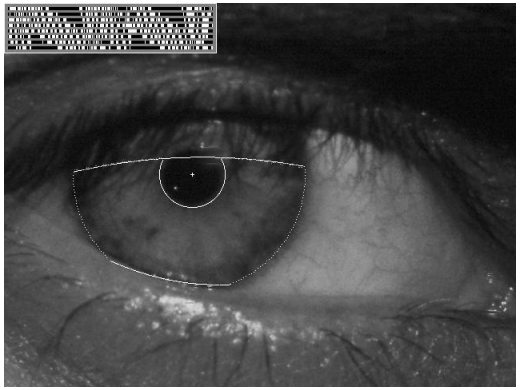
https://www.engadget.com/2016/10/24/apple-releases-ios-10-1-adds-portrait-mode-to-the-iphone-7-plus/

Biometrics



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

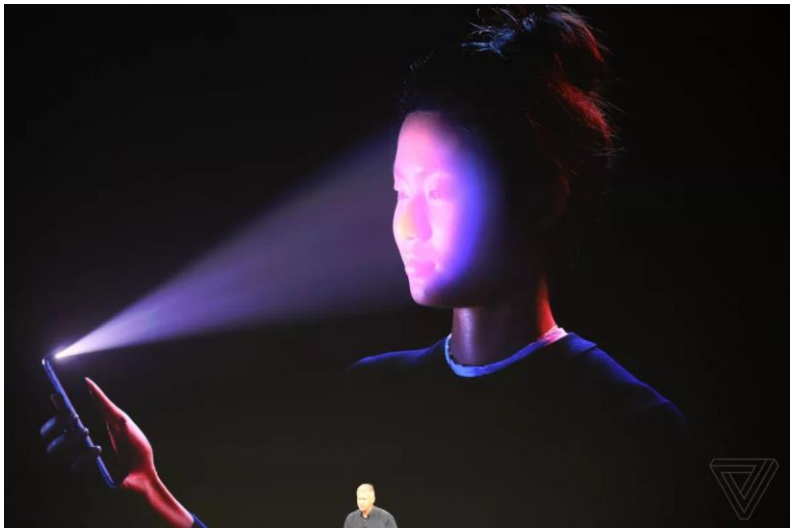




Source: S. Seitz

Biometrics

Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely

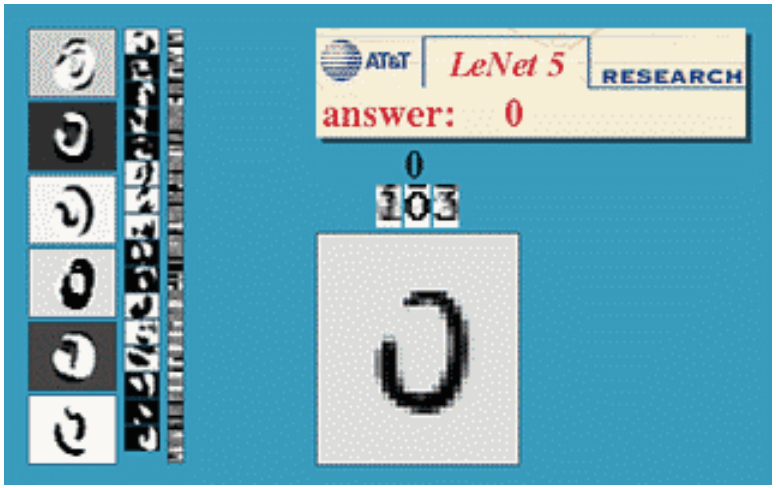
iphone X just introduced face recognition

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs

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

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Source: S. Seitz

Toys and Robots

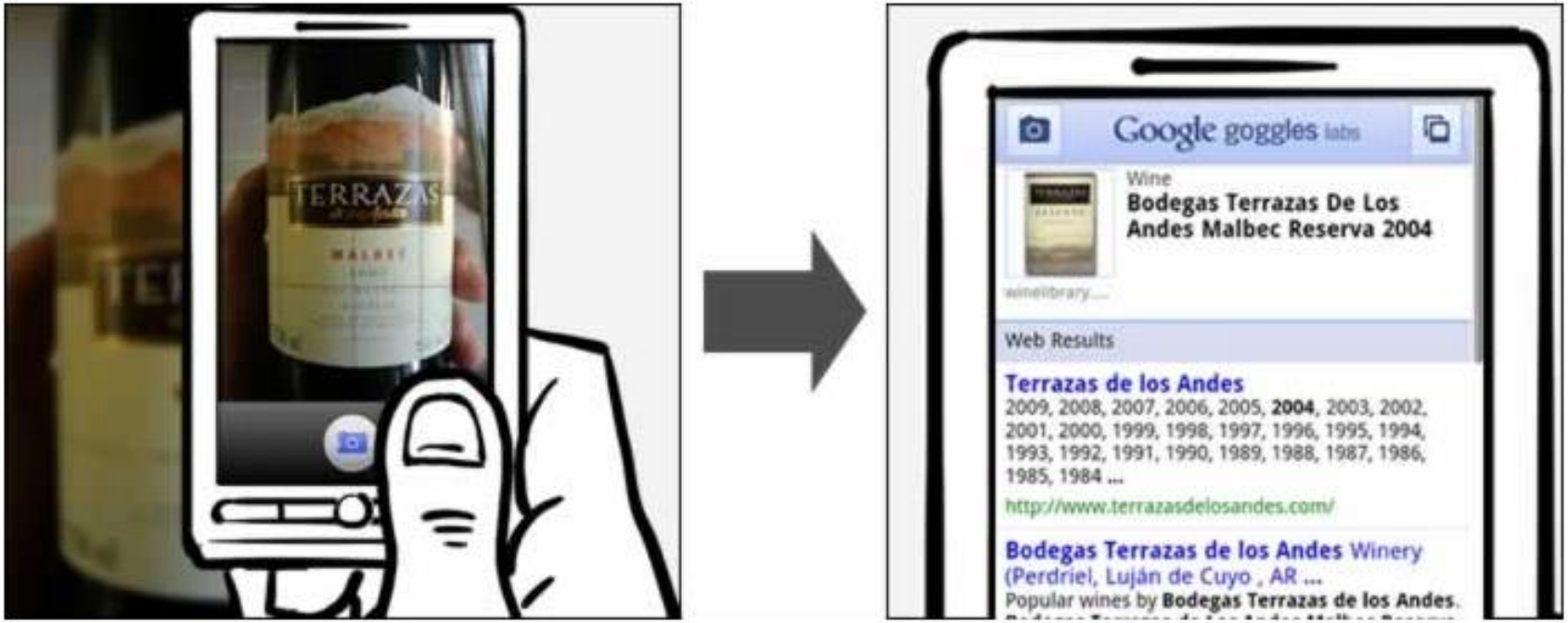
Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Mobile visual search: Google Goggles

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

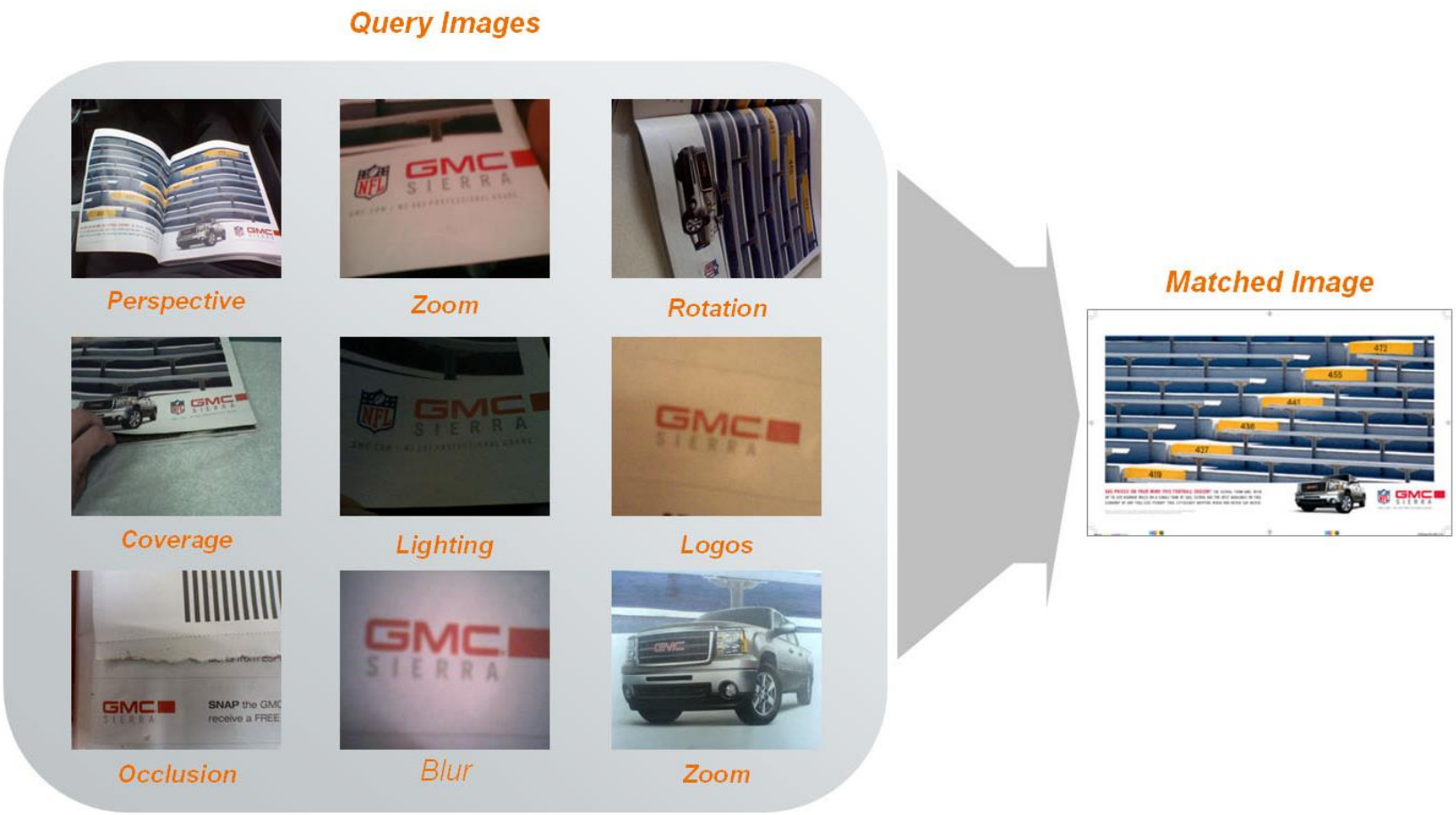
Google goggles labs



Mobile visual search: iPhone Apps

snaptell

kooaba



Adapt

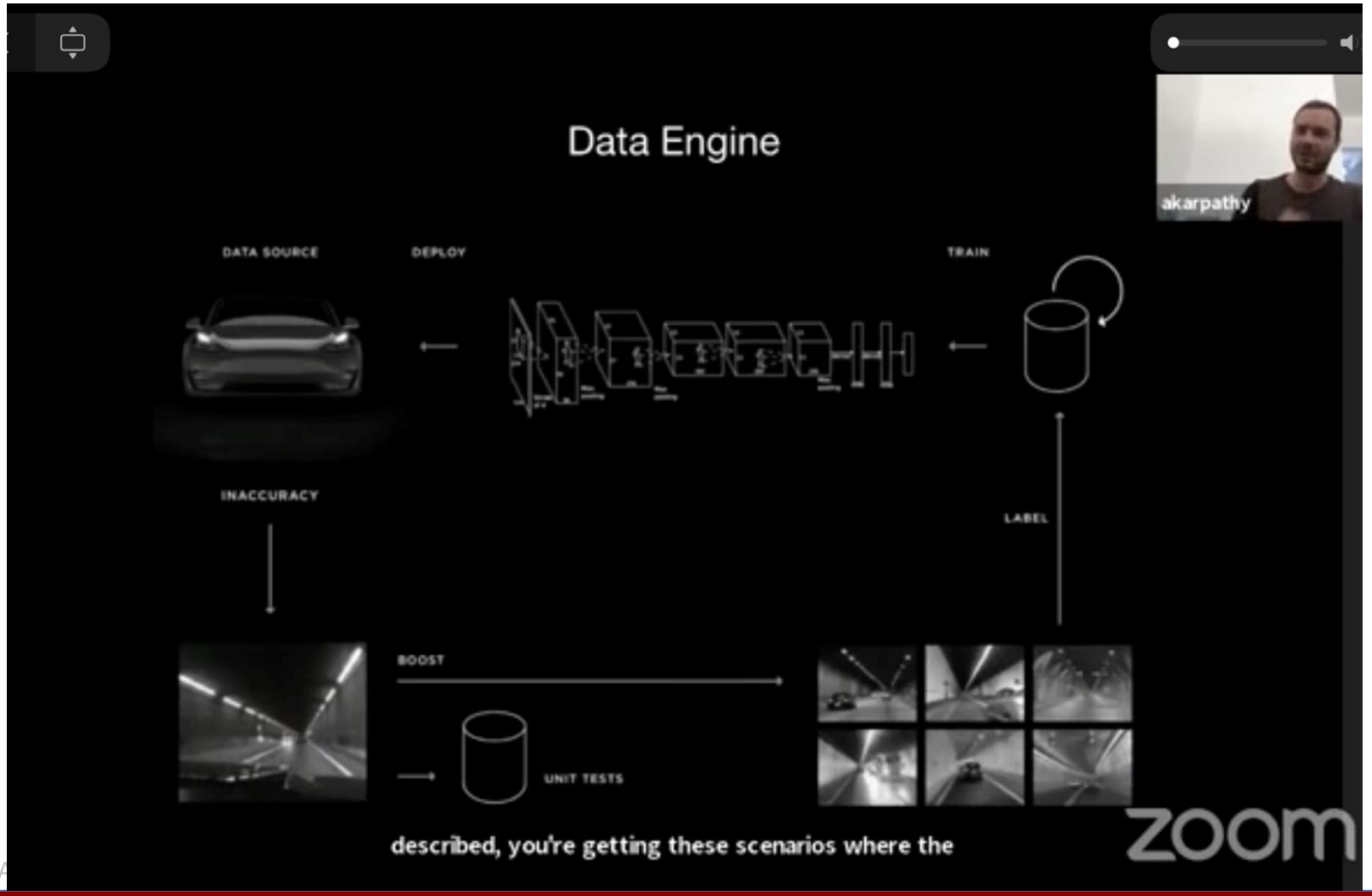
Automotive safety

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - "In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians."

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Source: A. Shashua, S. Seitz

Towards fully autonomous driving



Vision in supermarkets

[LaneHawk by EvolutionRobotics]

"A smart camera is flush-mounted in the checout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it… "

Source: S. Seitz

Amazon Go

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Vision-based interaction (and games)

Microsoft's Kinect

Sony EyeToy

Assistive technologies

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Augmented Reality

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Vision for robotics, space exploration

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

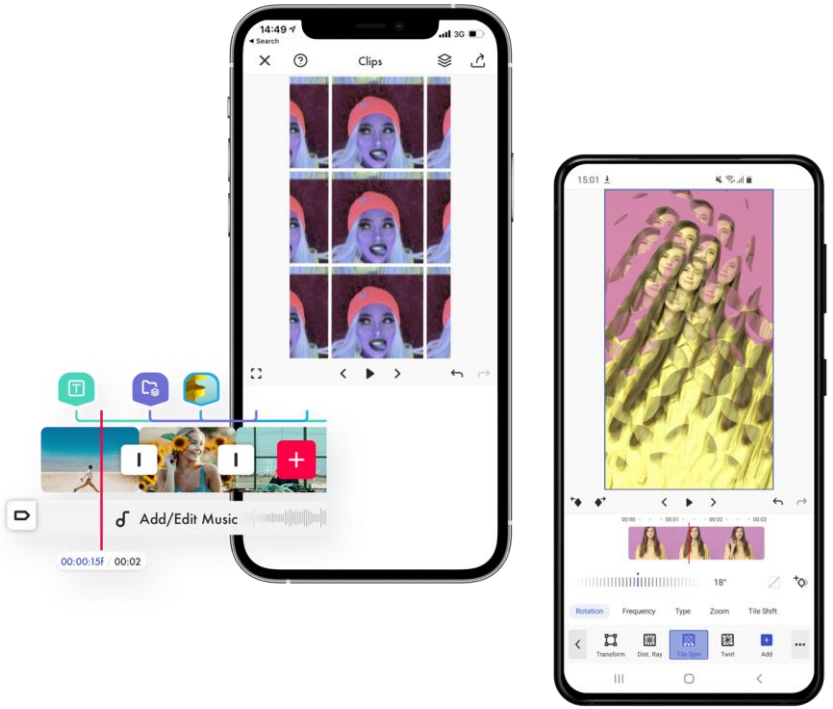
Vision systems (JPL) used for several tasks

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

Source: S. Seitz

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Creative content editing



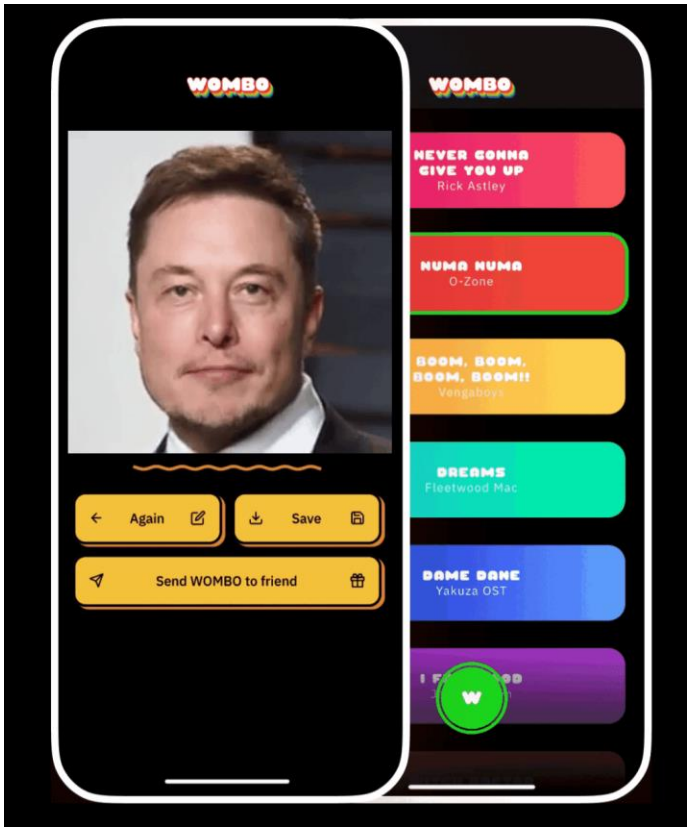
NVidia

https://www.pixerylabs.com

Content creation

https://ai.facebook.com/blog/generative-ai-text-to-video/
https://makeavideo.studio/
September 29, 2022

And *just* for fun :)

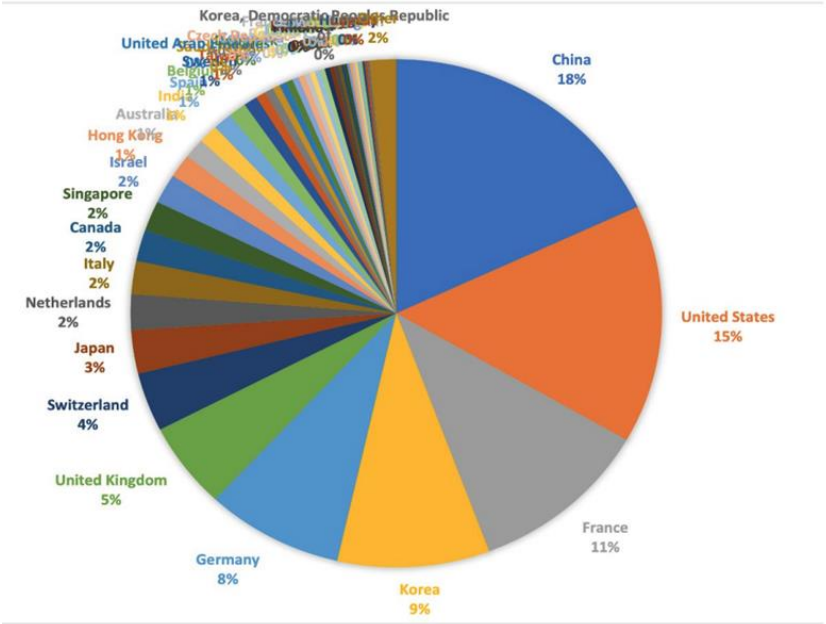


https://www.wombo.ai

Adapted from slides by Juan Carlos Niebles, and Ranjay Krishna

Computer Vision as a Research Field

Attendance



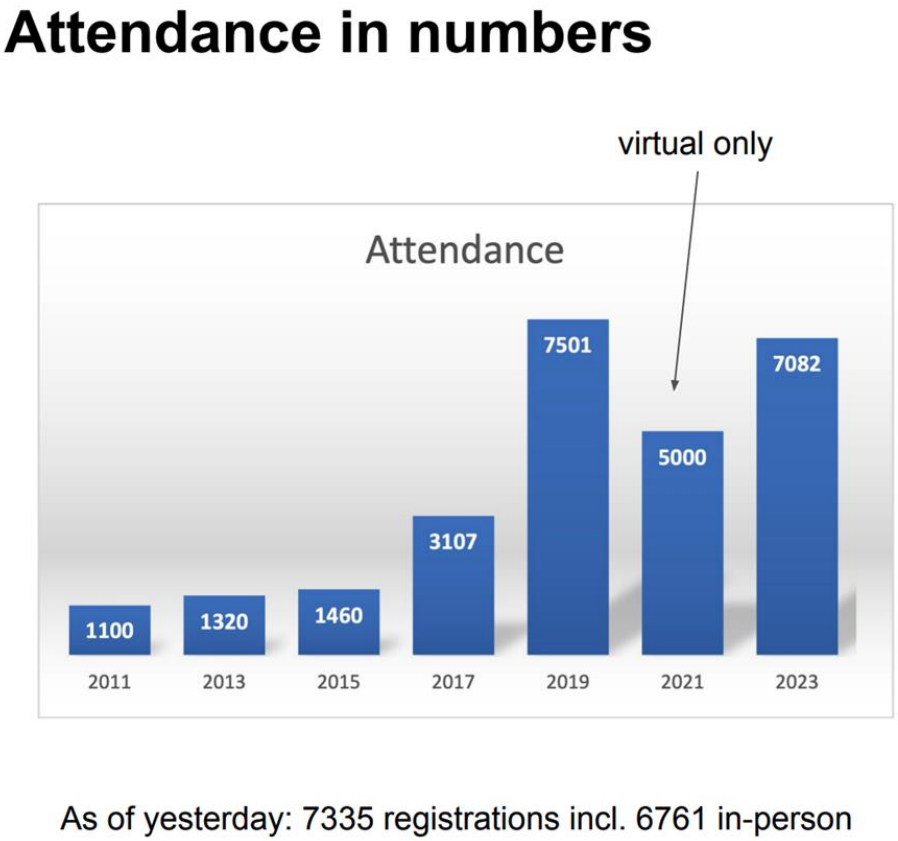
Attendees from 84 countries/regions

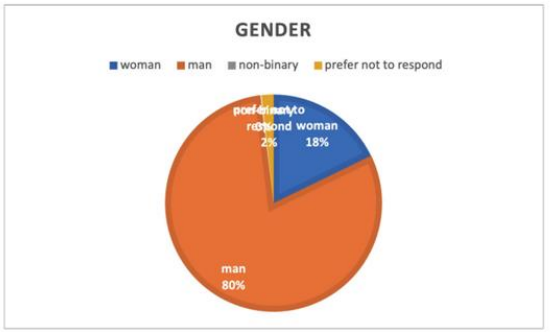
China 1233
United States 1034
France 749
Korea 630
Germany 569
UK 374
Switzerland 255
Japan 188
Netherlands 155
Italy 143
Canada 136
Singapore 130
Israel 128
Hong Kong 102
Australia 87
India 81
Spain 79
Belgium 75
Sweden 57
Taiwan 41
Austria 37
Saudi Arabia 30
United Arab Emirates

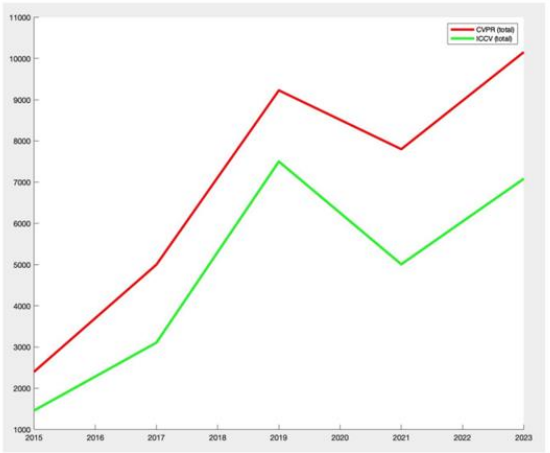
Denmark 29
Colombia 26
Czech Republic 26
Poland 23
Greece 22
Turkey 21
Mexico 21
Finland 21
Romania 19
Croatia 17
Armenia 12
Slovenia 12
Ethiopia 12
Luxembourg 11
Vietnam 11
Norway 10
Brazil 10
Hungary 10
Cyprus 10
Korea Dem 10
Other 37 countries < 10

https://iccv2023.thecvf.com/

Computer Vision as a Research Field







https://iccv2023.thecvf.com/