THE UNIVERSITY OF BRITISH COLUMBIA

CPSC 425: Computer Vision

Image Credit: Devi Parikh

Lecture 1: Introduction and Course Logistics

Course **logistics**

Times: Mon, Wed 3:30-5:00pm **Locations:** Friedman (FRDM), Room 153

Instructor: Leonid Sigal

E-mail: lsigal@cs.ubc.ca

Office: ICICS 119

Course **logistics**

Times: Mon, Wed 3:30-5:00pm

Locations: Friedman (FRDM), Room 153

Instructor: Leonid Sigal

Leon

E-mail: lsigal@cs.ubc.ca

Office: ICICS 119

About **me** ...

About **me** …

Software Engineer
1999 - 2001

COGNEX

About **me …**

Software Engineer
1999 - 2001

COGNEX

About **me** ...

Software Engineer
1999 - 2001

COGNEX

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

Disney Research

Postdoctoral Researcher
2007 - 2009

University of Toronto

PhD, MSc
2001 - 2008

Brown

Boston University

Software Engineer
1999 - 2001

Cognex

About **me ...**

Associate Professor
2017 -

THE UNIVERSITY OF BRITISH COLUMBIA

Senior Research Scientist
2009 - 2017

Disney Research

Postdoctoral Researcher
2007 - 2009

UNIVERSITY OF TORONTO

PhD, MSc
2001 - 2008

BROWN

BOSTON UNIVERSITY

Software Engineer
1999 - 2001

COGNEX

About **me …**

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

Associate Professor
2017 -

UBC THE UNIVERSITY OF BRITISH COLUMBIA

Senior Research Scientist
2009 - 2017

Disney Research

Postdoctoral Researcher
2007 - 2009

UNIVERSITY OF TORONTO

PhD, MSc
2001 - 2008

BROWN

BOSTON UNIVERSITY

Software Engineer
1999 - 2001

COGNEX

Course **logistics**

Times: Mon, Wed 3:30-5:00pm

Locations: Friedman (FRDM), Room 153

Instructor: Leonid Sigal

E-mail: lsigal@cs.ubc.ca

Office: ICICS 119

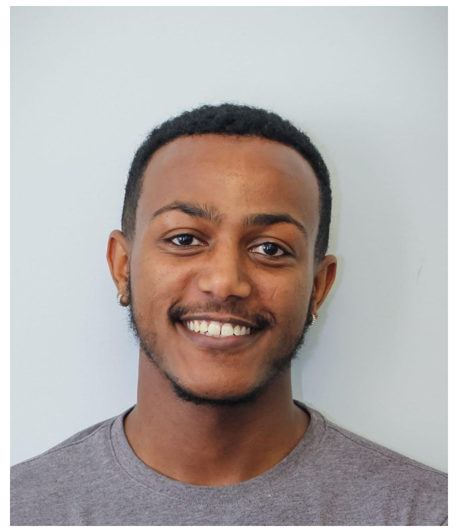
Course **logistics**

Times: Mon, Wed 3:30-5:00pm **Locations:** Friedman (FRDM), Room 153

Instructor: Leonid Sigal

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

TAs: Bereket Guta



bguta@cs.ubc.ca

Gaurav Bhatt

gbhatt@cs.ubc.ca

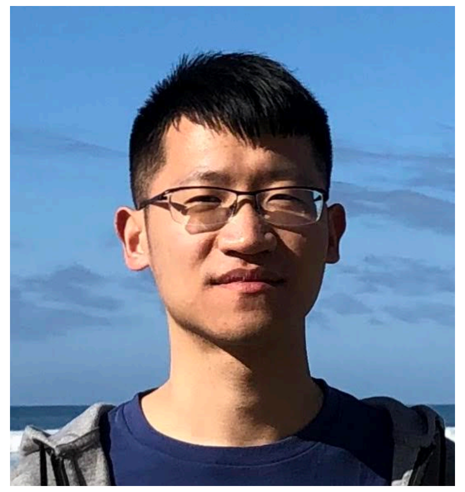
Aditya Chinchure

aditya10@cs.ubc.ca

Rayat Hossain

rayat137@cs.ubc.ca

Bicheng Xu



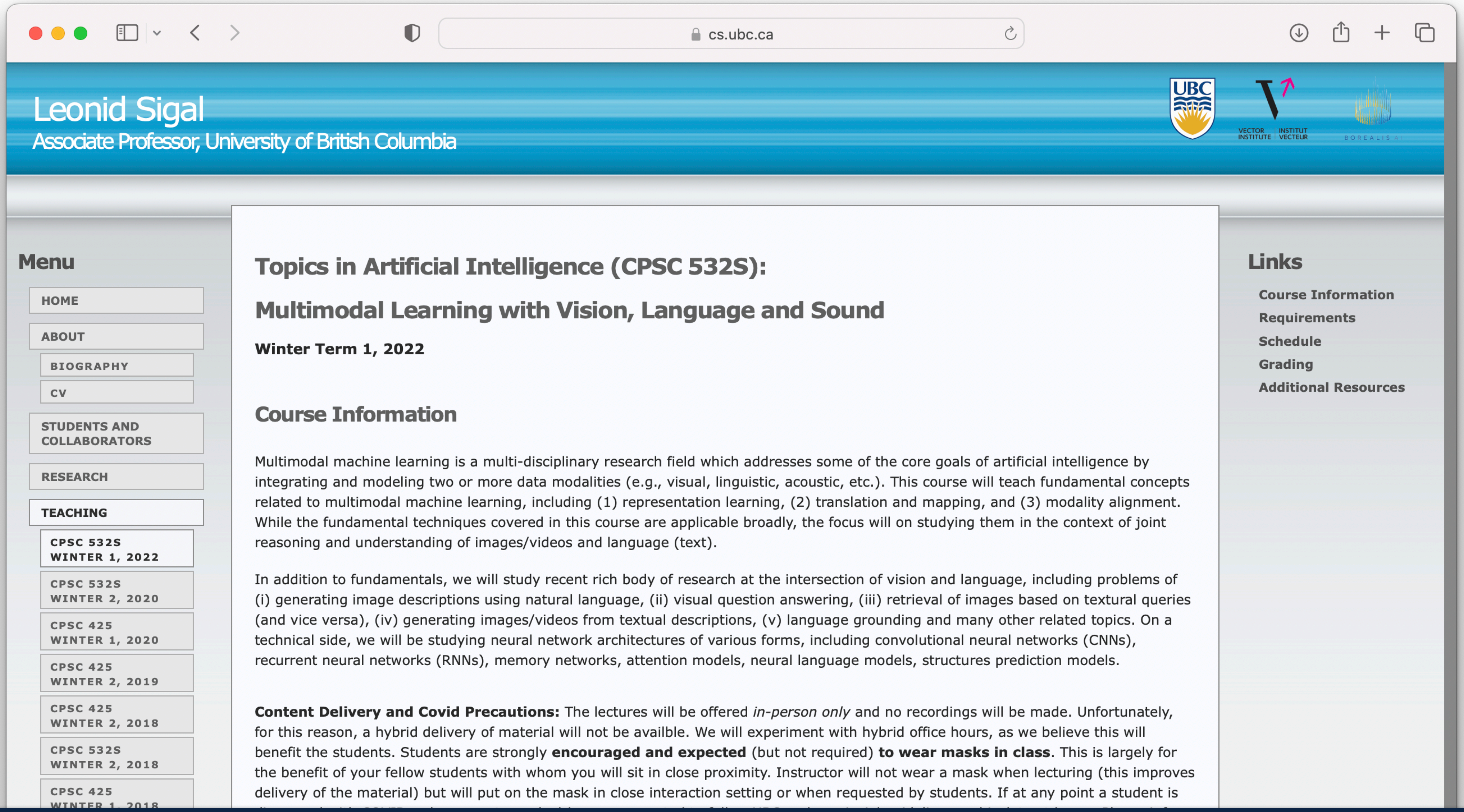
bichengx@cs.ubc.ca

Registration

- Capacity of room (160)
- Registered students (150)
- Wait List (21)

There is high likelihood everything will work out

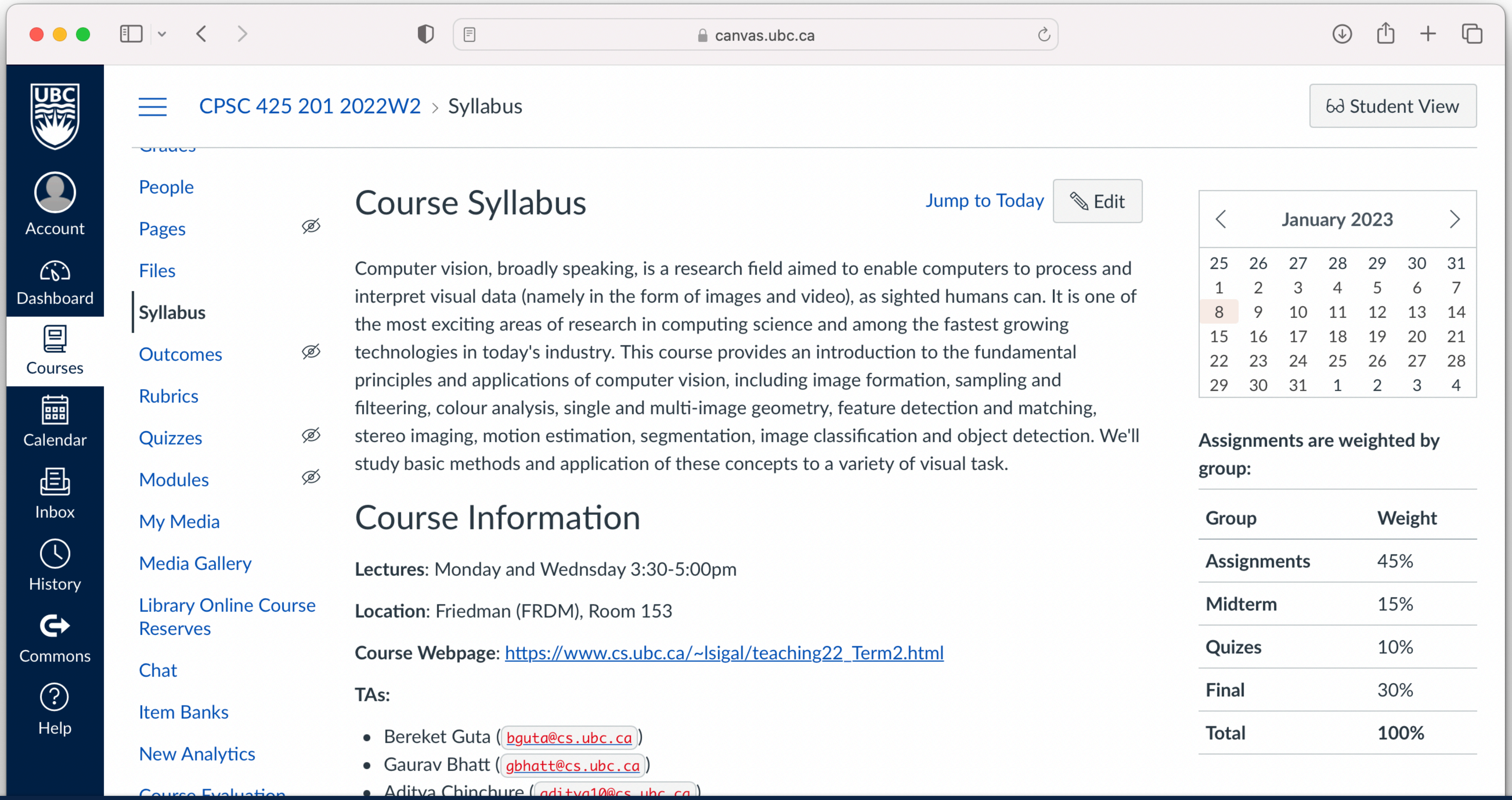
Course **Webpage**



- Schedule, Assignments
- Lecture Slides and Notes
- Readings
- Course Information (public)

https://www.cs.ubc.ca/~lsigal/teaching22_Term2.html

Canvas



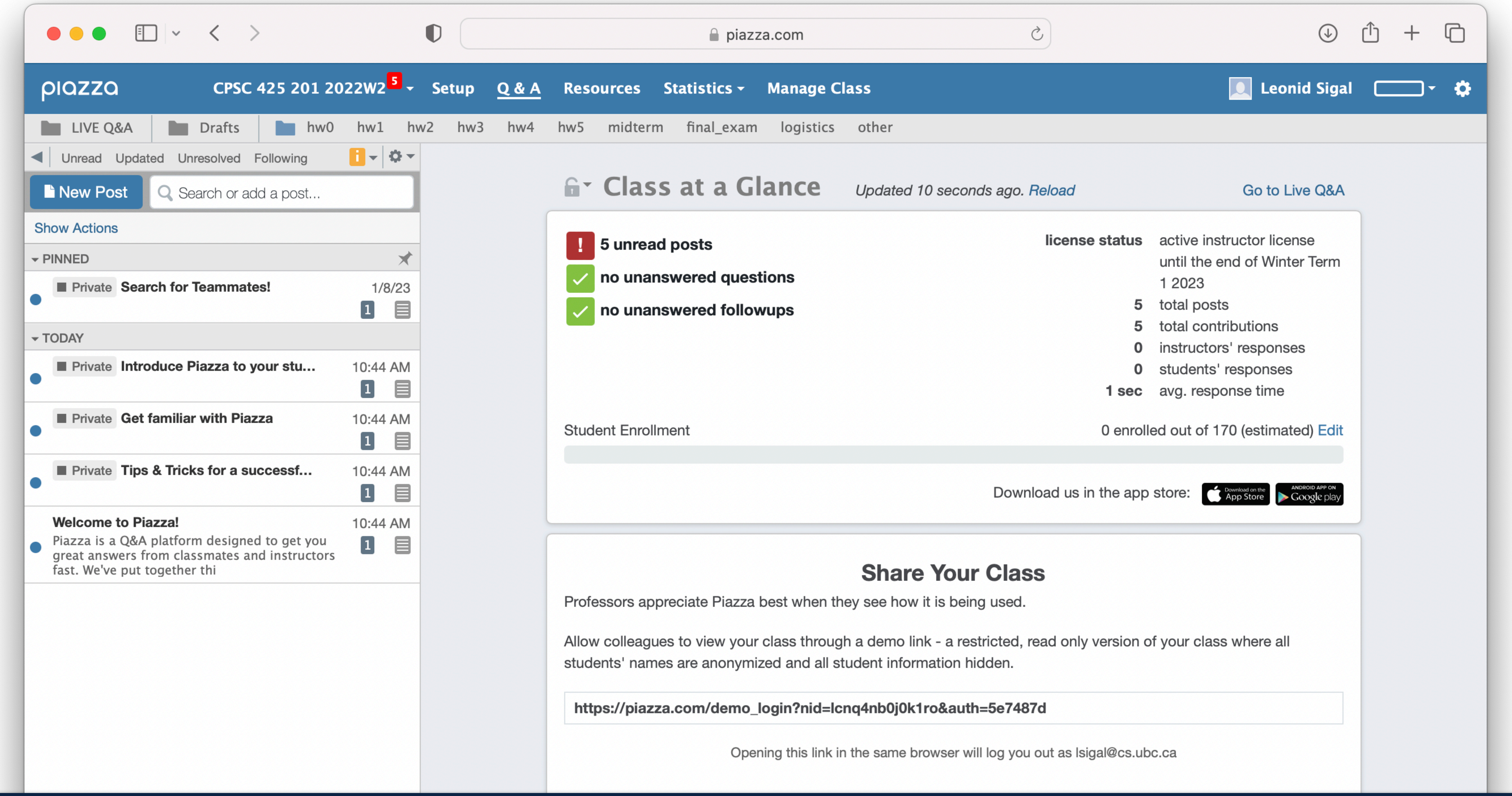
- Assignment hand-in
- Course Information (private)
- Piazza link

https://canvas.ubc.ca/courses/106387

Piazza

Discussion: piazza.com/ubc.ca/winterterm22023/cpsc4252012022w2/home

48 students were enrolled as of AM



- Discussions and Q+A
- Confused? Likely someone else has the same question as you!
- Lectures, Technical Issues, Assignments …

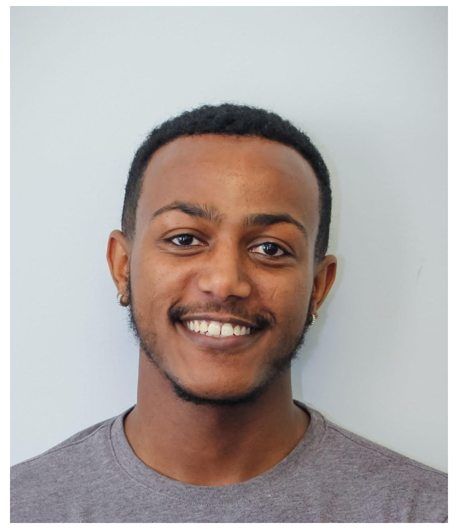
Sign up code in e-mail

Office **Hours**

Instructor: Leonid Sigal

Friday 12-1pm, Hybrid

TAs: Bereket Guta



Monday 2-3pm (in-person)

Gaurav Bhatt

TBD

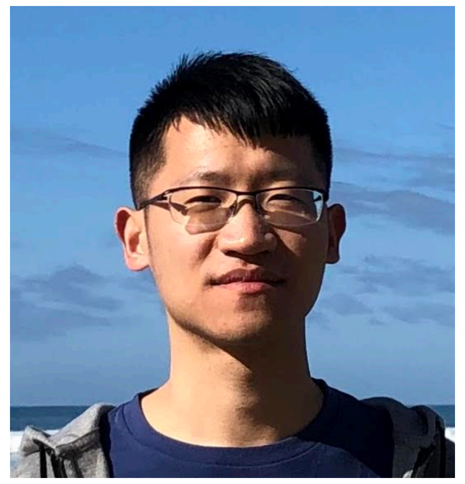
Aditya Chinchure

Tuesday noon-1pm

Rayat Hossain

TBD

Bicheng Xu



Friday 4-5pm
(in-person)

See Piazza for Links and Locations (mix of in-person and Zoom)

How important is **Vision**?

How important is **Vision**?

To answer this questions, we need to go back to about

…. **543 million years, B.C.**

How important is **Vision**?

To answer this questions, we need to go back to about

…. **543 million years, B.C.**

Vision is really fundamental to life and evolution

What is **Computer Vision**?

Image Credit: https://www.deviantart.com/infinitecreations/art/BioMech-Eye-168367549

What is **Computer Vision**?

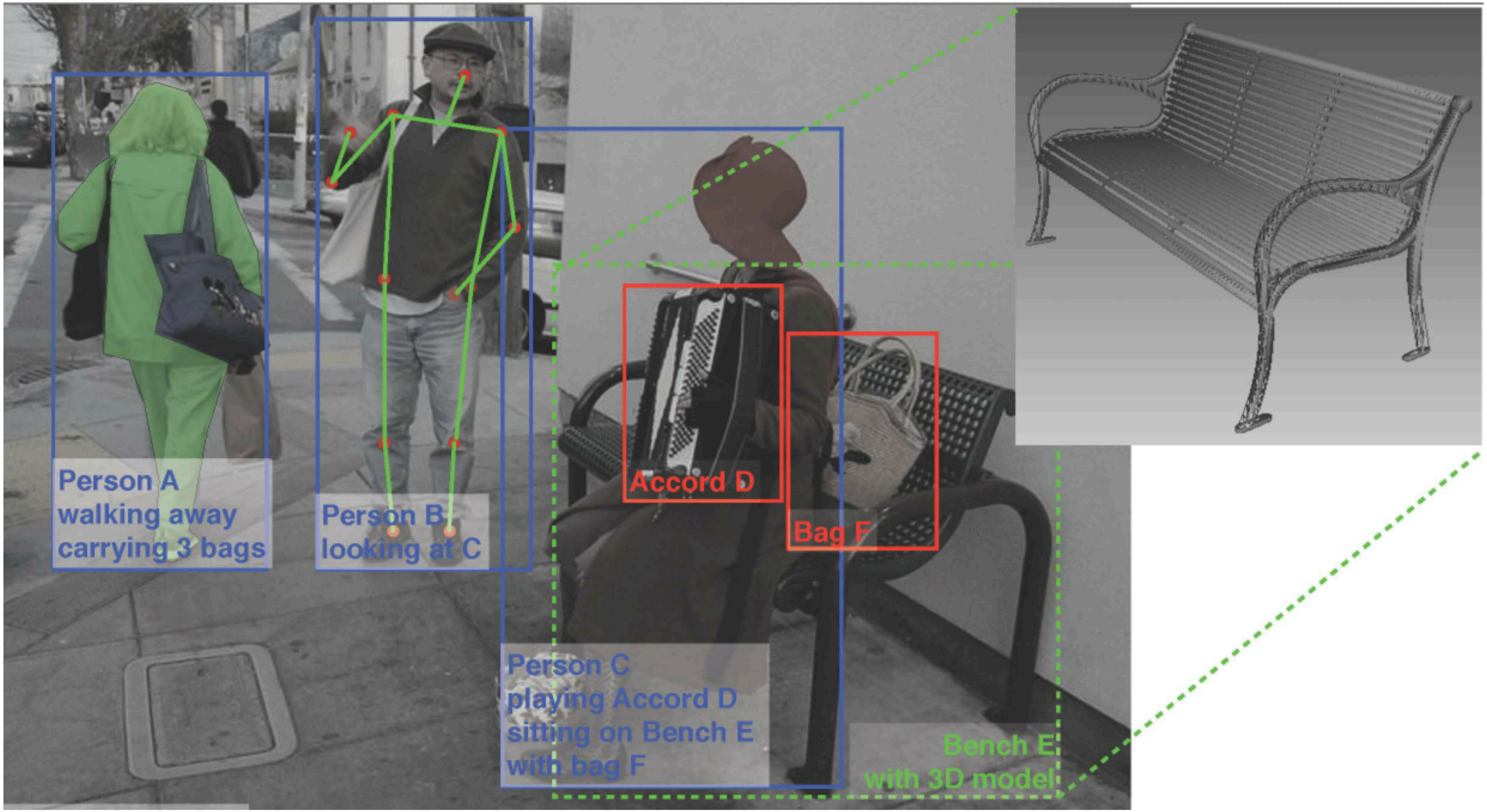
Compute vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.

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**?



Slide Credit: Jitendra Malik (UC Berkeley)

What we would like **computer to infer**?

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.

Sensing Device

Interpreting Device

Image (or video)

Interpretation

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.

Image (or video)

Sensing Device

Interpreting Device

Image Credit: https://www.flickr.com/photos/flamephoenix1991/8376271918

Interpretation

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.

Image (or video)

Sensing Device

Interpreting Device

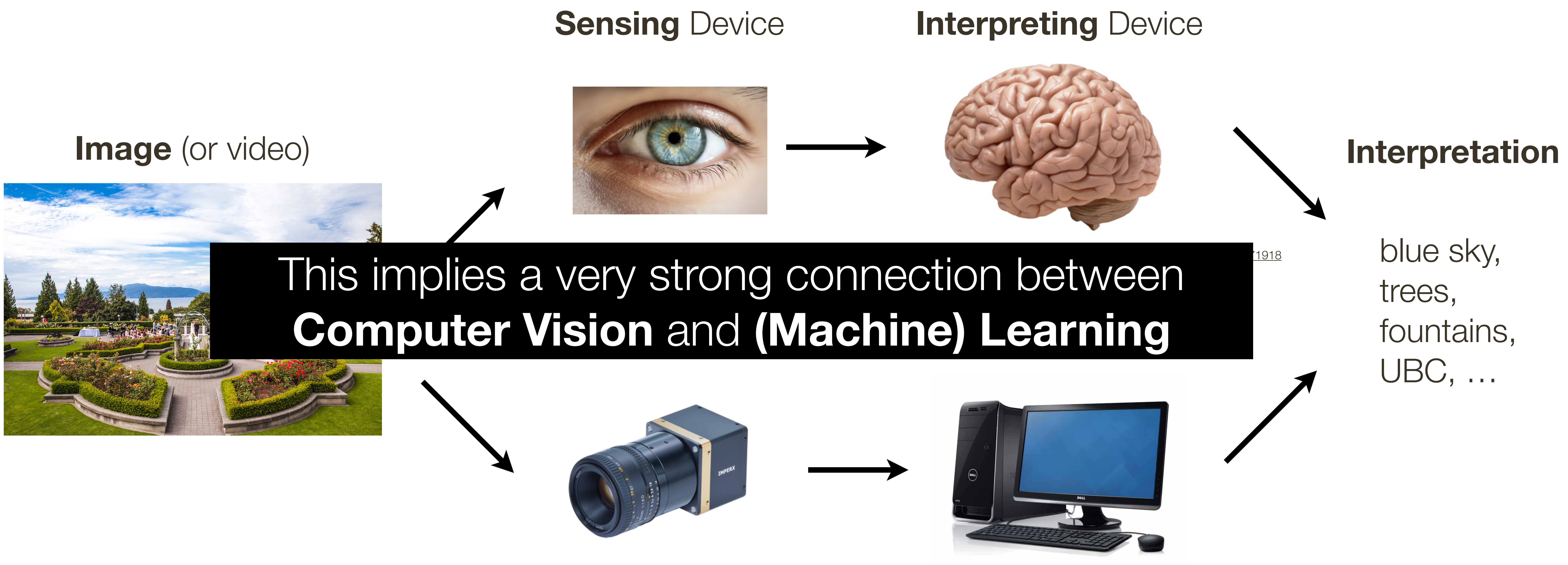
Image Credit: https://www.flickr.com/photos/flamephoenix1991/8376271918

Interpretation

blue sky,
trees,
fountains,
UBC, …

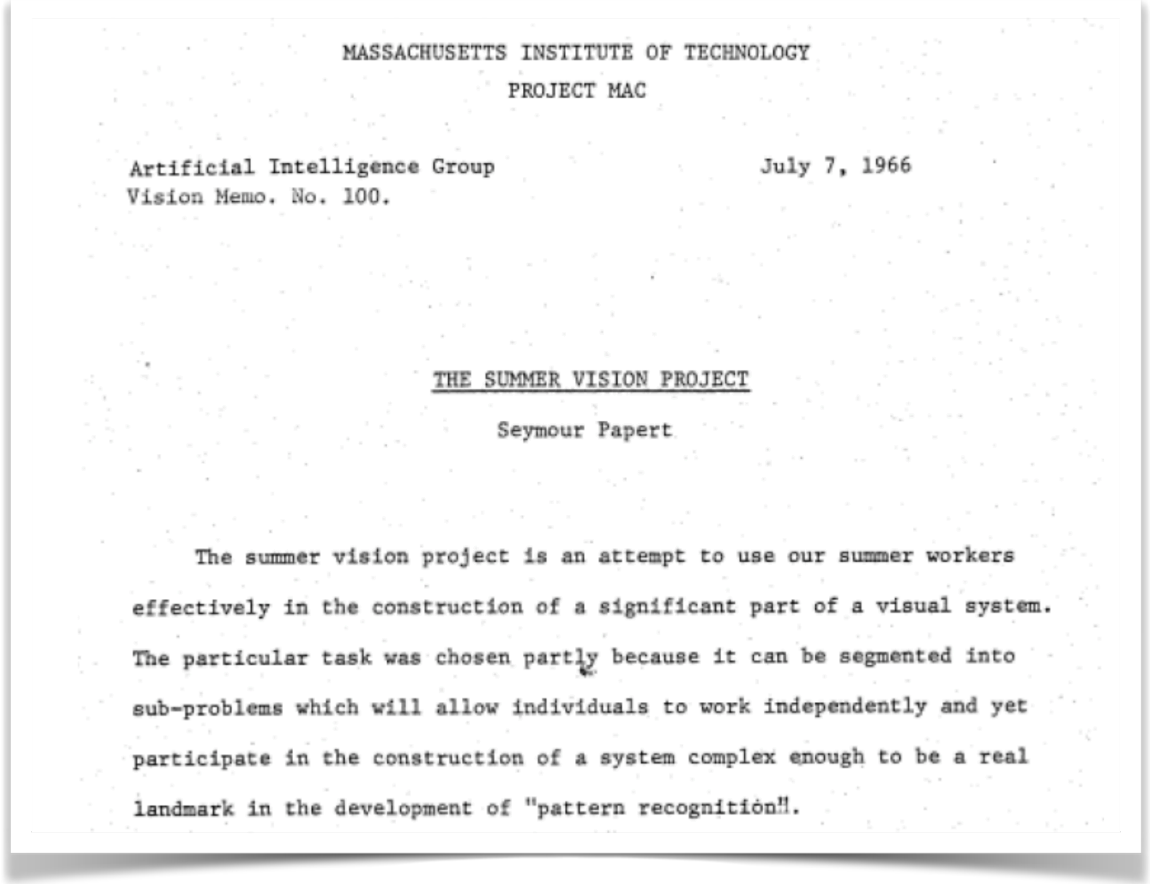
What is **Computer Vision**?

Compute vision, broadly speaking, is a research field aimed to enable computers to **process and interpret visual data**, as sighted humans can.



This implies a very strong connection between **Computer Vision** and **(Machine) Learning**

Computer vision … the beginning …

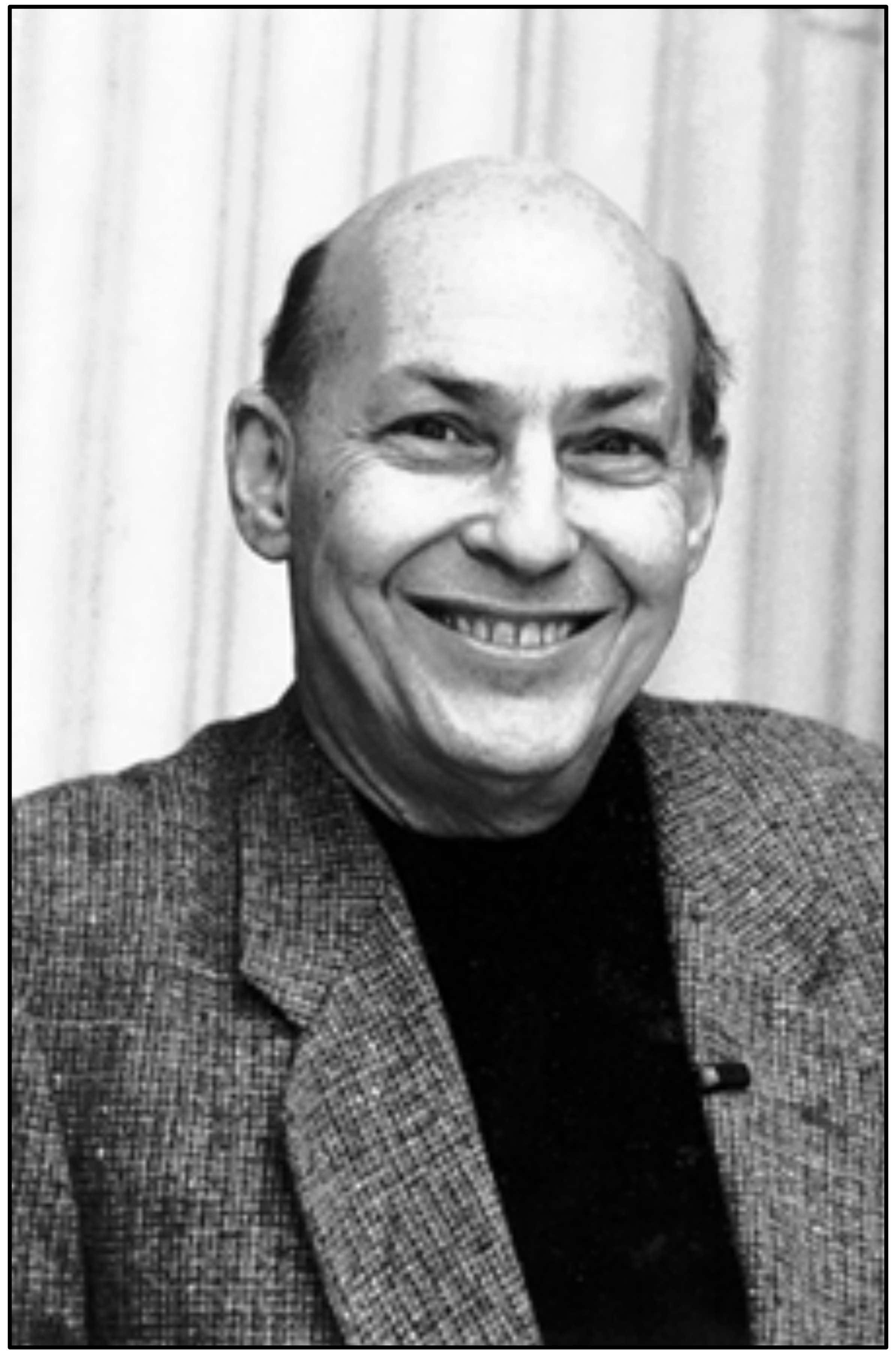

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

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".



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

Slide Credit: Devi Parikh (GA Tech)

Computer vision … the beginning …

Gerald Sussman, MIT

"You'll notice that **Sussman** never worked in vision again!" – Berthold Horn

Slide Credit: Devi Parikh (GA Tech)

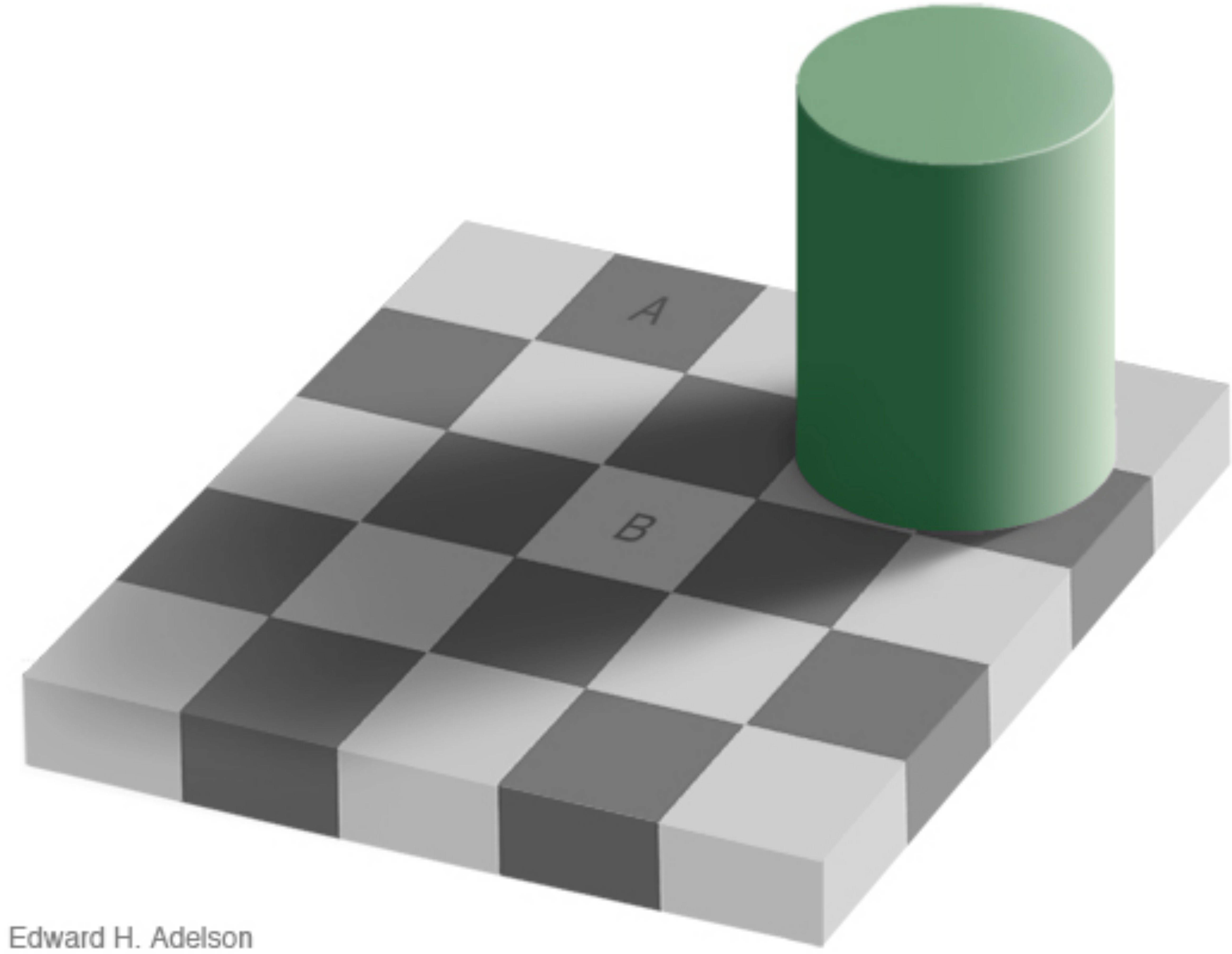
Can computers **match (or beat)** human vision?

- We've been at it for 50 years

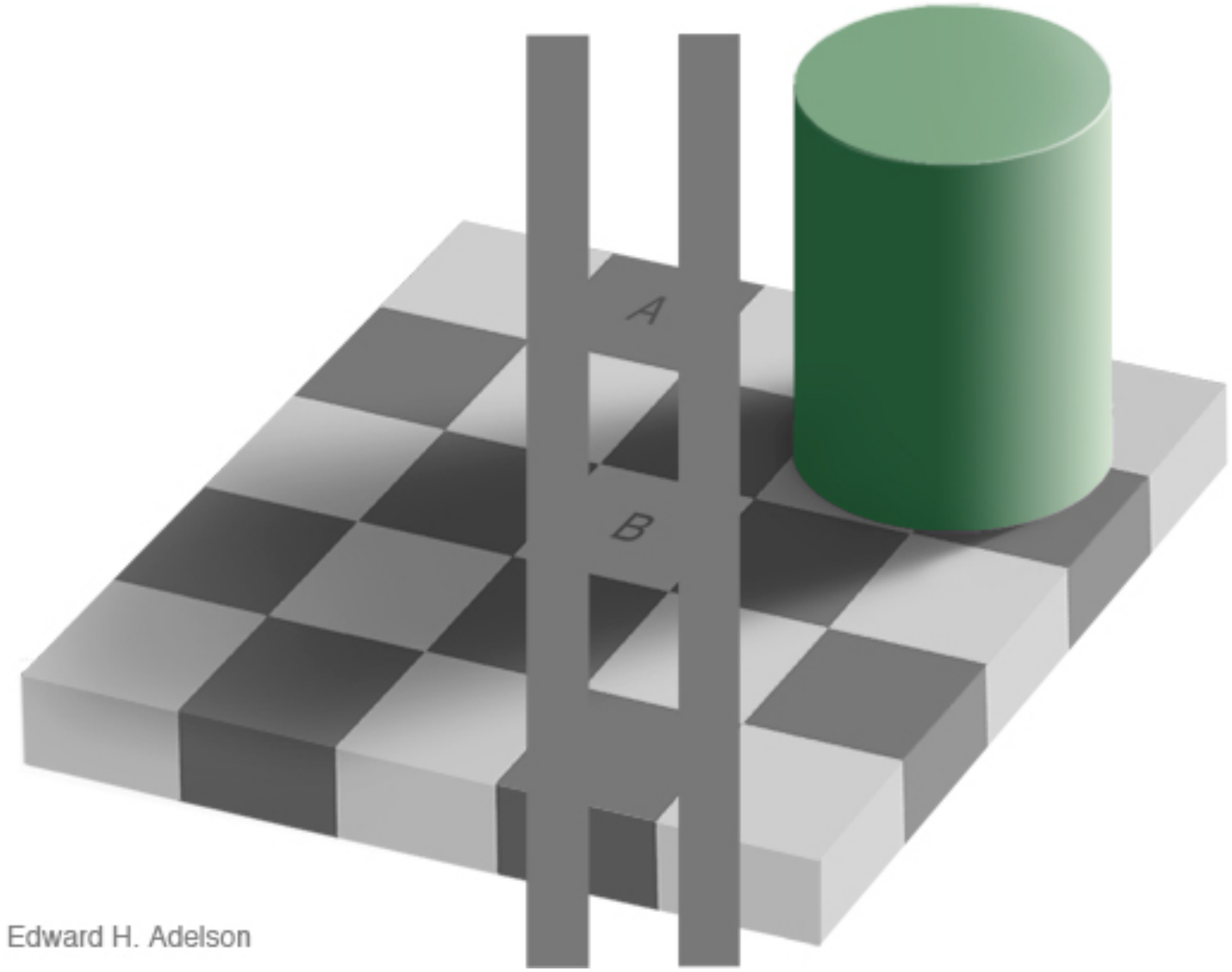
Can computers **match (or beat)** human vision?

- 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?

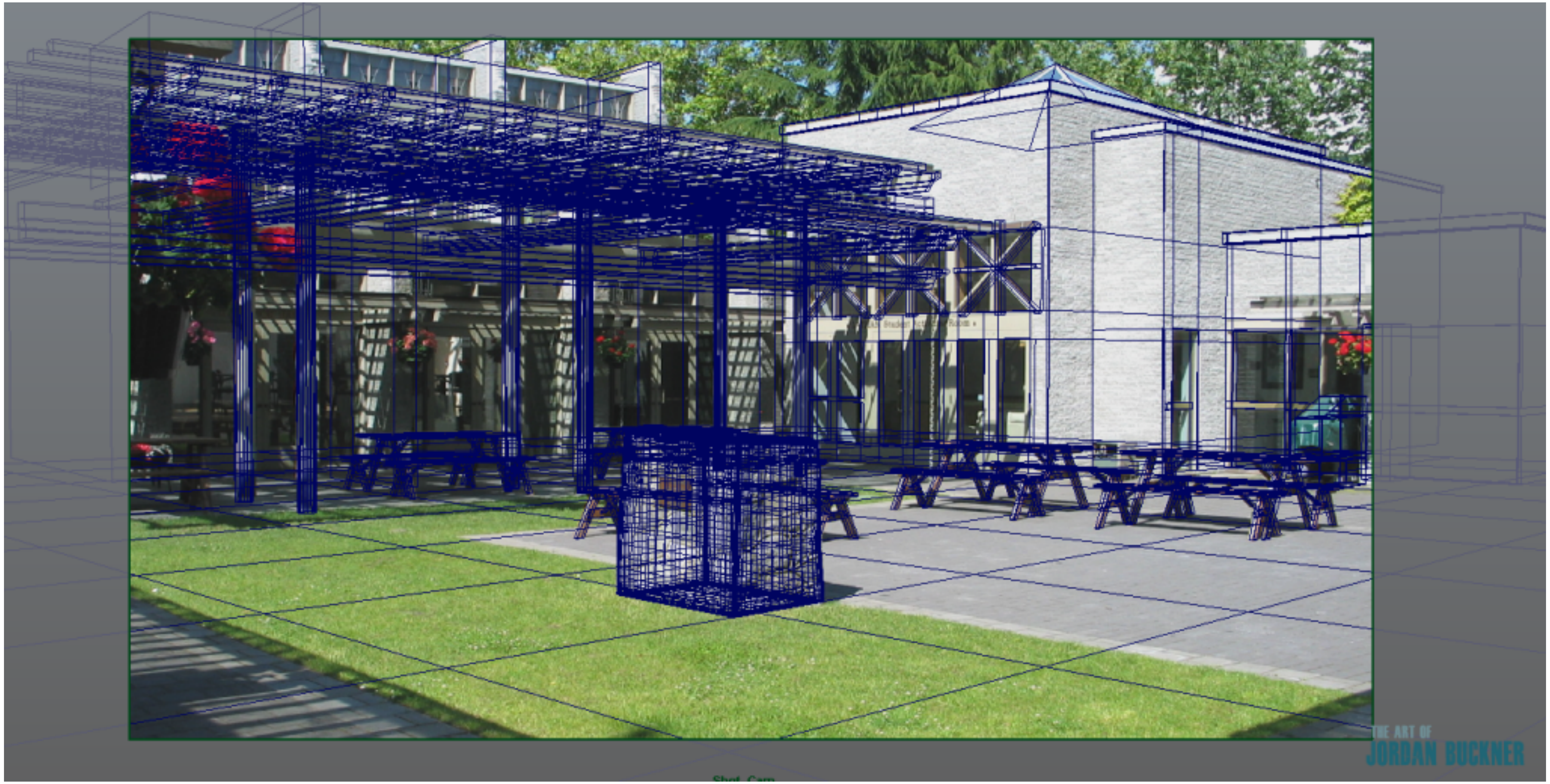
  As a measuring device not very good, as a functioning device really good

Can computers **match (or beat)** human vision?

- **Yes and No** (mostly NO)

Alternative definition of computer vision

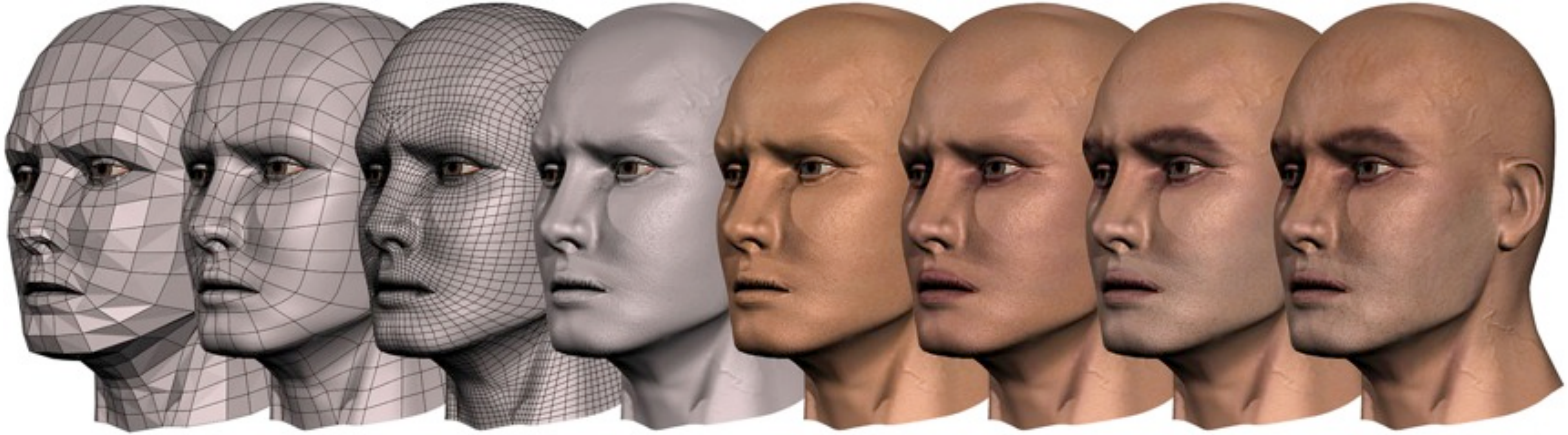
“Inverse Computer Graphics”

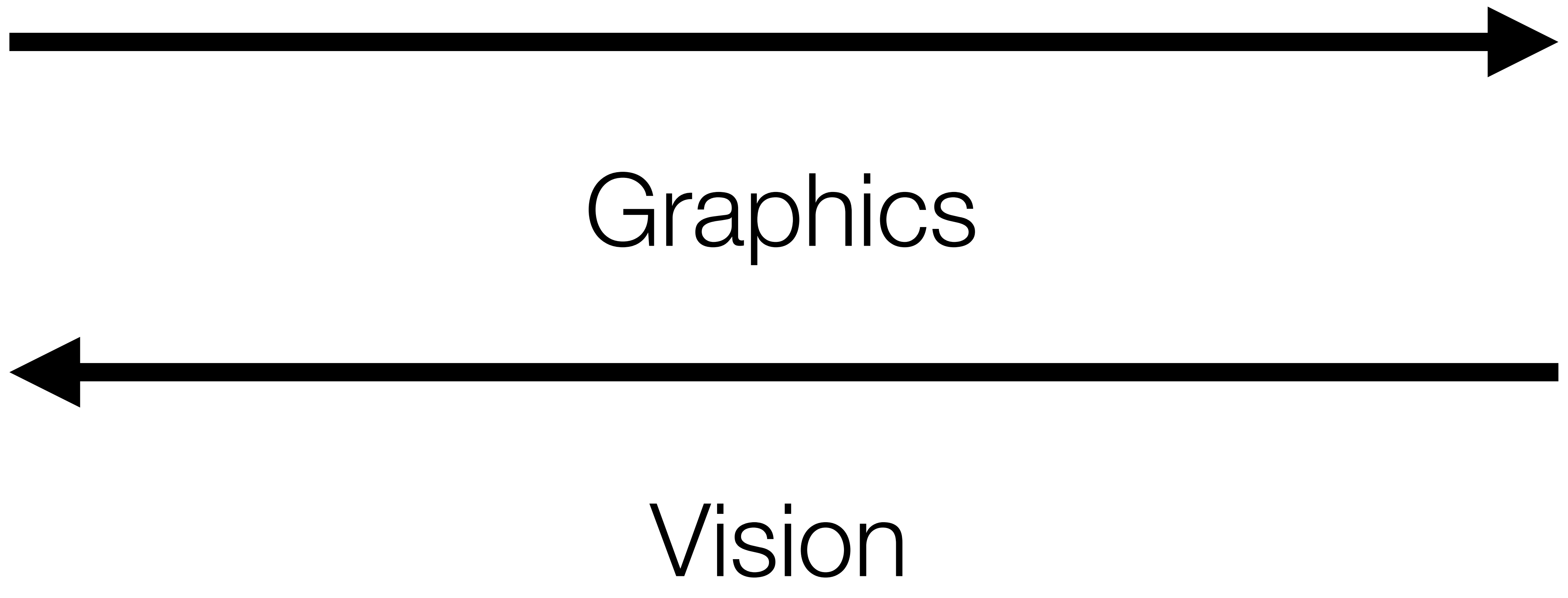


[J. Buckner]

Alternative definition of computer vision

“Inverse Computer Graphics”





Computer **Vision Problems**

Slide Credit: Kristen Grauman (UT Austin)

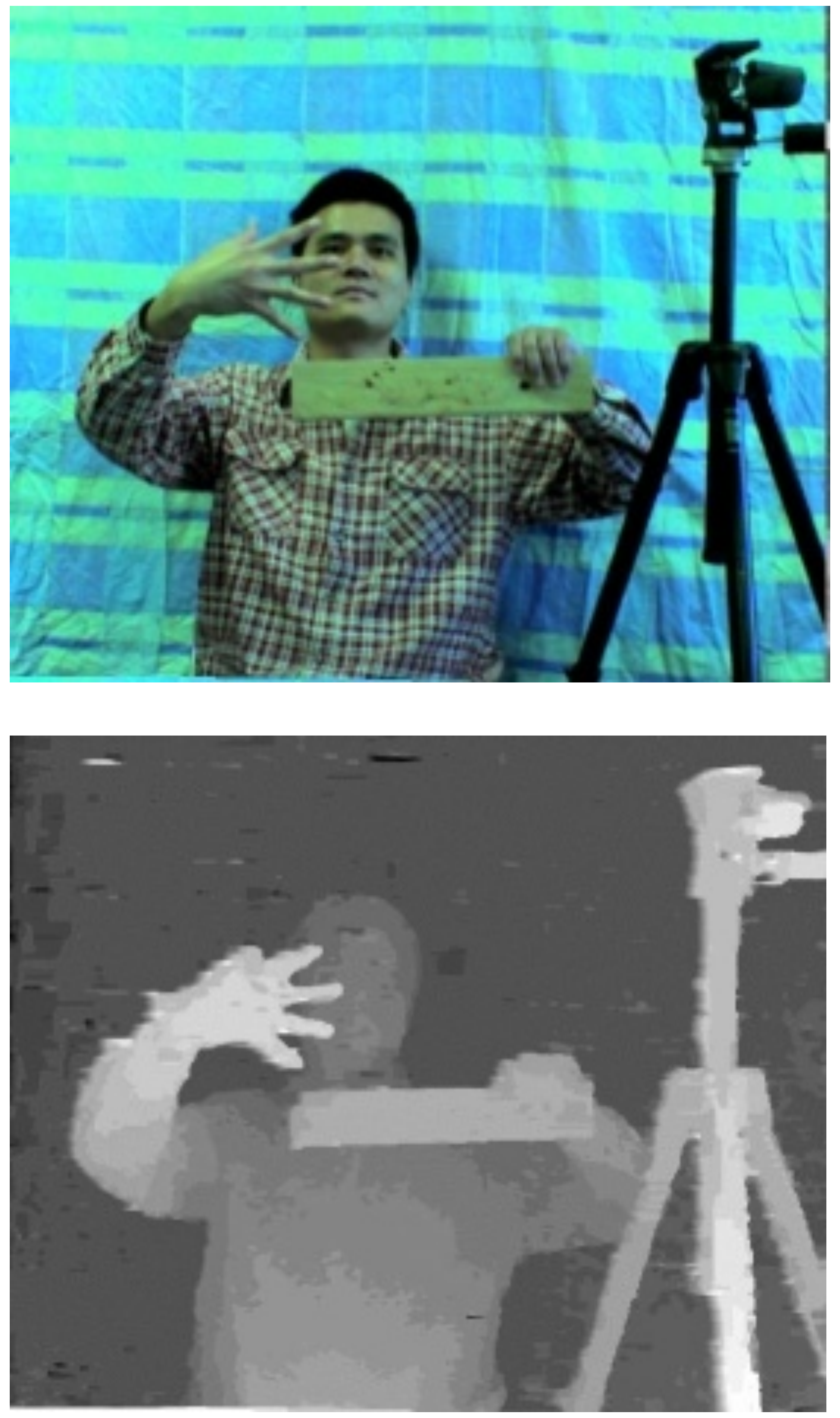
Computer **Vision Problems**

1. Computing properties of the 3D world from visual data (***measurement***)

Slide Credit: Kristen Grauman (UT Austin)

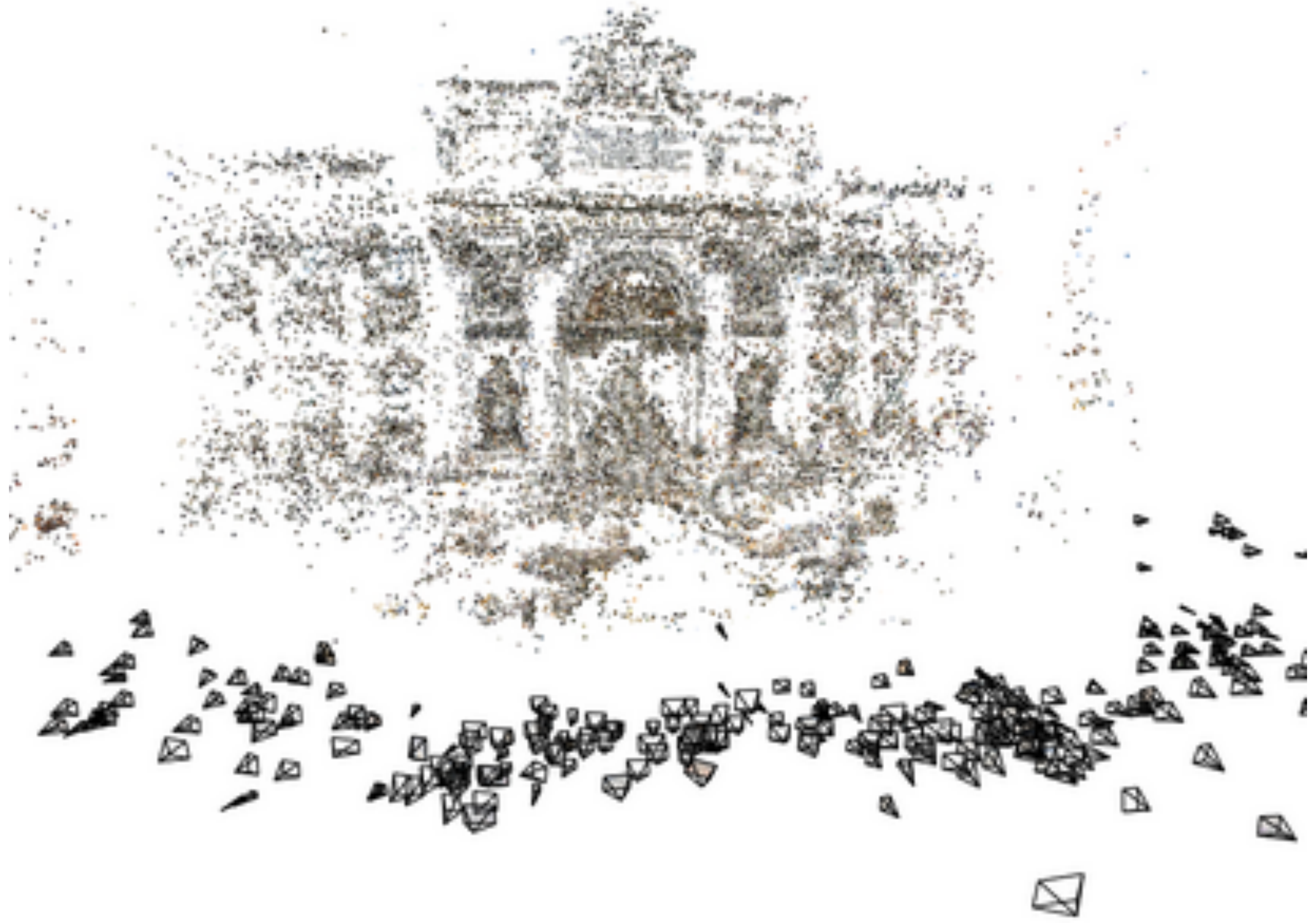
1. Vision for **Measurement**

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

Slide Credit: Kristen Grauman (UT Austin)

Computer **Vision Problems**

1. Computing properties of the 3D world from visual data (***measurement***)

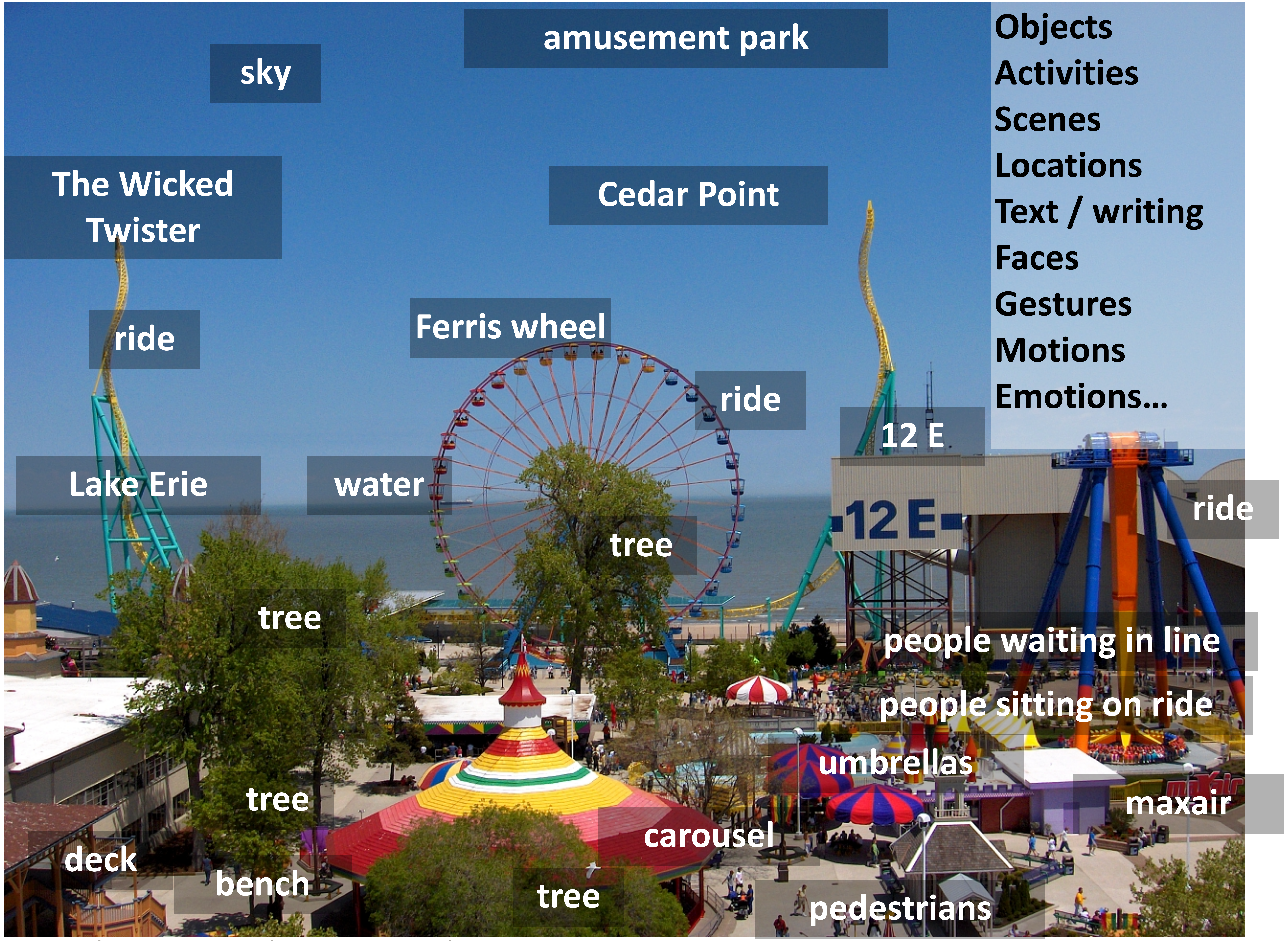
2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (***perception and interpretation***)

Slide Credit: Kristen Grauman (UT Austin)

2. Vision for **Perception and Interpretation**

Slide Credit: Kristen Grauman (UT Austin)

2. Vision for **Perception and Interpretation**



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

Slide Credit: Kristen Grauman (UT Austin)

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

Slide Credit: Kristen Grauman (UT Austin)

Computer **Vision Problems**

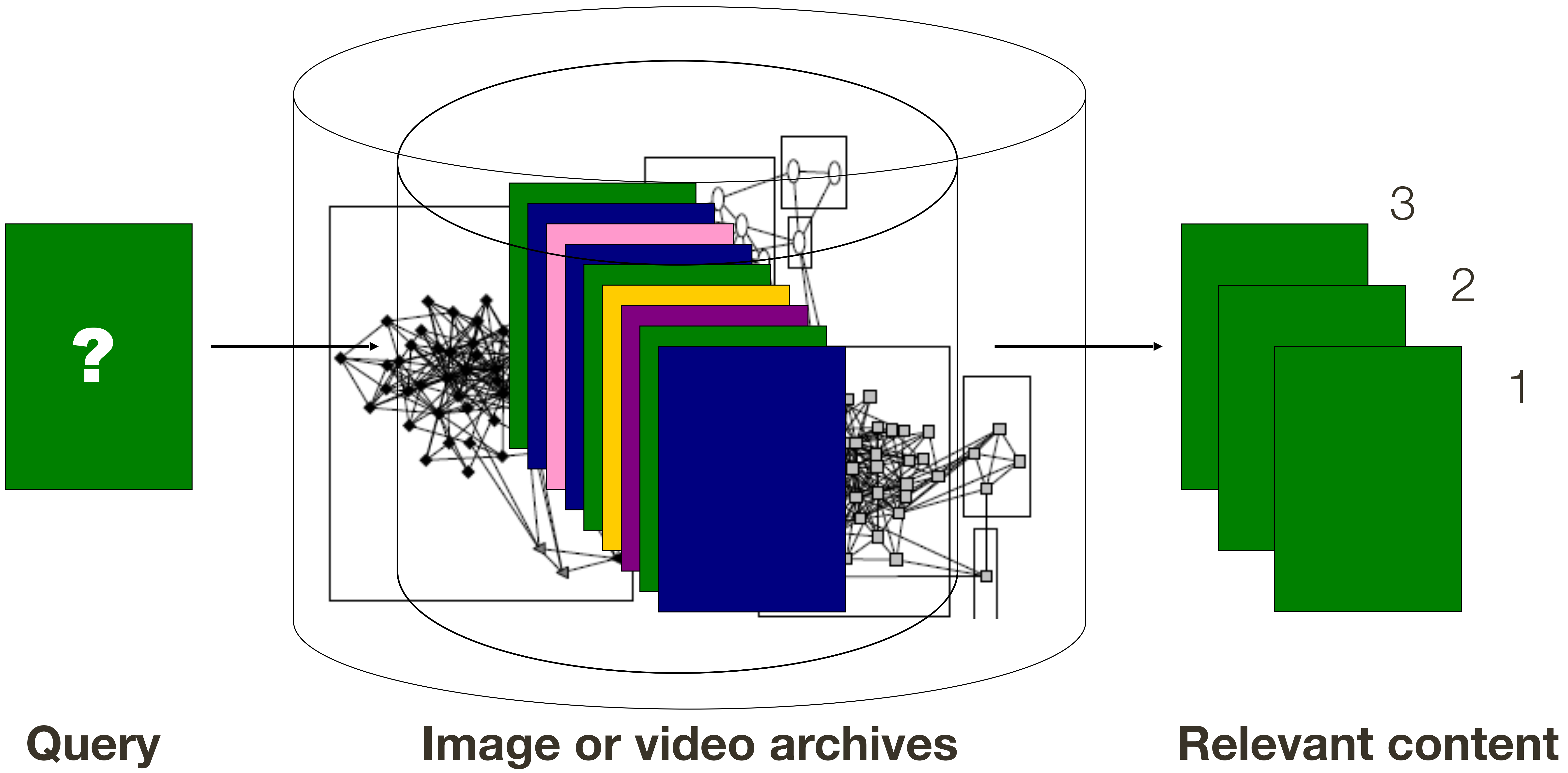
1. Computing properties of the 3D world from visual data (***measurement***)

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***)

Slide Credit: Kristen Grauman (UT Austin)

3. **Search and Organization**



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

Slide Credit: Kristen Grauman (UT Austin)

3. **Search and Organization**

*from iStock by GettyImages

3. **Search and Organization**

*from iStock by GettyImages

Snapchat

31.7 Million
/ hour

WhatsApp

29.2 Million
/ hour

Facebook

14.6 Million
/ hour

Instagram

2.9 Million
/ hour

Flickr

0.2 Million
/ hour

YouTube

18K hours
/ hour

*based on article by Kimberlee Morrison in Social Times (2015)

3. **Search and Organization**

*from iStock by GettyImages

> 85% of all web content is multimedia content of visual form

Snapchat	WhatsApp	Facebook
31.7 Million	**29.2 Million**	**14.6 Million** our

Instagram	Flickr	YouTube
2.9 Million / hour	**0.2 Million** / hour	**18K hours** / hour

*based on article by Kimberlee Morrison in Social Times (2015)

Computer **Vision Problems**

1. Computing properties of the 3D world from visual data (***measurement***)

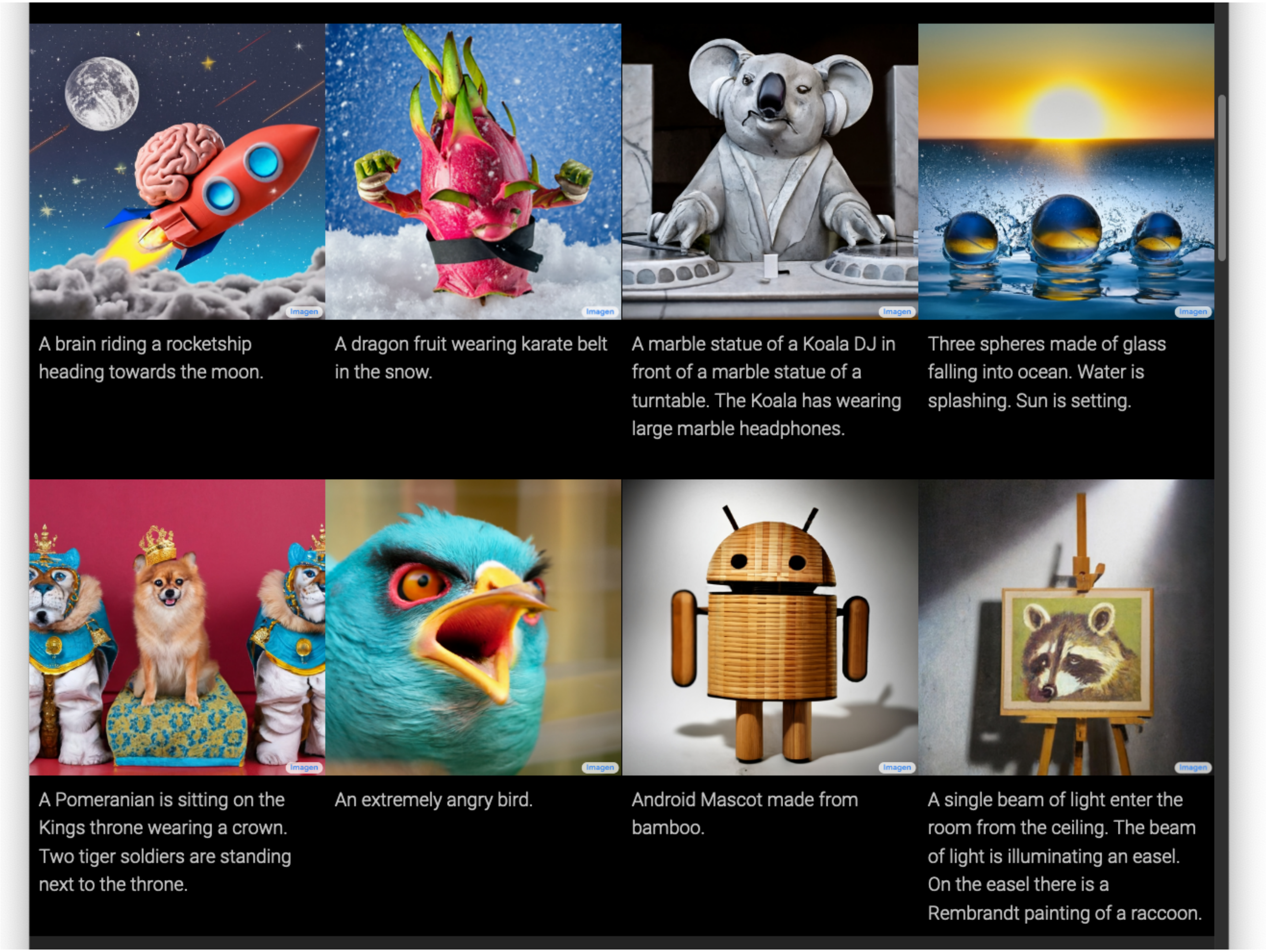
2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (***perception and interpretation***)

3. Algorithms to mine, search, and interact with visual data (***search and organization***)

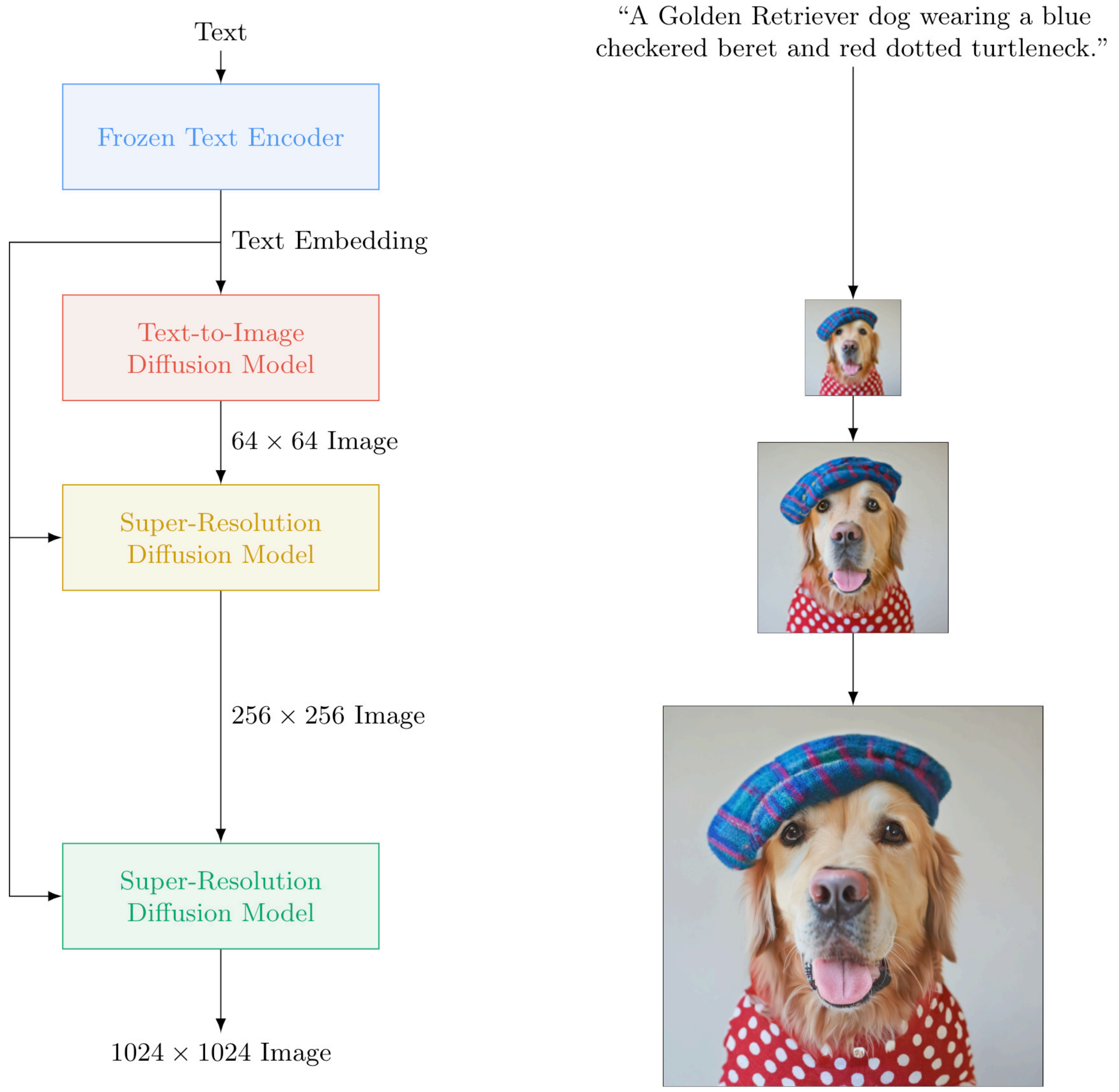
4. Algorithms for manipulation or creation of image or video content (***visual imagination***)

Slide Credit: Kristen Grauman (UT Austin)

4. Visual **Imagination**



- imagen.research.google
- Text to image generation
- Uses diffusion process, training using large dataset of text (web scale) and image-text (400M) pairs



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

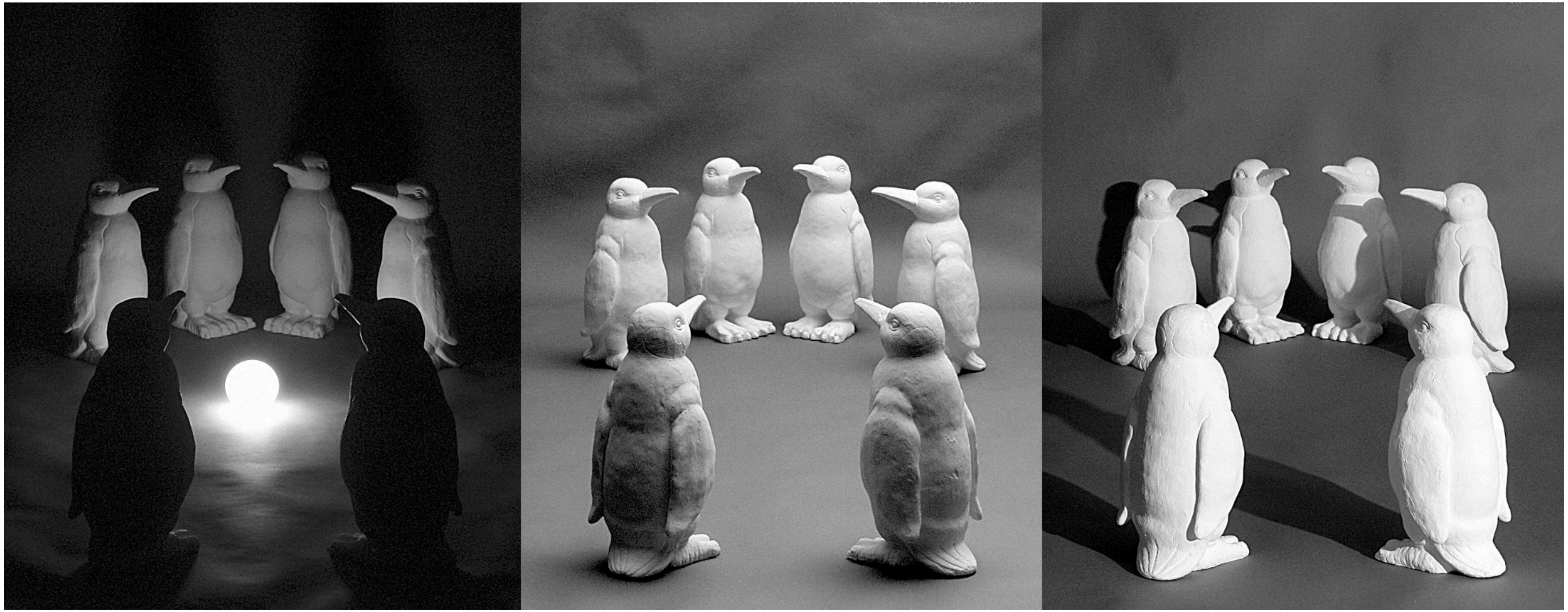


*image credit J. Koenderink

Challenges: Scale

*slide credit Fei-Fei, Fergus & Torralba

Challenges: Deformation

*image credit Peter Meer

Challenges: Occlusions

Rene Magritte 1965

Challenges: Background clutter

Kilmeny Niland 1995

Challenges: Local ambiguity and context

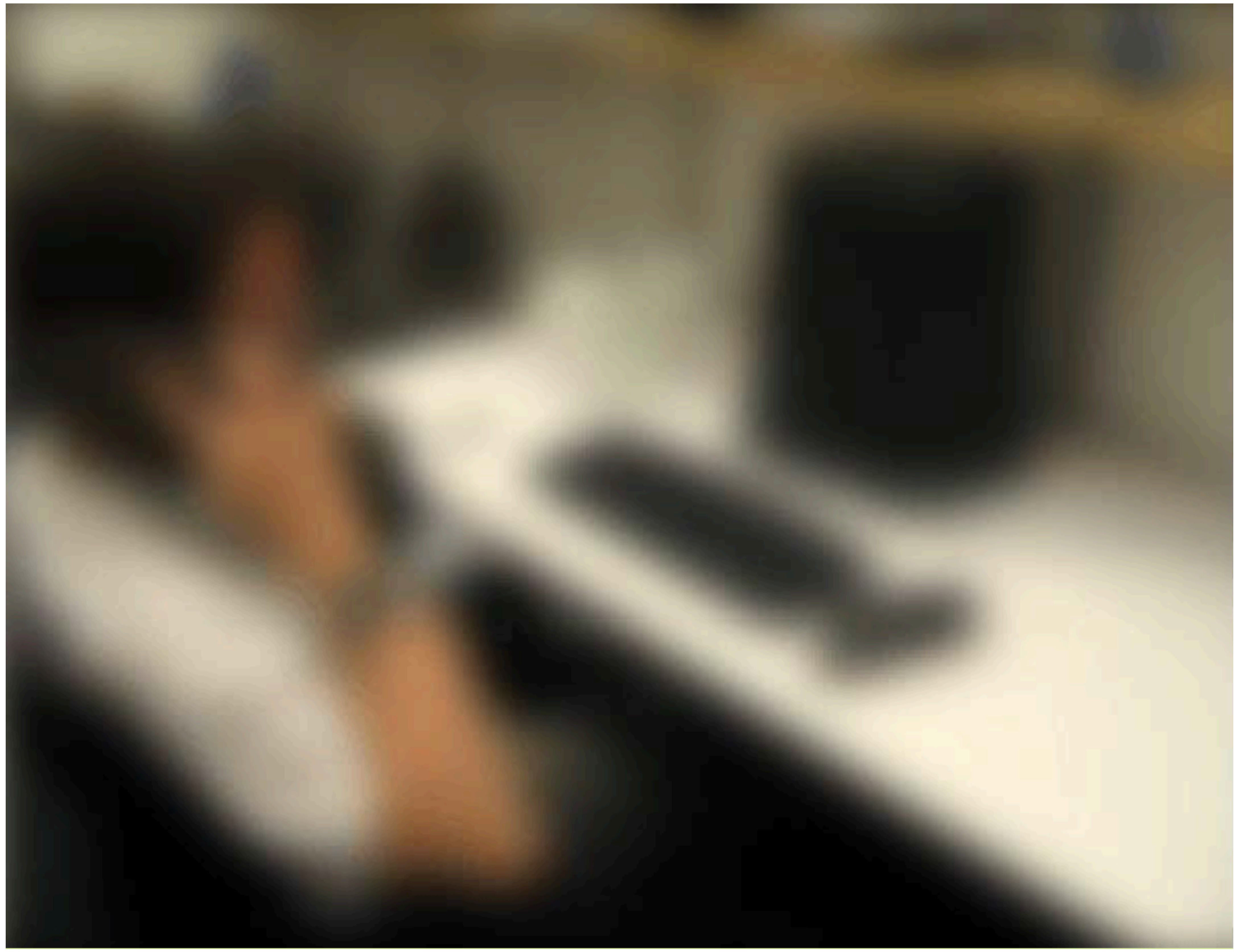


*image credit Fergus & Torralba

Challenges: Local ambiguity and context

*image credit Fergus & Torralba

Challenges: Motion

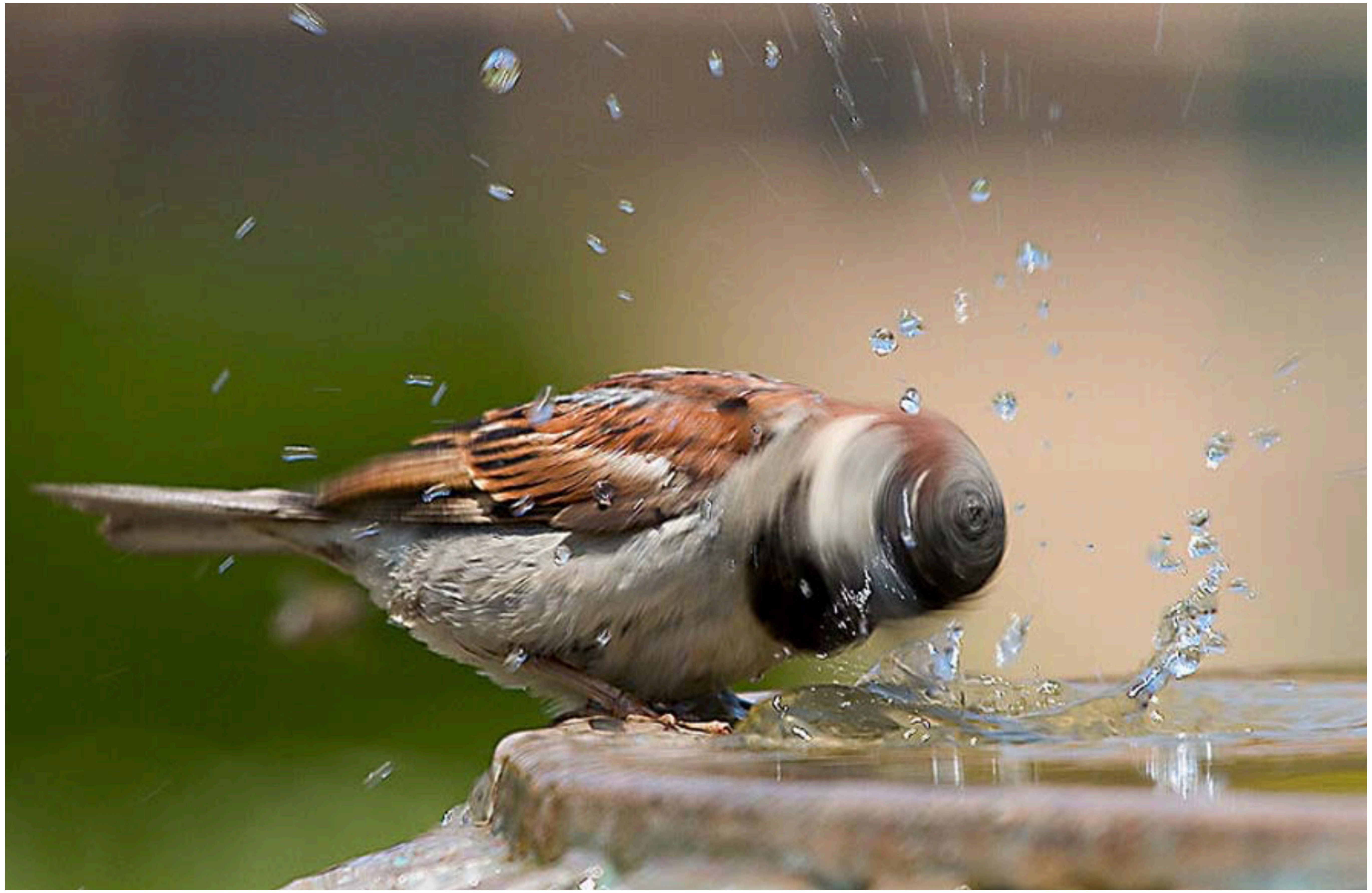


*image credit Peter Meer

Challenges: Object inter-class variation






*slide credit Fei-Fei, Fergus & Torralba

Computer Vision **Applications**

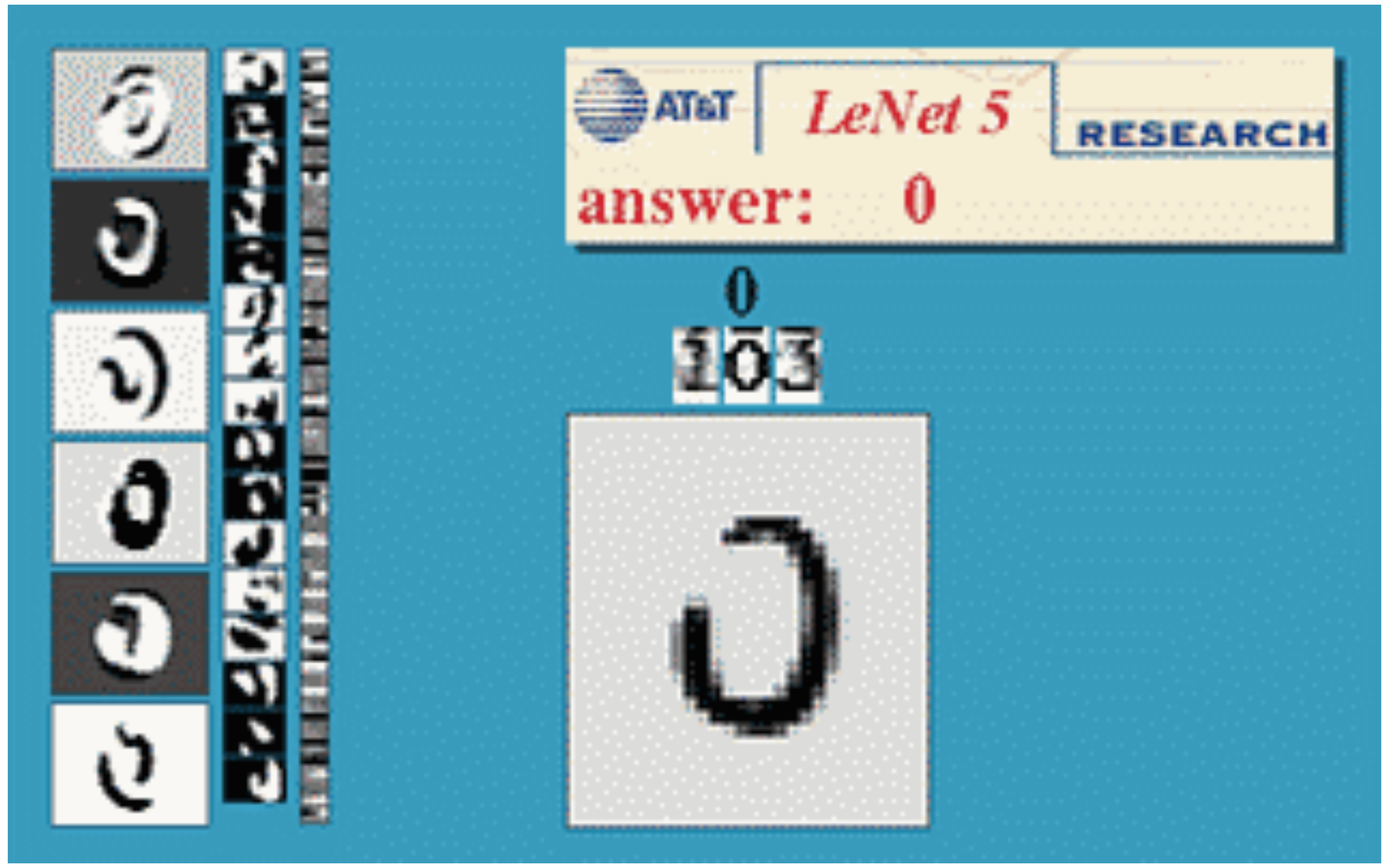
- Let’s see some examples of state-of-the-art and where it is used

Optical Character Recognition (**OCR**)

Technology to convert **scanned documents to text**

(comes with any scanner now days)

Yann LeCun



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

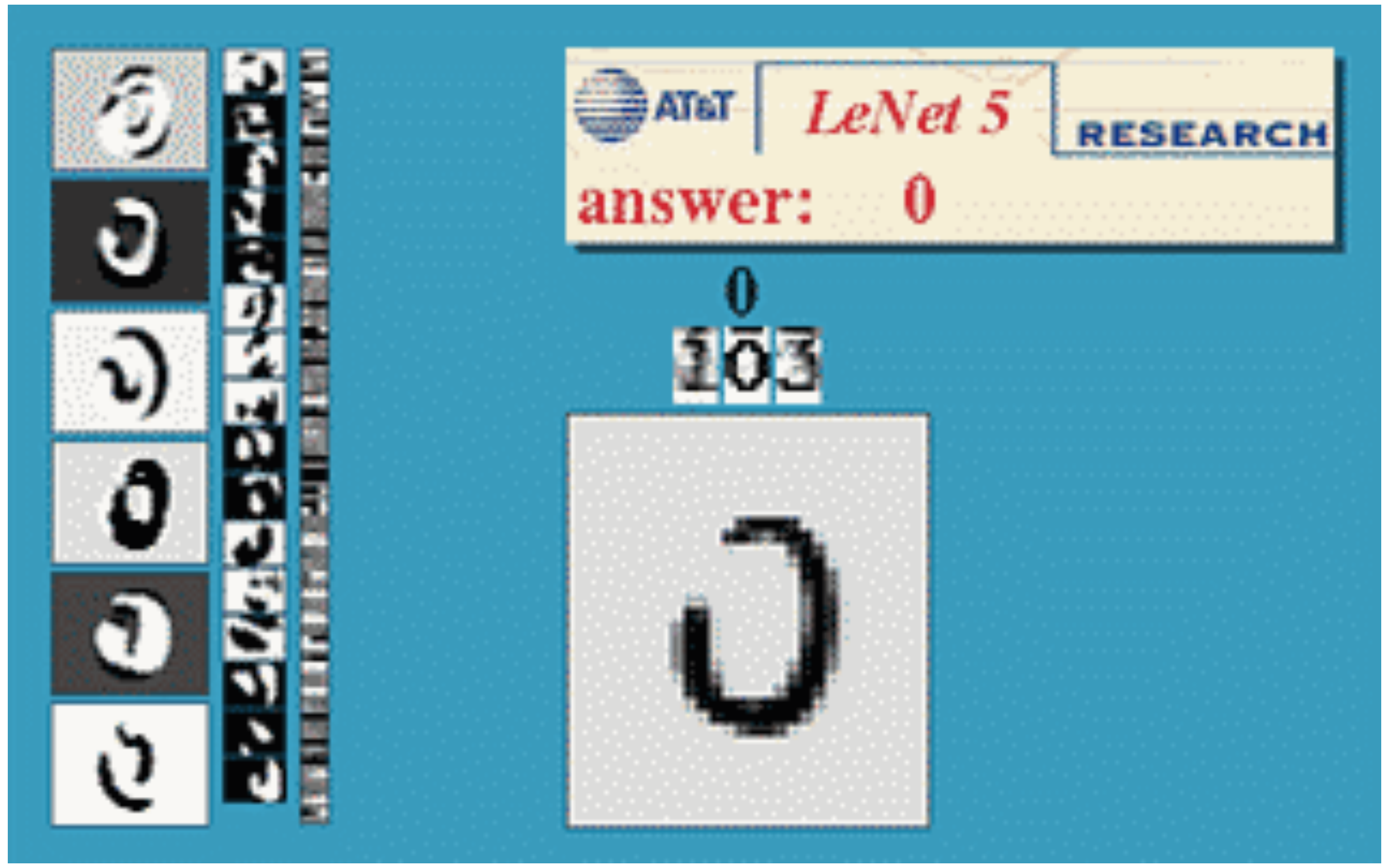
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Optical Character Recognition (**OCR**)

Technology to convert **scanned documents to text**

(comes with any scanner now days)

Yann LeCun



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

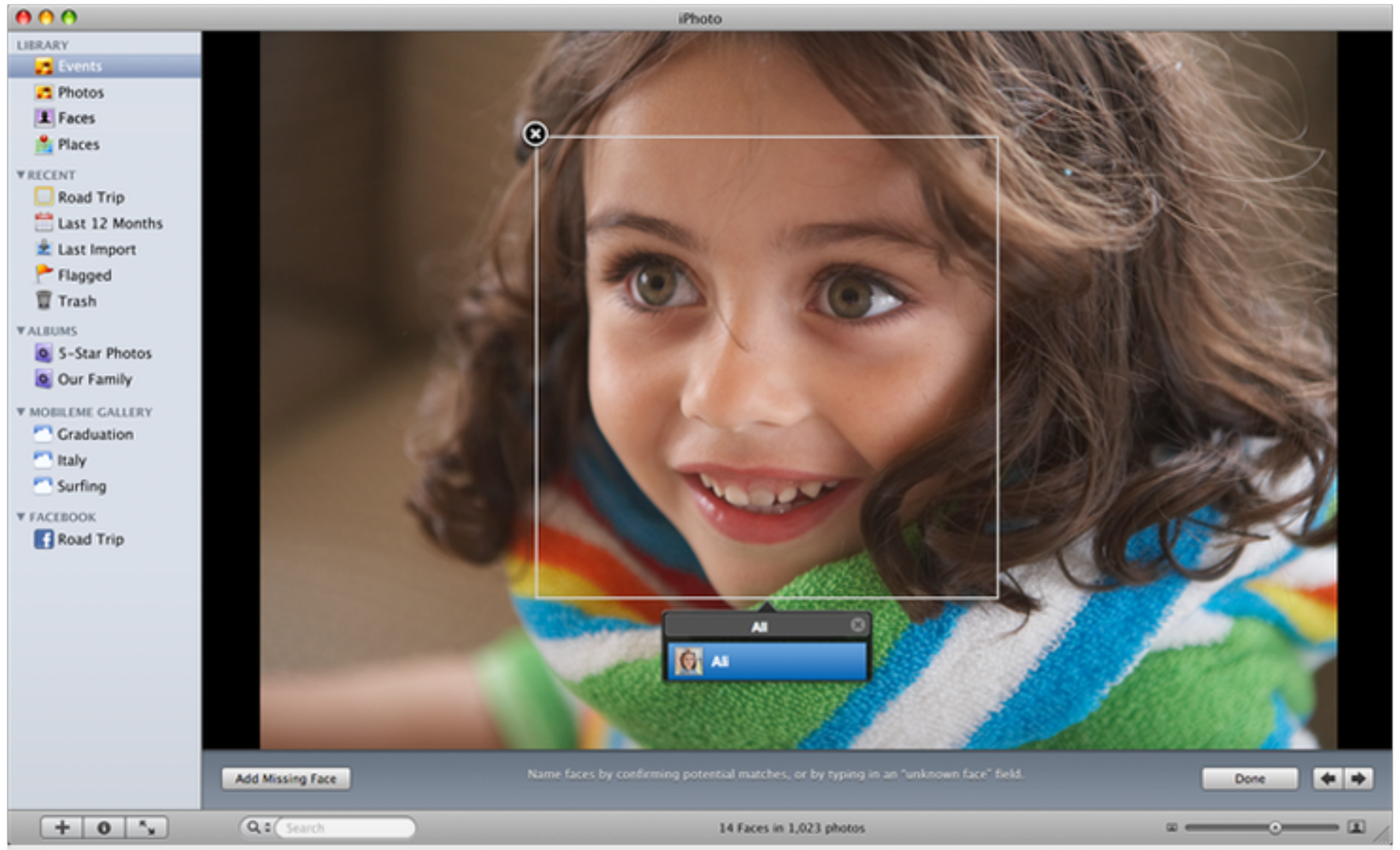
(one of the first big commercial successes of vision algorithms)

[Motorola]

Face **Recognition**

Facebook

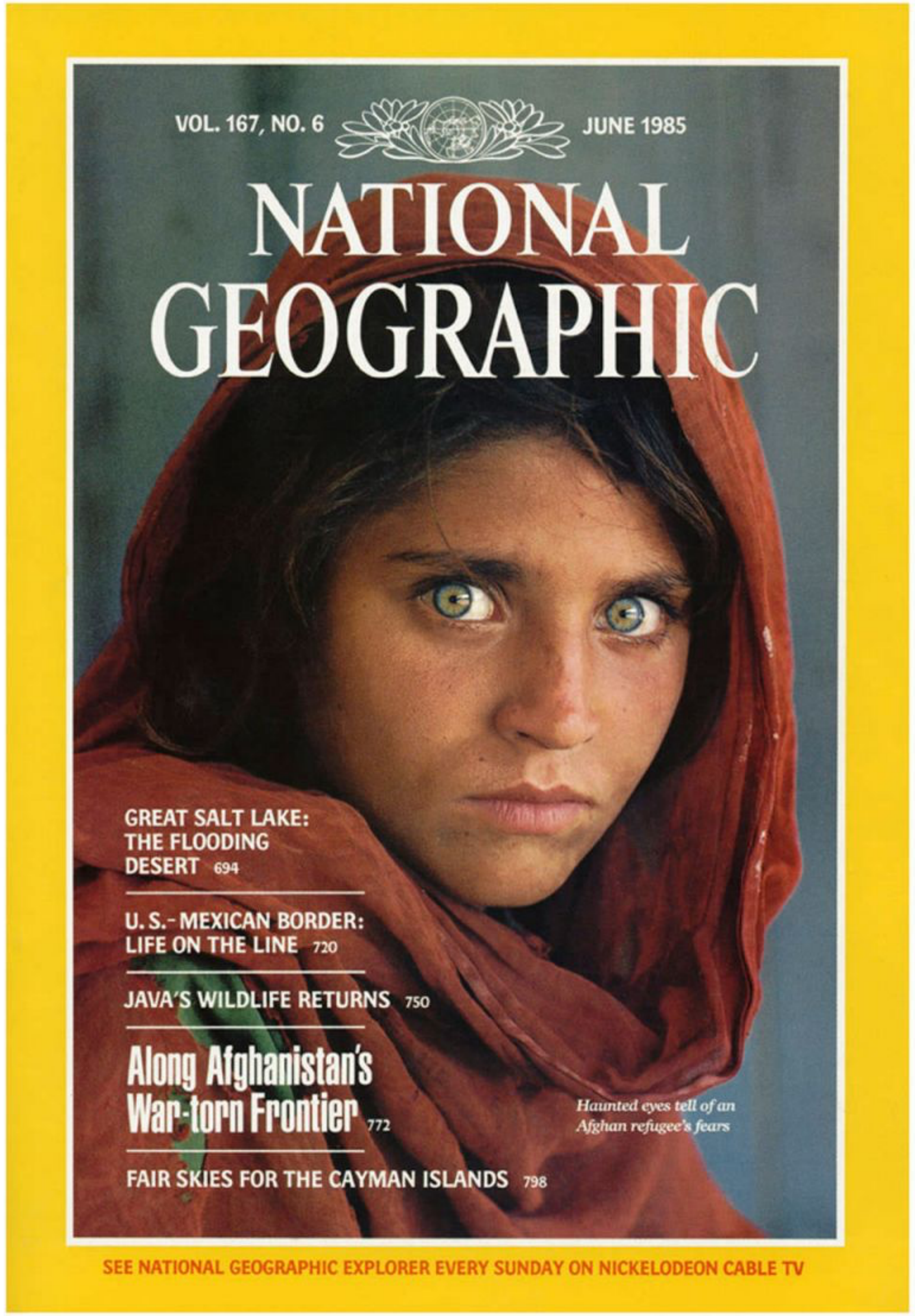
Apple's iPhoto

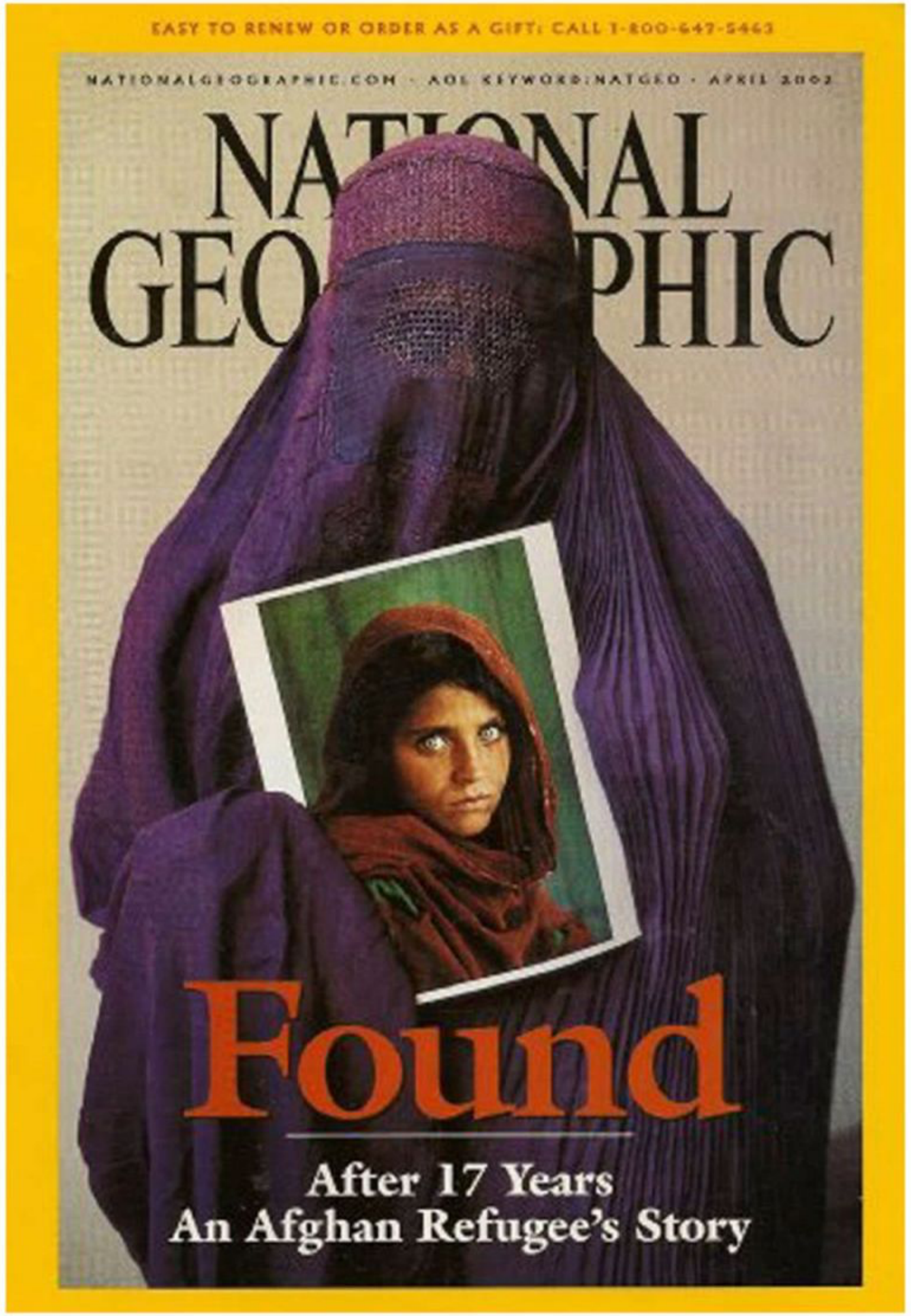


http://www.apple.com/ilife/iphoto/

Slide Credit: Devi Parikh (GA Tech) and Fei-Fei Li (Stanford)

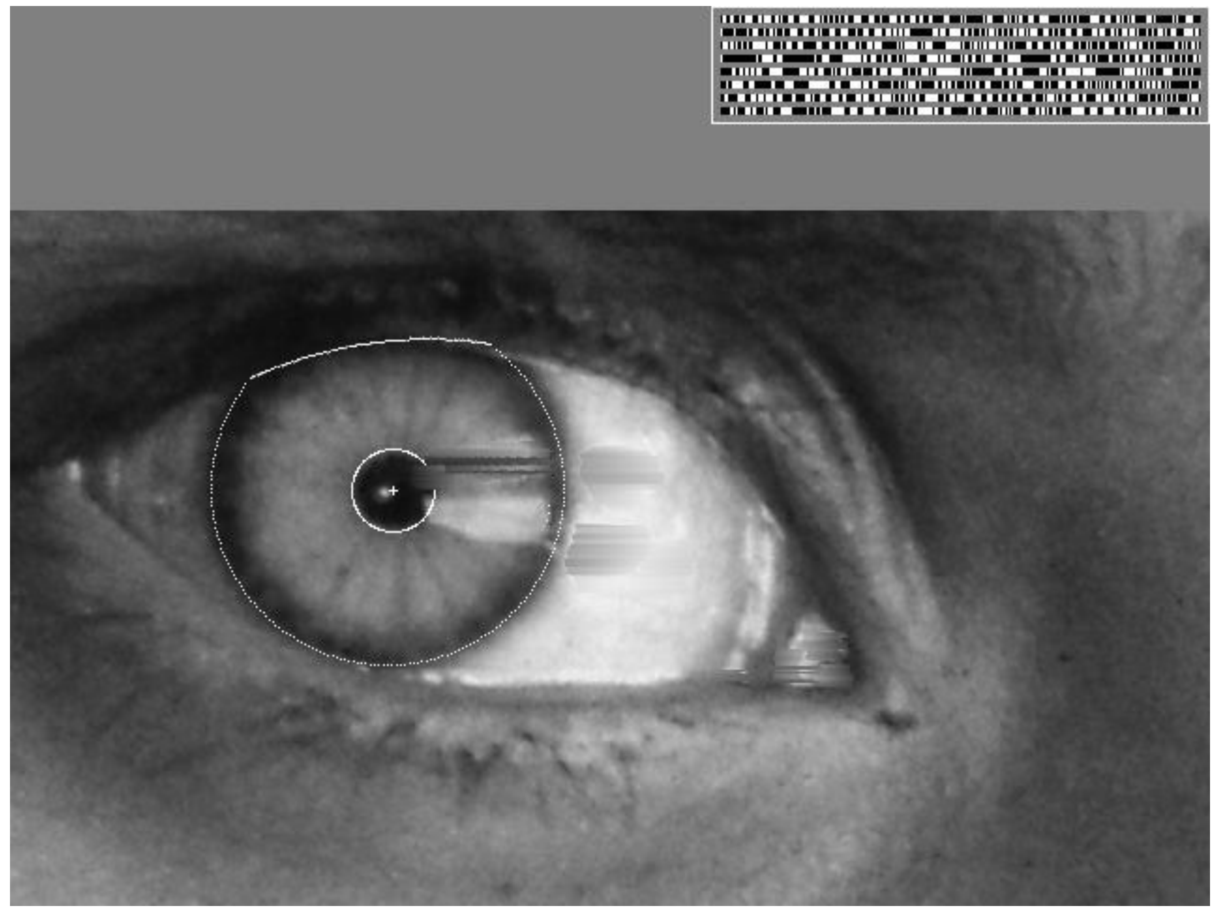
Vision for **Biometrics**

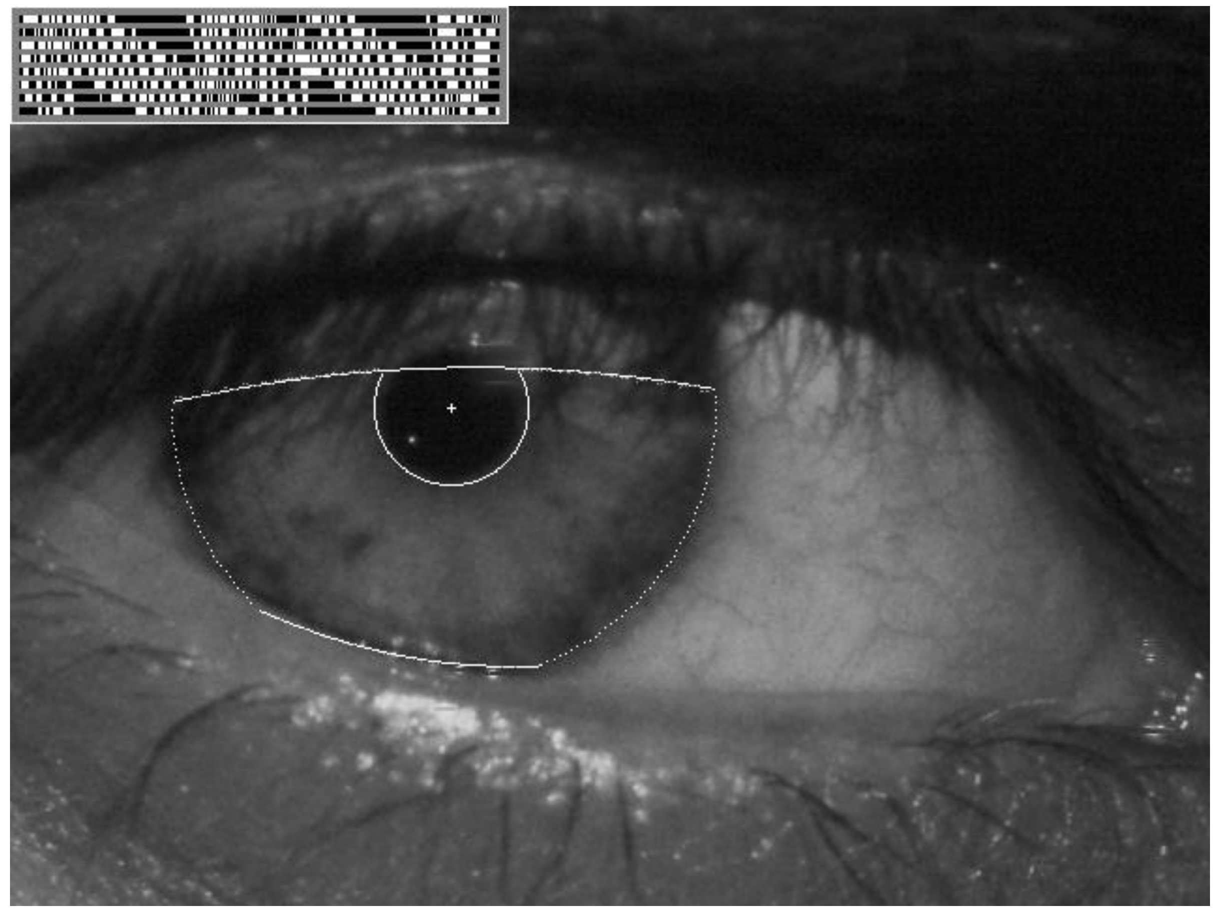




Vision for **Biometrics**

"How the Afghan Girl was Identified by Her Iris Patterns" Read the story wikipedia





Slide Credit: James Hays (GA Tech)

Vision for **Biometrics**

Fingerprint scanners on many new laptops, other devices

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

Image Credit: James Hays (GA Tech)

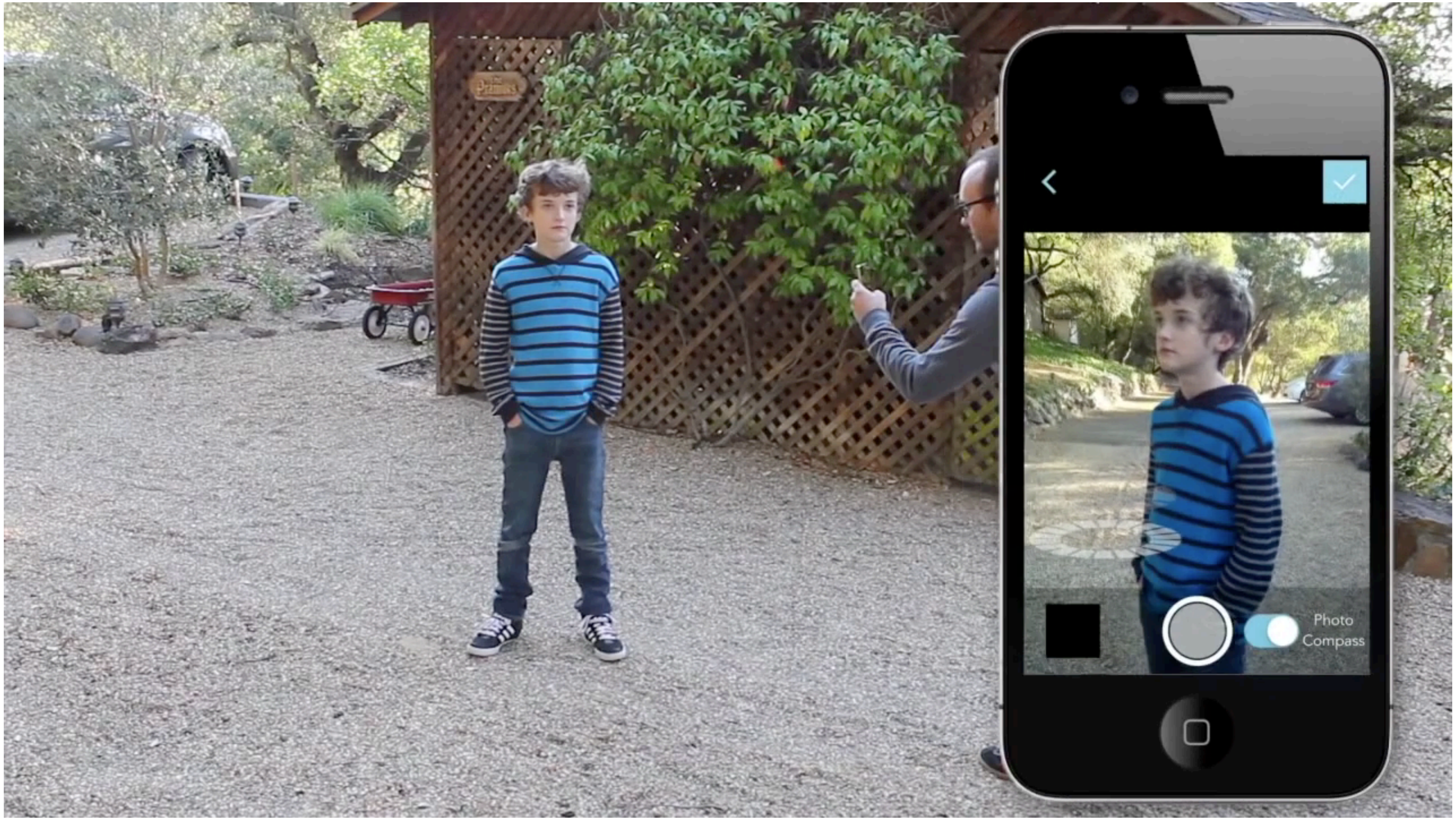
Camera **Tracking**

[Boujou — Vicon / OMG]

Camera **Tracking**

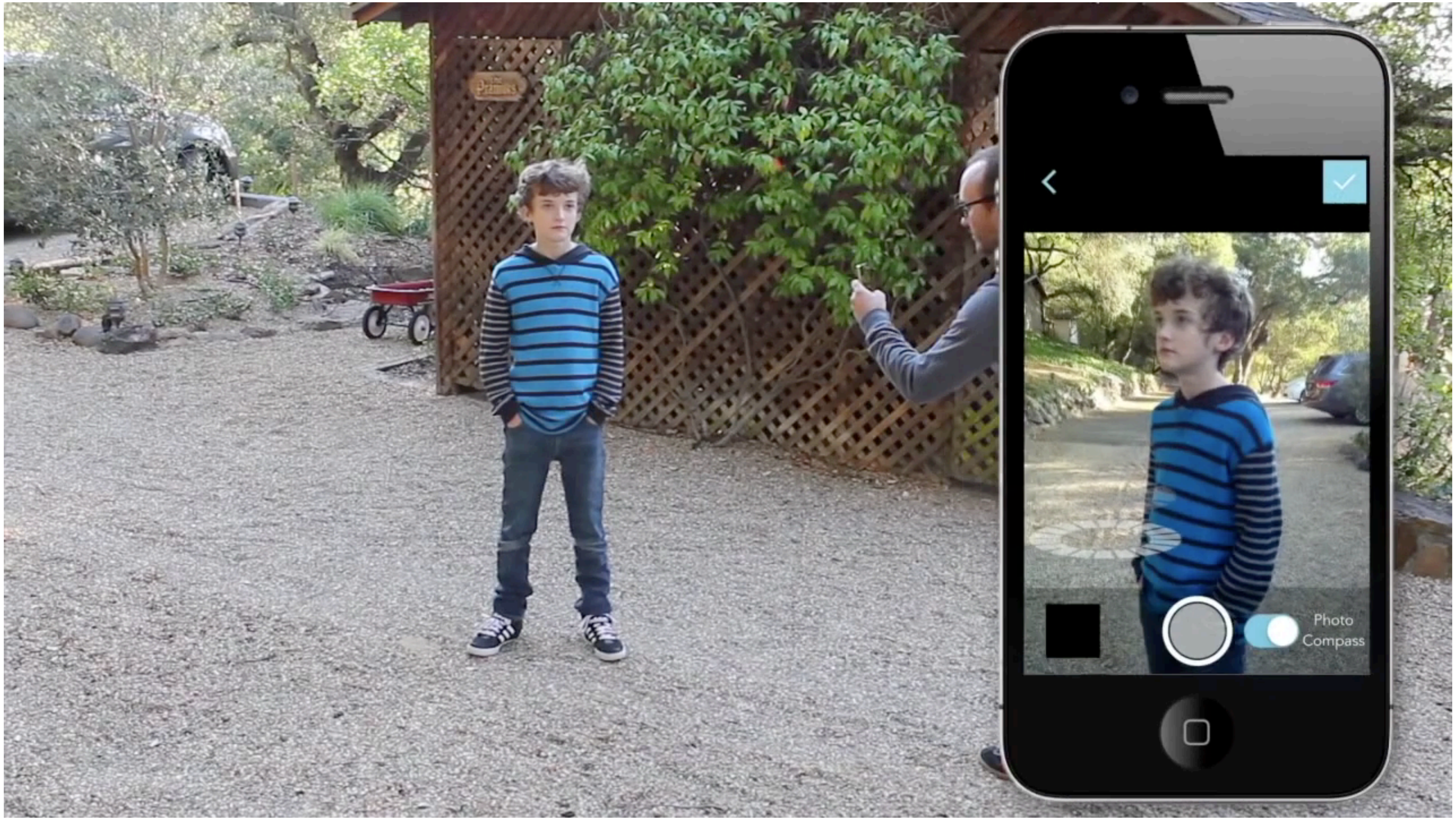
[Boujou — Vicon / OMG]

3D **Reconstruction**



[Autodesk 123D Catch]

3D **Reconstruction**

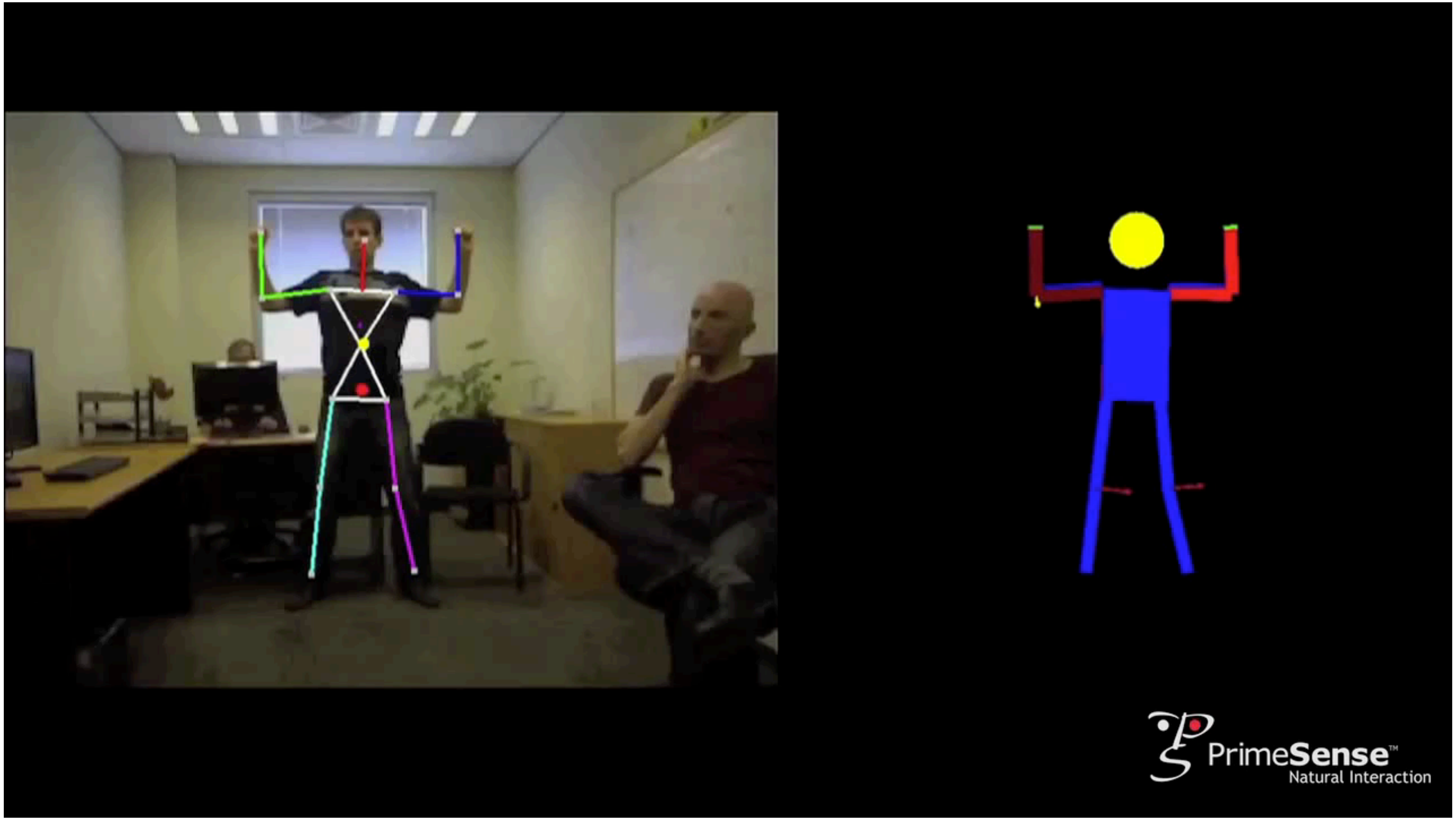


[Autodesk 123D Catch]

Body Pose Tracking

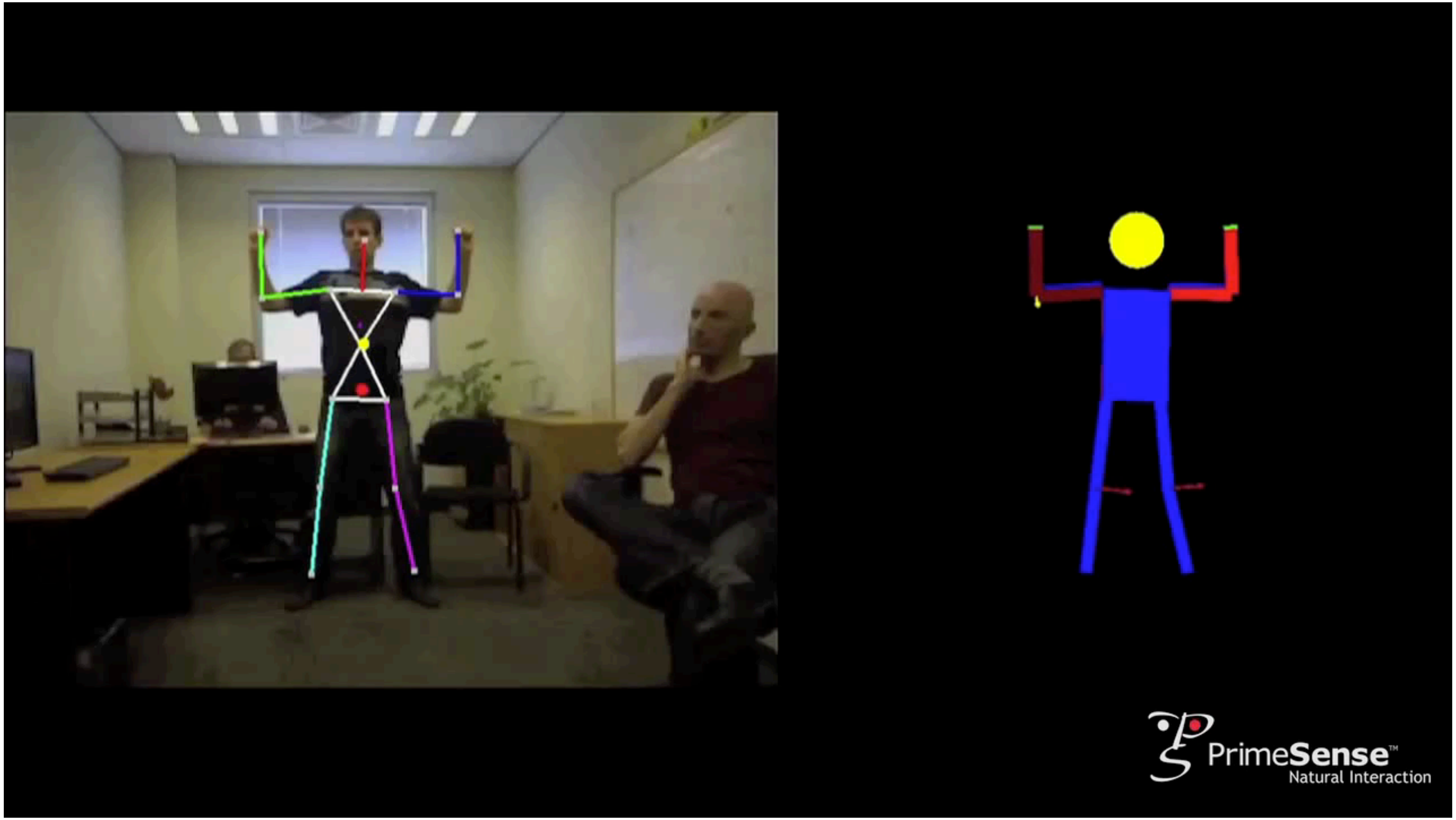
[Microsoft Xbox Kinect]

Body Pose Tracking



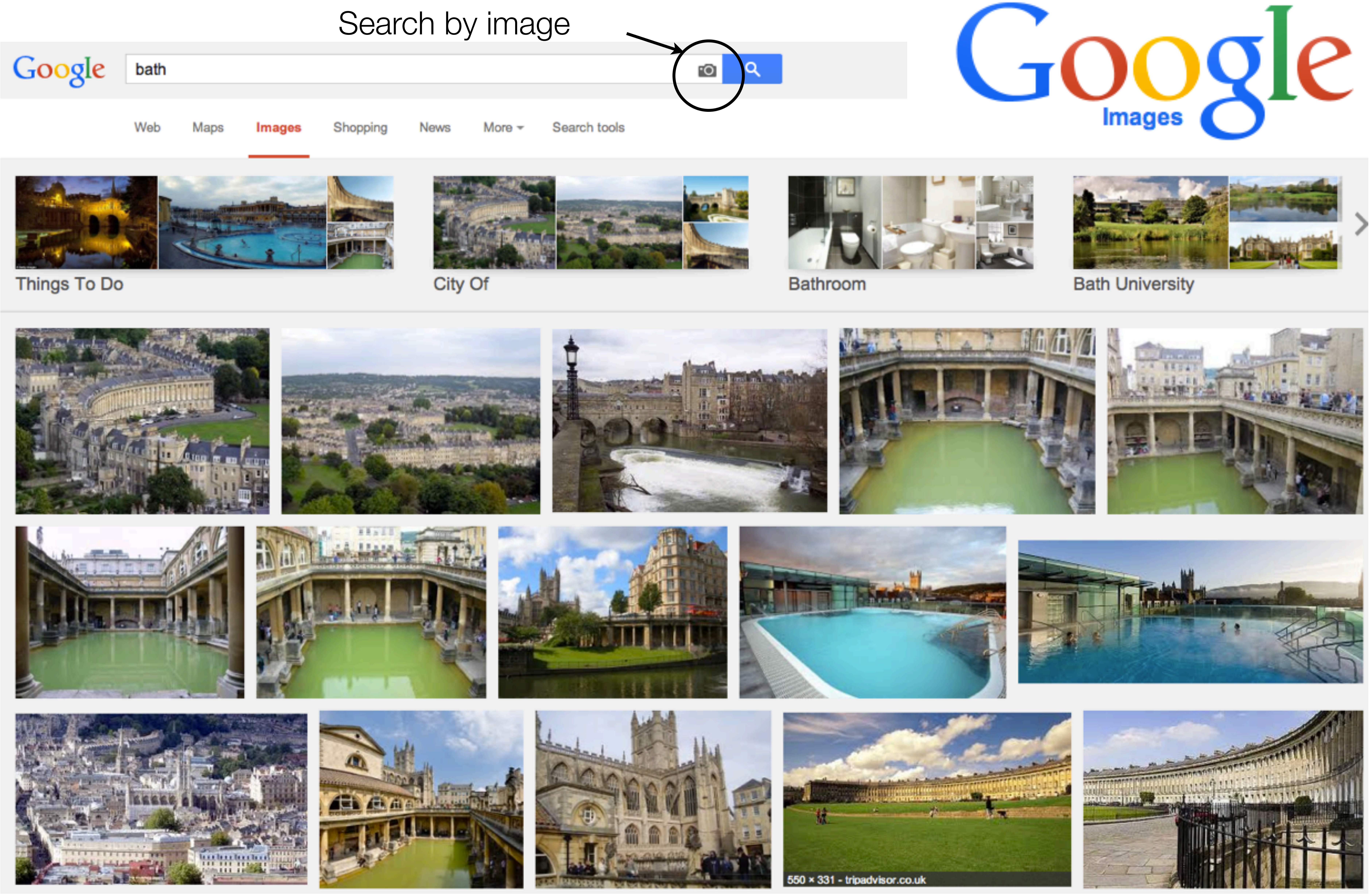
[PrimeSense]

Body Pose Tracking



[PrimeSense]

Image **Recognition** and **Search**



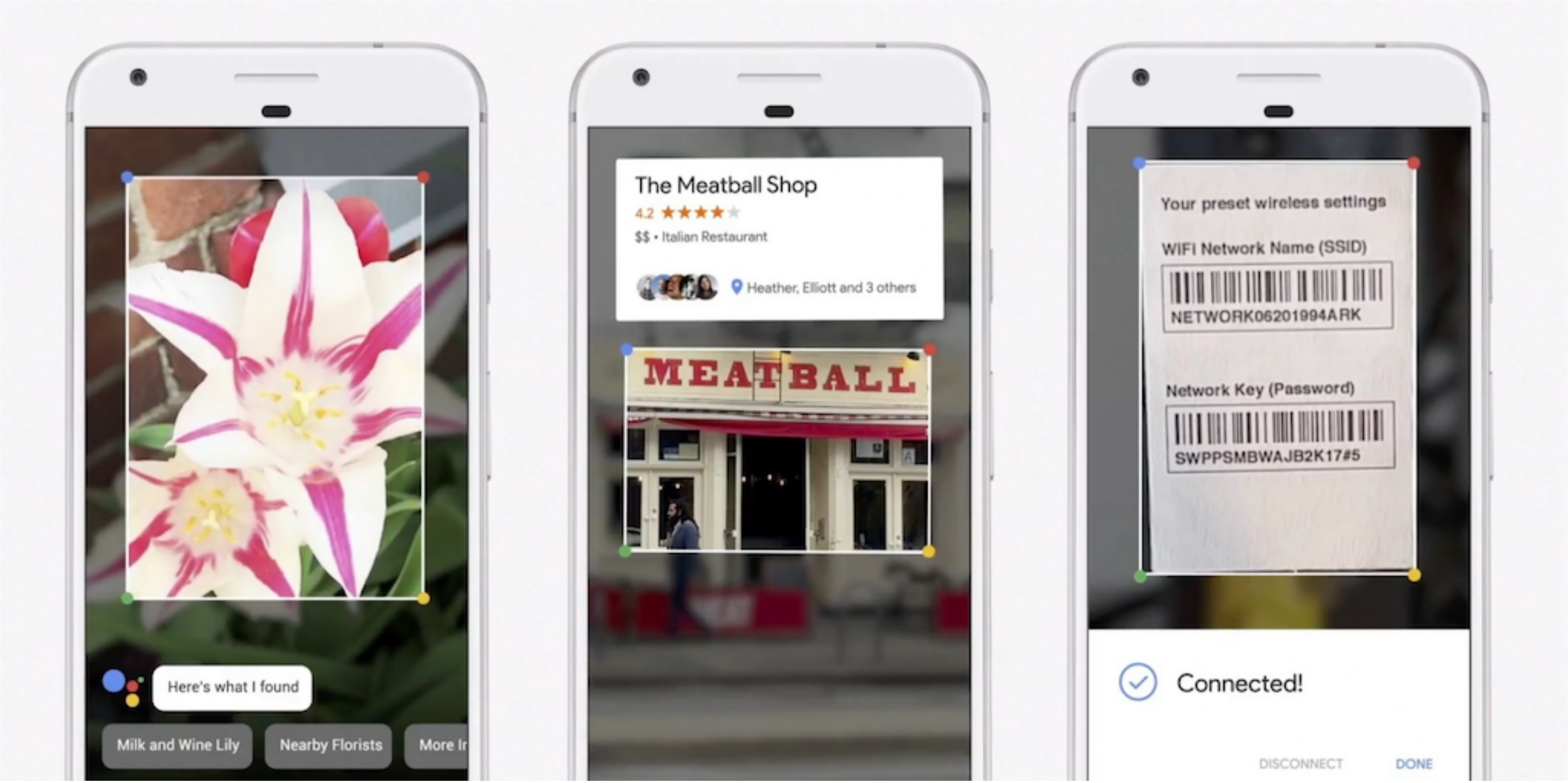
Self-Driving Cars

[Google]

AR / VR

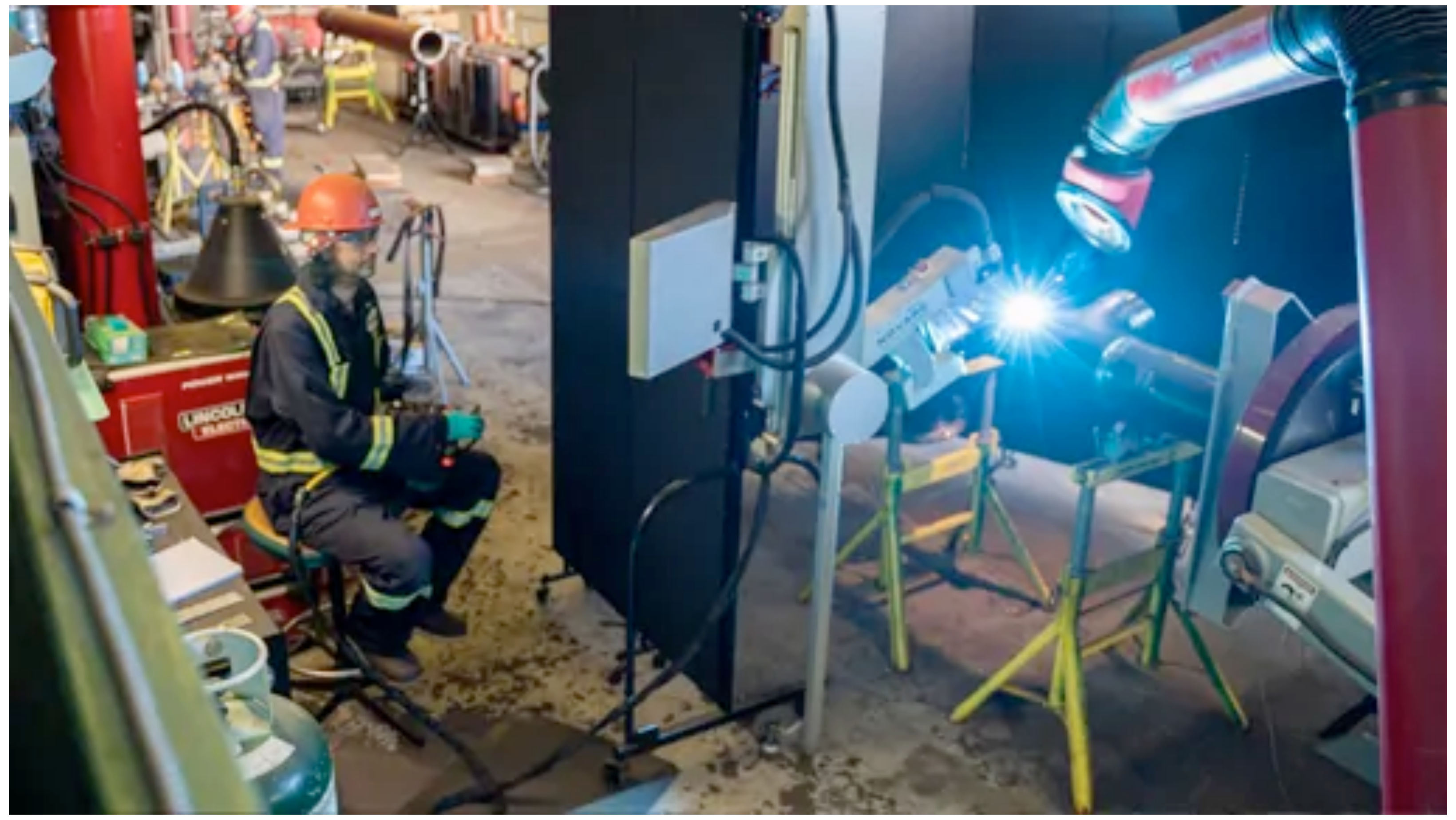
[Microsoft HoloLens]

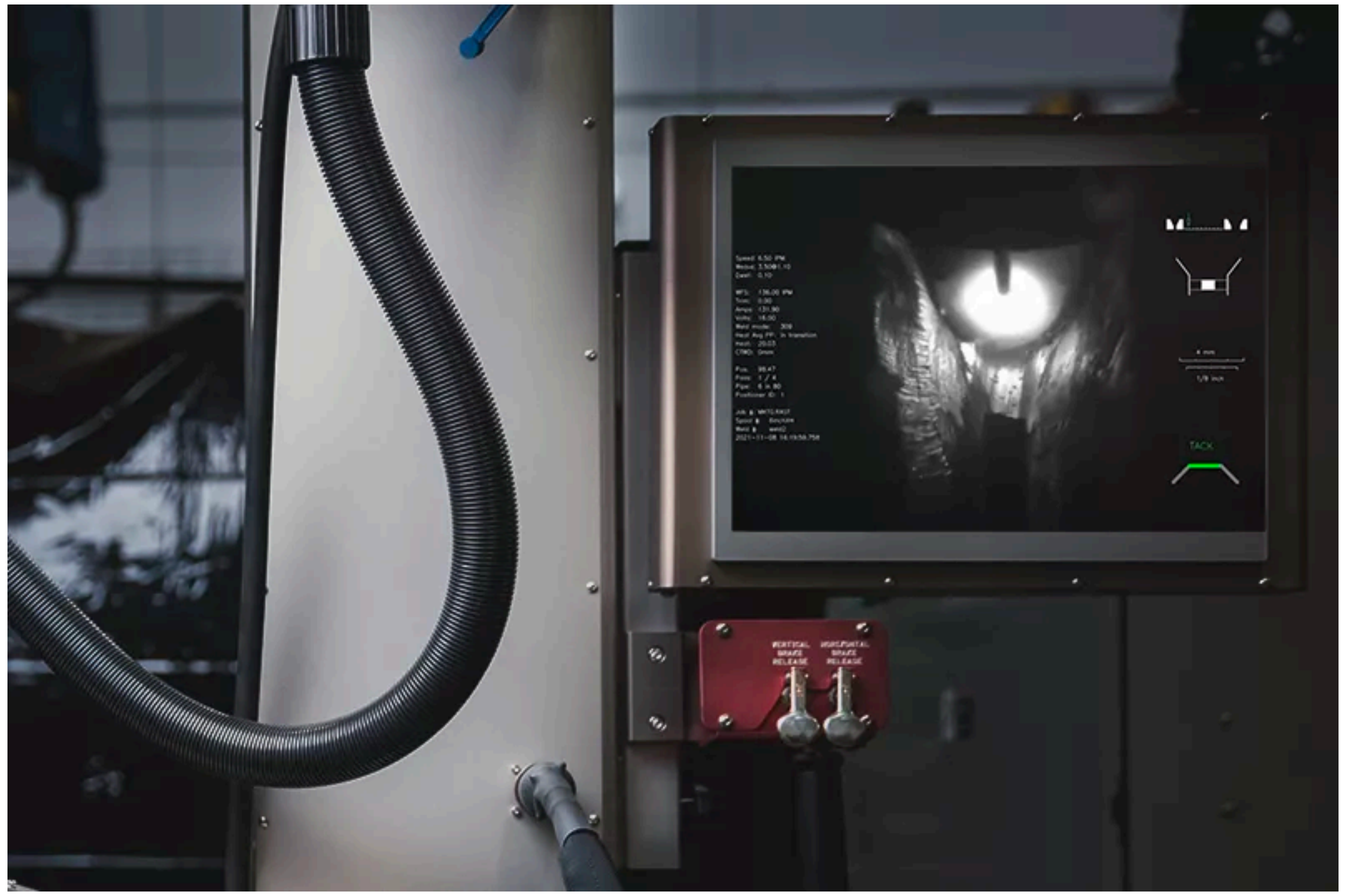
Mobile Apps



[Google Lens]

Industrial

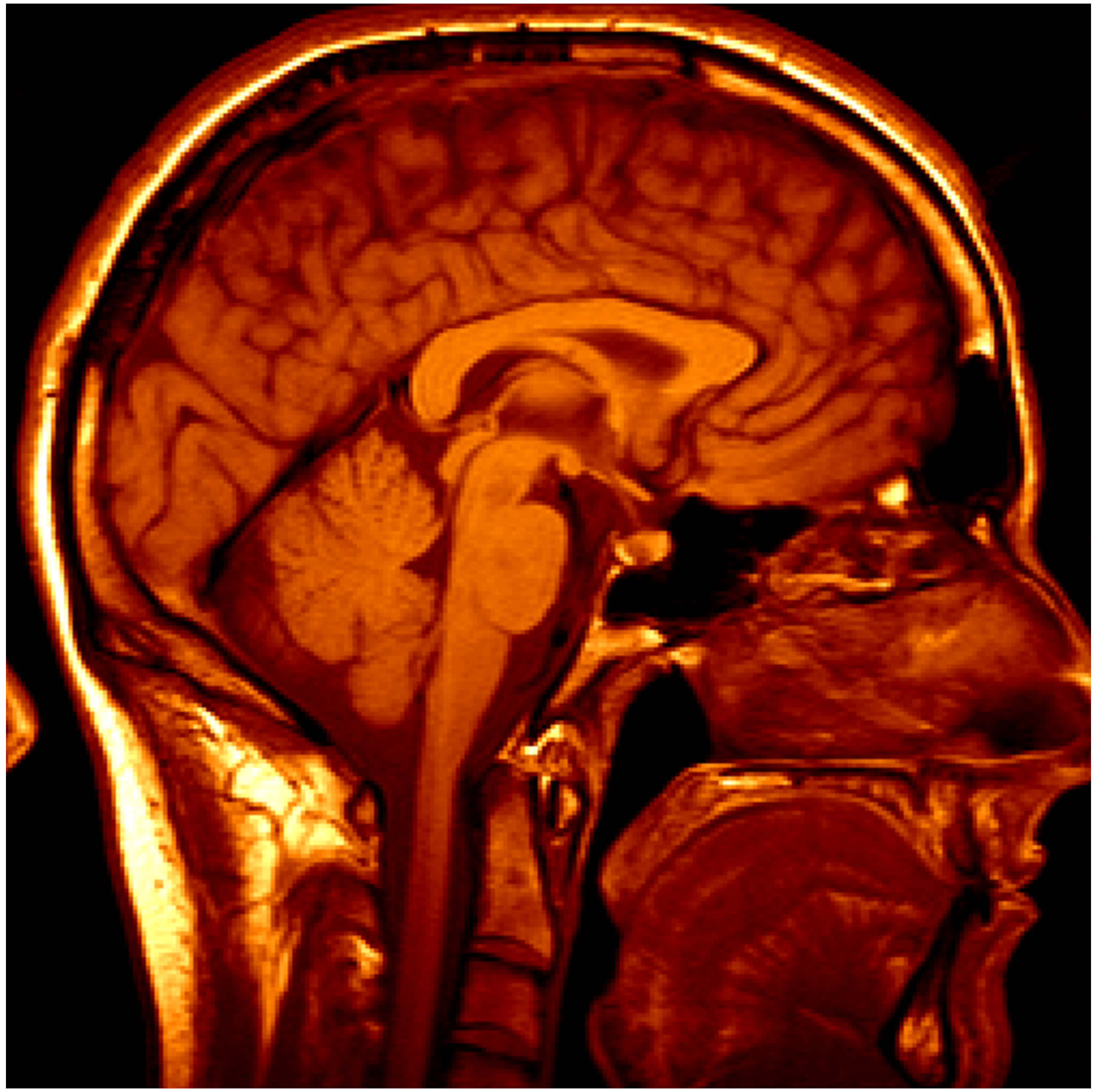




Machine Vision controlled welding robotics

NOVARC
TECHNOLOGIES

Medicine

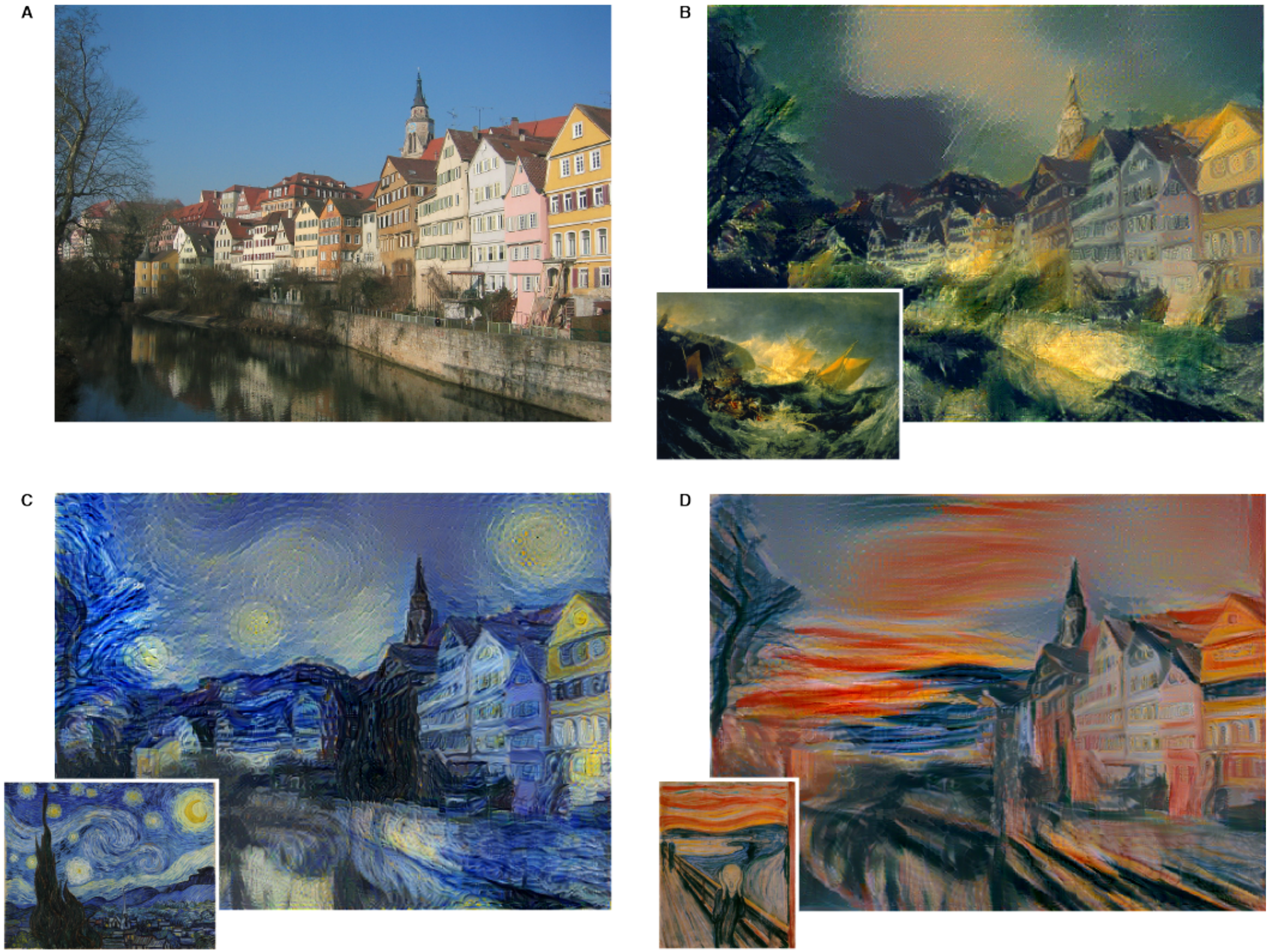


3D imaging
MRI, CT

Image guided surgery
[Grimson et al., MIT]

Slide Credit: James Hays (GA Tech)

Art



[Gatys, Ecker, Bethge 2015]

Art

TEXT DESCRIPTION

An astronaut Teddy bears A bowl of soup

riding a horse lounging in a tropical resort in space playing basketball with cats in space

in a photorealistic style in the style of Andy Warhol as a pencil drawing

→

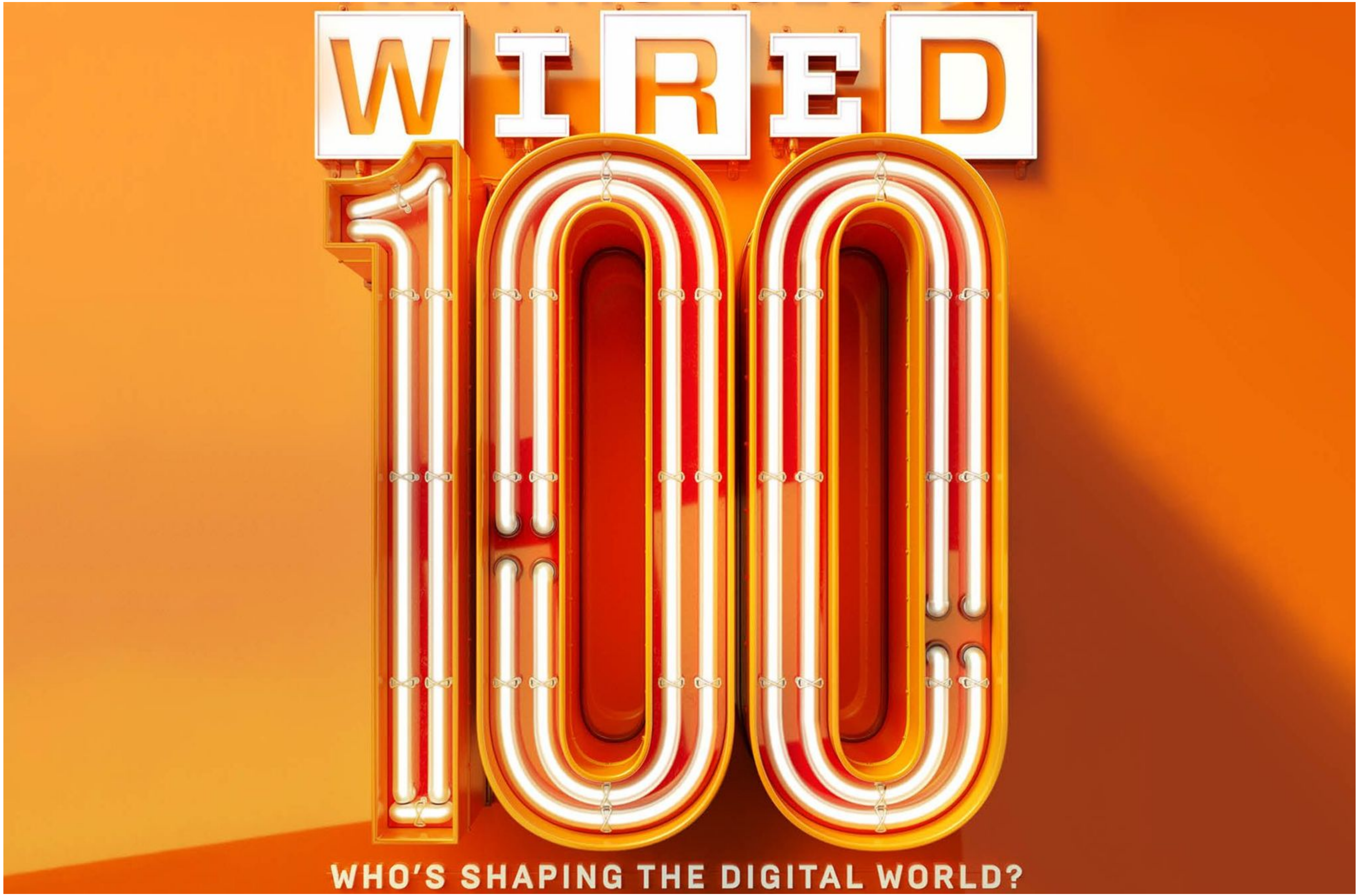
DALL·E 2

[Dall-E v2]

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


WIRED
100
WHO'S SHAPING THE DIGITAL WORLD?

Wired's 100 **Most Influential People** in the World

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

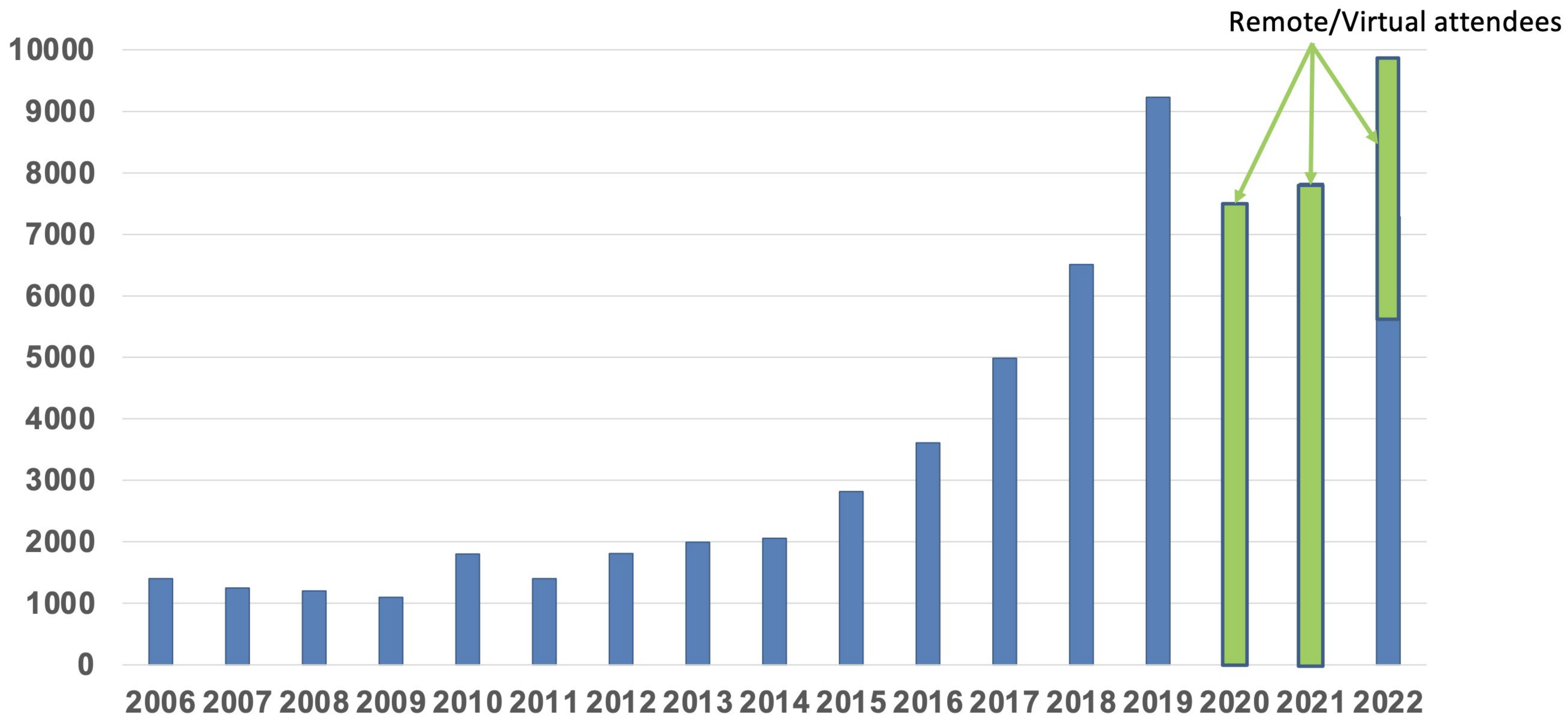
Branson saw his personal fortune grow £550 million when Alaska Air bought Virgin America for $2.6 billion in April. He is pressing on with civilian space travel with Virgin Galactic.

61. Taylor Swift

Entertainer, Los Angeles

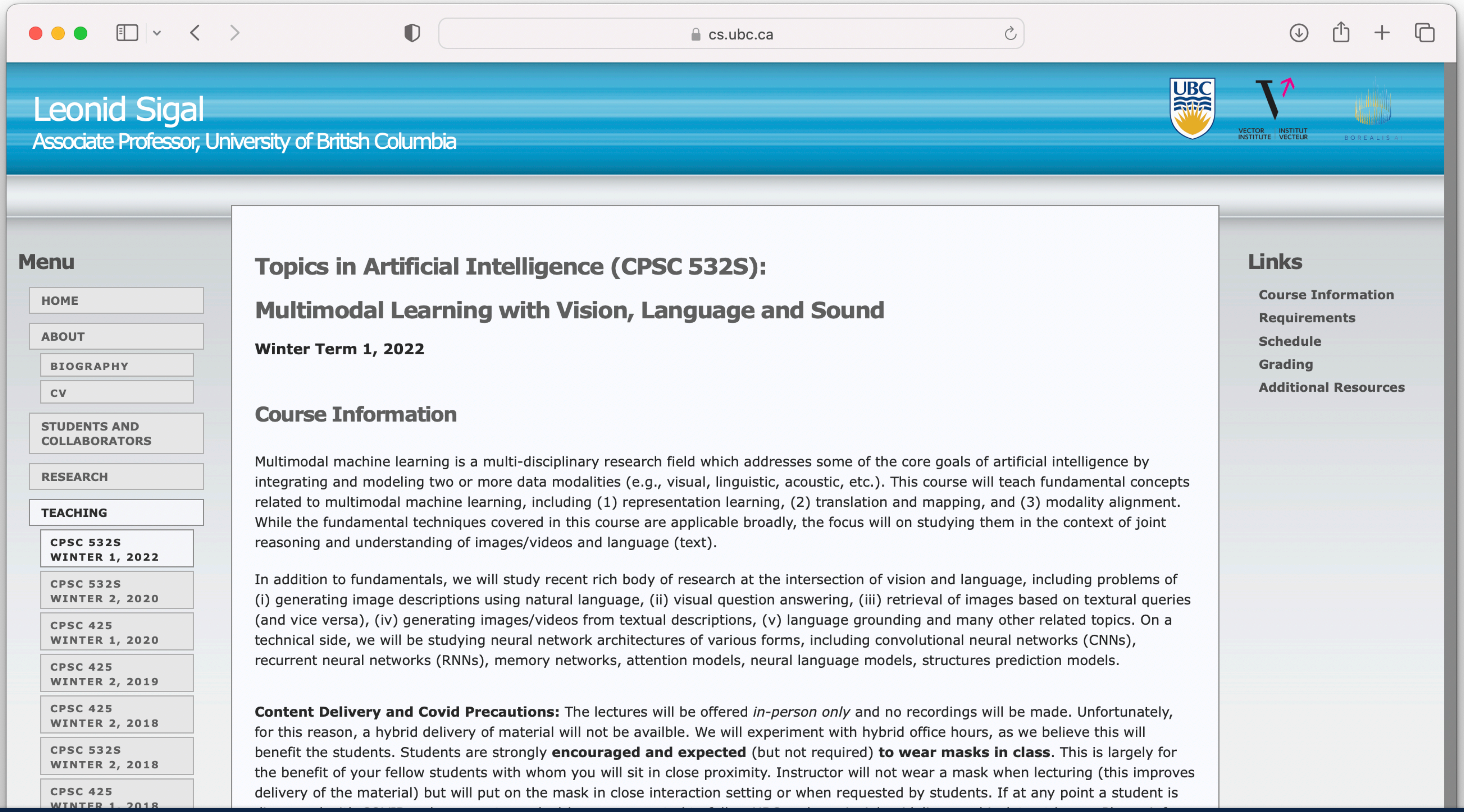

CVPR
JUNE
19-24
2022
NEW ORLEANS
LOUISIANA

CVPR **Attendance**




VANCOUVER
CVPR 2023
June 17-23, 2023

Course **Schedule**



- Schedule, Assignments
- Lecture Slides and Notes
- Readings
- Course Information (public)

https://www.cs.ubc.ca/~lsigal/teaching22_Term2.html

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

<u>Previously taught by:</u>

— 2022-2023 Term 1 by **Matthew Brown**
— 2021-2022 Term 1 & 2 by **Jim Little**
— 2020-2021 Term 1 by **Leonid Sigal**
— 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**

How to **Learn** from the **Course**?

— 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.

— 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: 45%

6 graded and 1 ungraded (optional) assignment

Midterm Exam (October 19th): 15%

Final Exam (TBD): 30%

Clicker Questions

Bring your i>Clickers to class

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

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 (done individually)

There will be **7 assignments** in total (6 marked)

— Approximately 1 every 2 weeks

— 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 45% to your final score

Midterm Exam

Scheduled for **February 27th** (right after the midterm 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 15% to your final score

Final Exam

The Final exam is held during the regular examination period, **April 17-28, 2023**, 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 <u>30%</u> to your final score

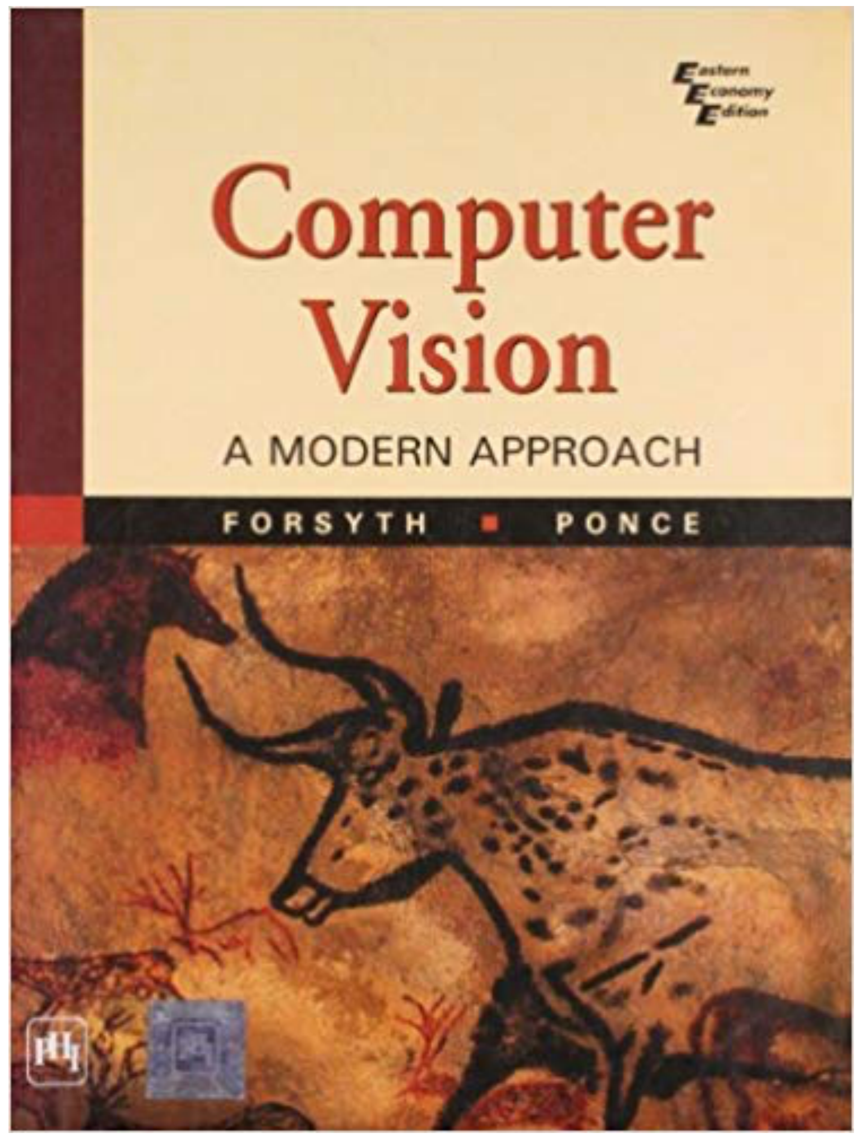
Grading issues & **Academic** Misconduct

Strict policy: Grading mistakes happen, it's just a nature of life. If you see an issue with your grade, you have 1 week from the release of any assignment grade to bring a specific issue to our attention.

Academic Misconduct: Please don't do it. Trust me it is not worth it.

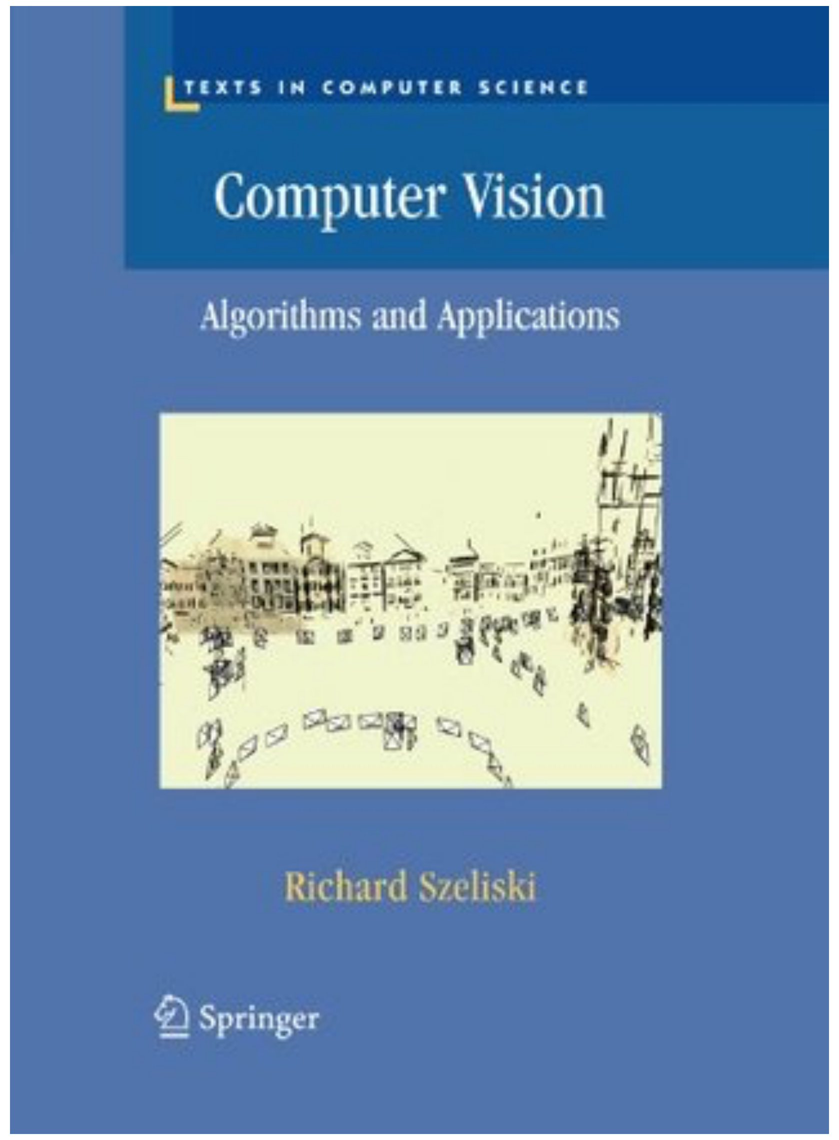
Textbooks

The course uses the following textbooks, which are recommended (but **not required**):



Computer Vision: A Modern Approach (2nd ed)

By: D. Forsyth & J. Ponce
Publisher: Pearson 2012



Computer Vision: Algorithms and Applications (2nd ed)

By: R. Szeliski
Publisher: Springer 2022

https://szeliski.org/Book/

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: Szeliski Chapter 2, Forsyth & Ponce (2nd ed.) 1.1.1 — 1.1.3

Reminders:

— Start working on **Assignment 0** (ungraded) suggest complete by Jan 16

— **[optional]** Watch TED talk by Prof. Fei-Fei Li

https://www.youtube.com/watch?v=40riCqvRoMs