

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

Instructor: Leonid Sigal



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

Locations: Friedman (FRDM), Room 153



Course logistics

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

Instructor: Leonid Sigal



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

Locations: Friedman (FRDM), Room 153



Software Engineer 1999 - 2001







Software Engineer 1999 - 2001

COGNEX



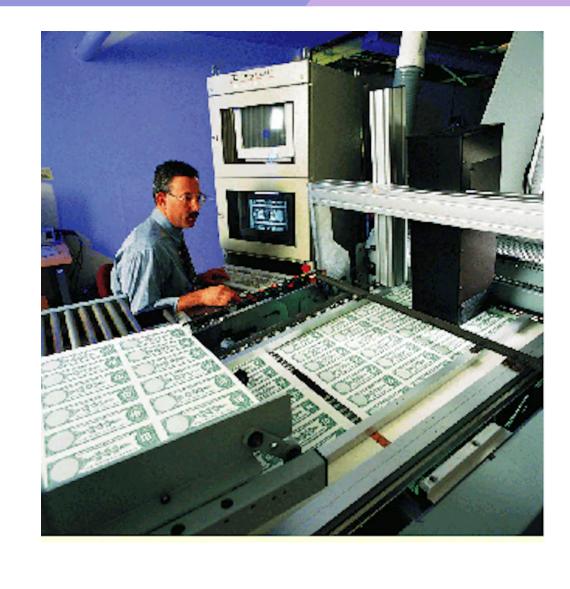




Software Engineer 1999 - 2001

COGNEX



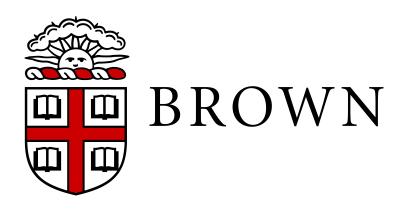




PhD, MSc 2001 - 2008

Software Engineer 1999 - 2001







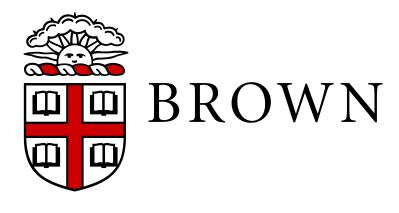


Postdoctoral Researcher 2007 - 2009

PhD, MSc 2001 - 2008

Software Engineer 1999 - 2001









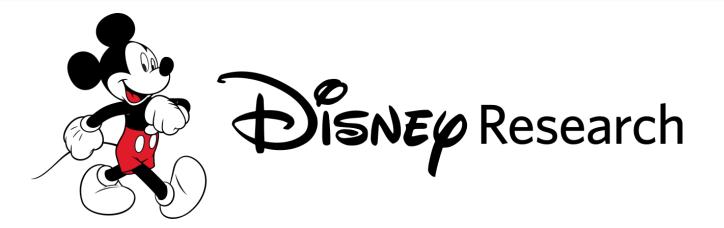


Senior Research Scientist 2009 - 2017

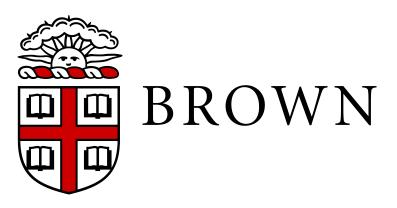
Postdoctoral Researcher 2007 - 2009

PhD, MSc 2001 - 2008

Software Engineer 1999 - 2001













Associate Professor 2017 -

Senior Research Scientist 2009 - 2017

Postdoctoral Researcher 2007 - 2009

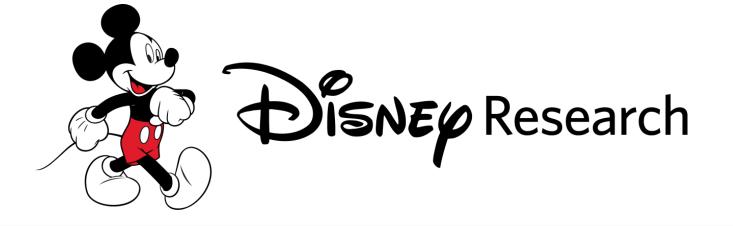
PhD, MSc 2001 - 2008

Software Engineer 1999 - 2001

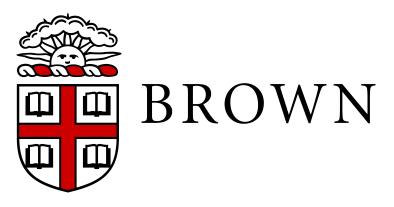




THE UNIVERSITY **OF BRITISH COLUMBIA**













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

Associate Professor 2017 -

Senior Research Scientist 2009 - 2017

Postdoctoral Researcher 2007 - 2009

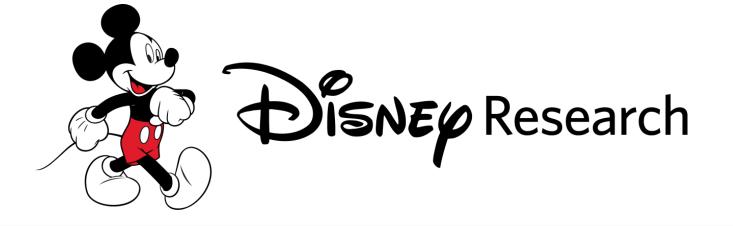
PhD, MSc 2001 - 2008

Software Engineer 1999 - 2001

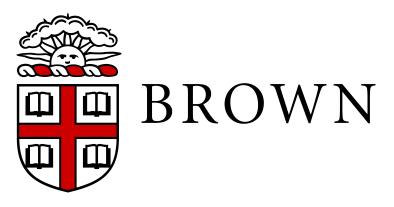




THE UNIVERSITY **OF BRITISH COLUMBIA**













Course logistics

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

Instructor: Leonid Sigal



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

Locations: Friedman (FRDM), Room 153



Course logistics

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

Instructor: Leonid Sigal

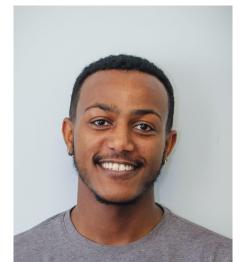


E-mail: lsigal@cs.ubc.ca Office: ICICS 119



Locations: Friedman (FRDM), Room 153

TAs: Bereket Guta



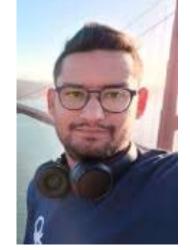
bguta@cs.ubc.ca

Aditya Chinchure



aditya10@cs.ubc.ca

Gaurav Bhatt



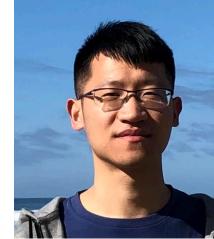
gbhatt@cs.ubc.ca

Rayat Hossain



rayat137@cs.ubc.ca

Bicheng Xu

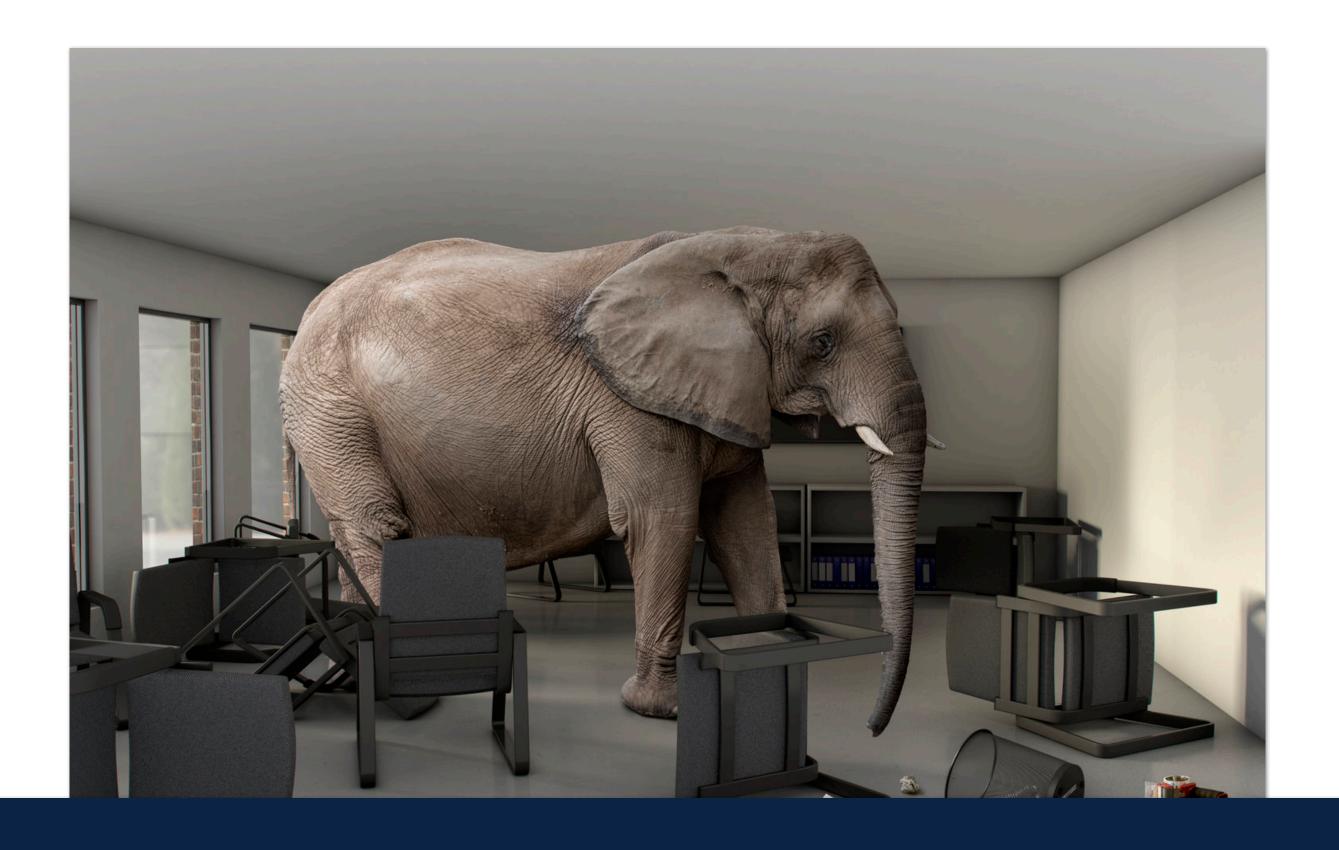


bichengx@cs.ubc.ca





Registration



There is high likelihood everything will work out

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



Course Webpage

Leonid Sigal

Associate Professor, University of British Columbia

Menu

HOME
ABOUT
BIOGRAPHY
cv
STUDENTS AND
COLLABORATORS
RESEARCH
TEACHING
CPSC 532S
WINTER 1, 2022
CPSC 532S WINTER 2, 2020
WINTER 2, 2020
CPSC 425
WINTER 1, 2020
CPSC 425
WINTER 2, 2019
CPSC 425
WINTER 2, 2018
-
CPSC 532S
WINTER 2, 2018
CPSC 425
WINTER 1 2018

Topics in Artificial Intelligence (CPSC 532S):

Multimodal Learning with Vision, Language and Sound

Winter Term 1, 2022

Course Information

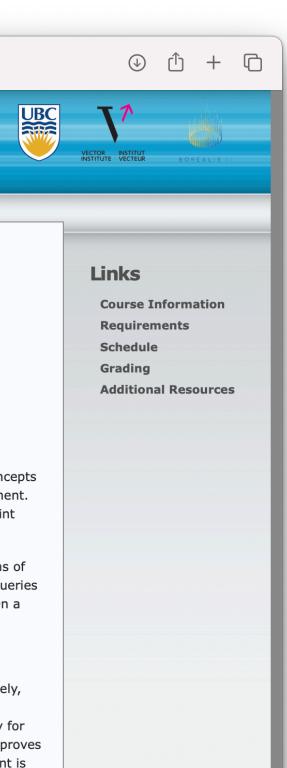
Multimodal machine learning is a multi-disciplinary research field which addresses some of the core goals of artificial intelligence by integrating and modeling two or more data modalities (e.g., visual, linguistic, acoustic, etc.). This course will teach fundamental concepts related to multimodal machine learning, including (1) representation learning, (2) translation and mapping, and (3) modality alignment. While the fundamental techniques covered in this course are applicable broadly, the focus will on studying them in the context of joint reasoning and understanding of images/videos and language (text).

🔒 cs.ubc.ca

In addition to fundamentals, we will study recent rich body of research at the intersection of vision and language, including problems of (i) generating image descriptions using natural language, (ii) visual question answering, (iii) retrieval of images based on textural queries (and vice versa), (iv) generating images/videos from textual descriptions, (v) language grounding and many other related topics. On a technical side, we will be studying neural network architectures of various forms, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), memory networks, attention models, neural language models, structures prediction models.

Content Delivery and Covid Precautions: The lectures will be offered *in-person only* and no recordings will be made. Unfortunately, for this reason, a hybrid delivery of material will not be available. We will experiment with hybrid office hours, as we believe this will benefit the students. Students are strongly **encouraged and expected** (but not required) **to wear masks in class**. This is largely for the benefit of your fellow students with whom you will sit in close proximity. Instructor will not wear a mask when lecturing (this improves delivery of the material) but will put on the mask in close interaction setting or when requested by students. If at any point a student is

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



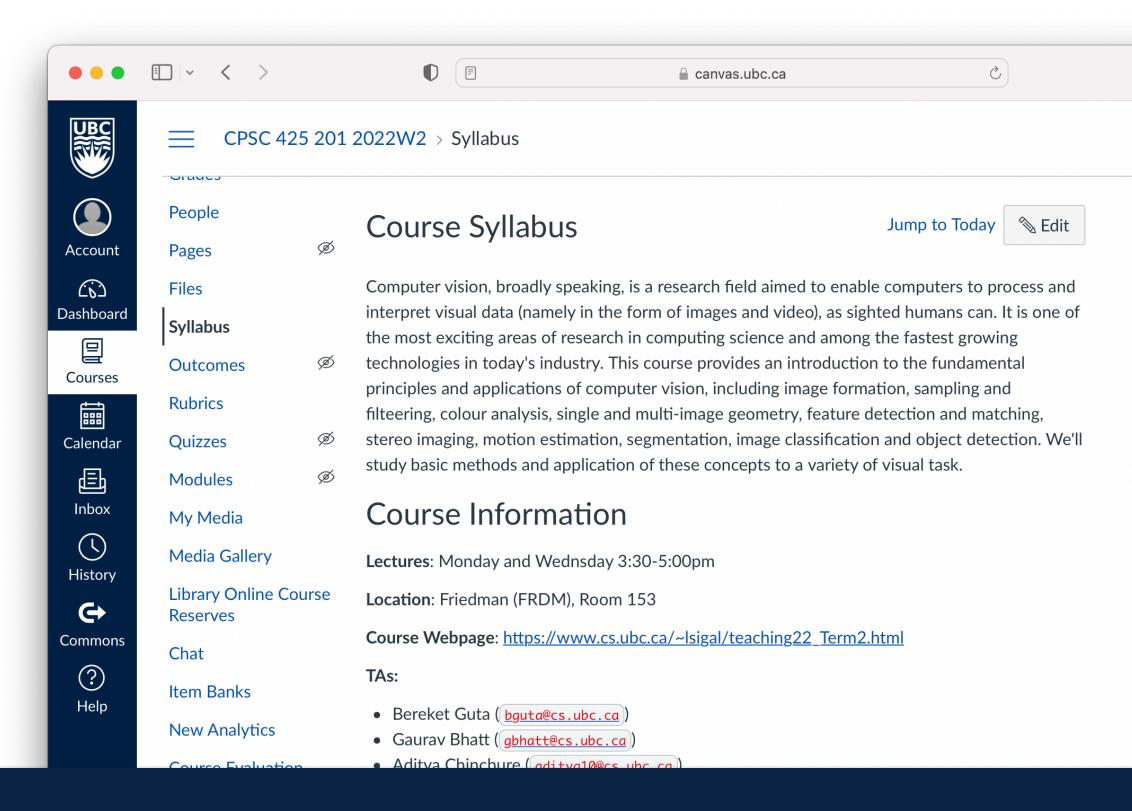
C

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



;)

Canvas



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

Total

				Û	+	ſĊ				
6d Student View										
< January 2023 >										
		January 2025 /								
25	26	27	28	29	30	31				
1	2	3	4	5	6	7				
8	9	10	11	12	13	14				
15	16	17	18	19	20	21				
22	23	24	25	26	27	28				
29	30	31	1	2	3	4				
Assignments are weighted by group: Group Weight										
Assignments				45%						
Midterm 15%										
Qui	zes			10%						
Fina	al			10% 30%						

100%

*

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



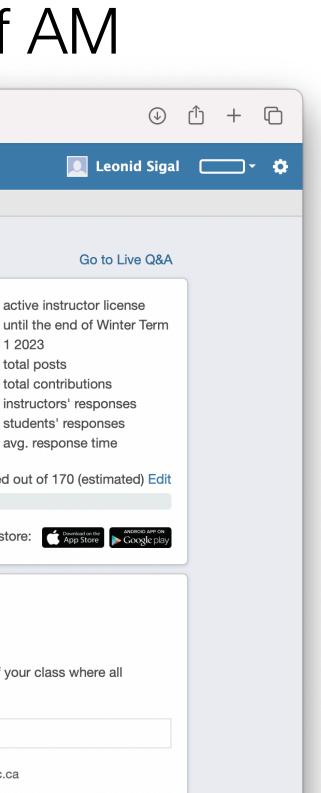
Piazza

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

48 students were enrolled as of AM

••• • • •	0			🔒 piazz	a.com		Ś	
ριαΖΖα CPSC 425 201 20	022W2 ⁵ - Setup	<u>Q & A</u>	Resources S	Statistics -	Manage C	lass		
LIVE Q&A Drafts hw0 Unread Updated Unresolved Following New Post Q Search or add a post	hw1 hw2 hw3		w5 midterm	final_exam	logistics ance	other Updated 10 secon	nds ago. Reload	
Show Actions PINNED Private Search for Teammates! TODAY	★ 1/8/23 1 ■			posts wered quest wered follov			license status 5 5 0	acti unti 1 20 tota tota inst
 Private Introduce Piazza to your stu Private Get familiar with Piazza Private Tips & Tricks for a successf 	10:44 AM 1 10:44 AM 10:44 AM 10:44 AM	5	Student Enrollm	ent			0 1 sec 0 enroll	0
 Welcome to Piazza! Piazza is a Q&A platform designed to get you great answers from classmates and instructors fast. We've put together thi 	1 🗐		Download us in the app sto Share Your Class					o store
		ŀ	Professors appreciate Piazza best when they see how it is being used. Allow colleagues to view your class through a demo link - a restricted, read only version of you students' names are anonymized and all student information hidden. https://piazza.com/demo_login?nid=lcnq4nb0j0k1ro&auth=5e7487d					
			nttps://piazza				n=5e7487d	oc.ca





- 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**

TAs: Bereket Guta

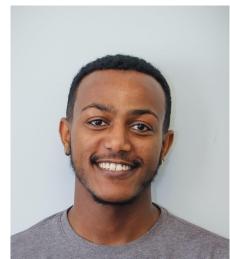
Instructor: Leonid Sigal



Monday 2-3pm (in-person)

Friday 12-1pm, Hybrid

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

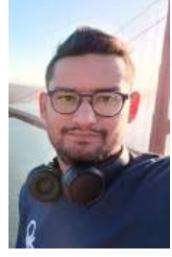


Aditya Chinchure



Tuesday noon-1pm

Gaurav Bhatt



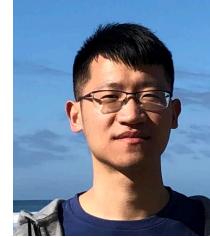
TBD

Rayat Hossain



TBD

Bicheng Xu



Friday 4-5pm (in-person)



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









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

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



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

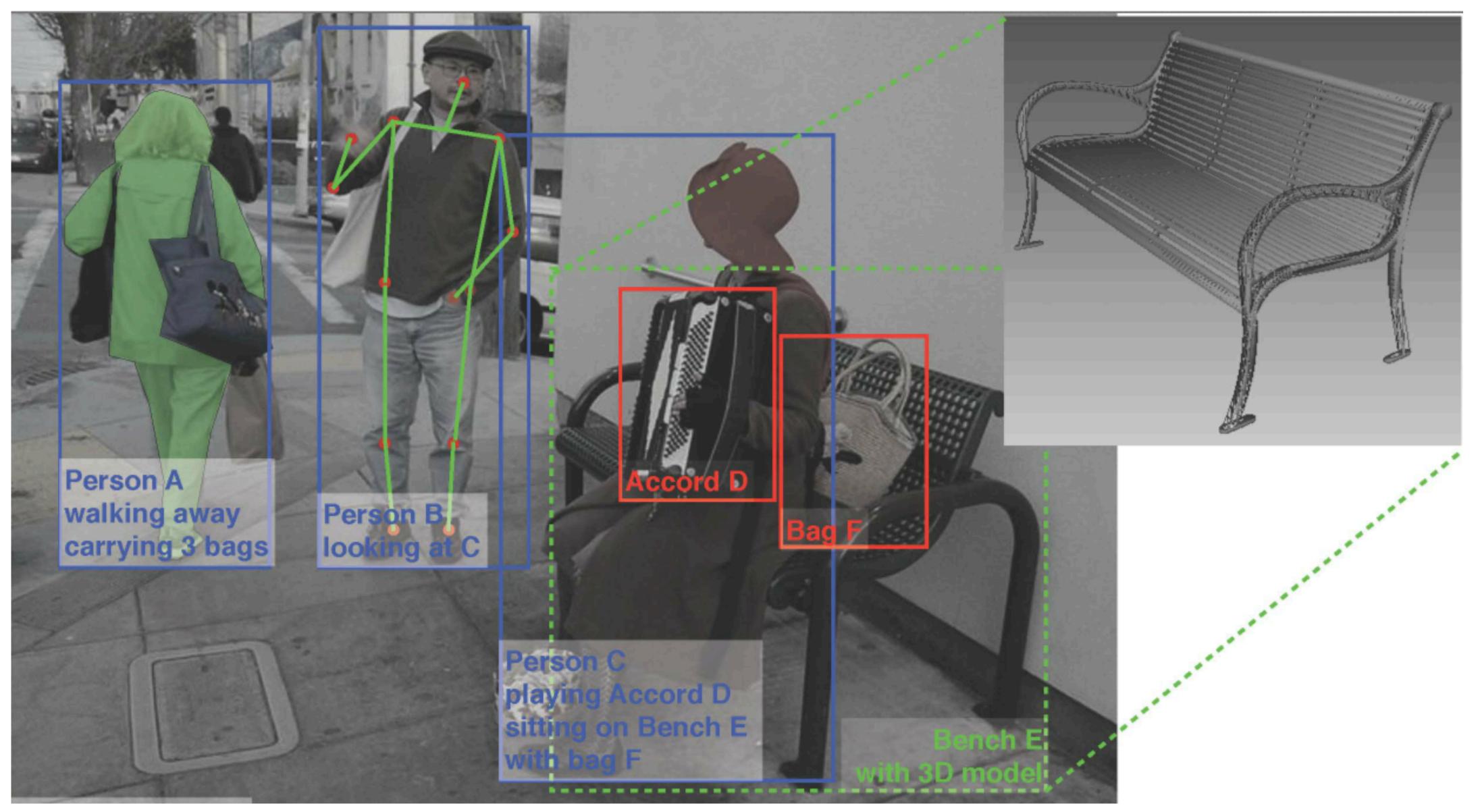


What do you see?



Slide Credit: Jitendra Malik (UC Berkeley)

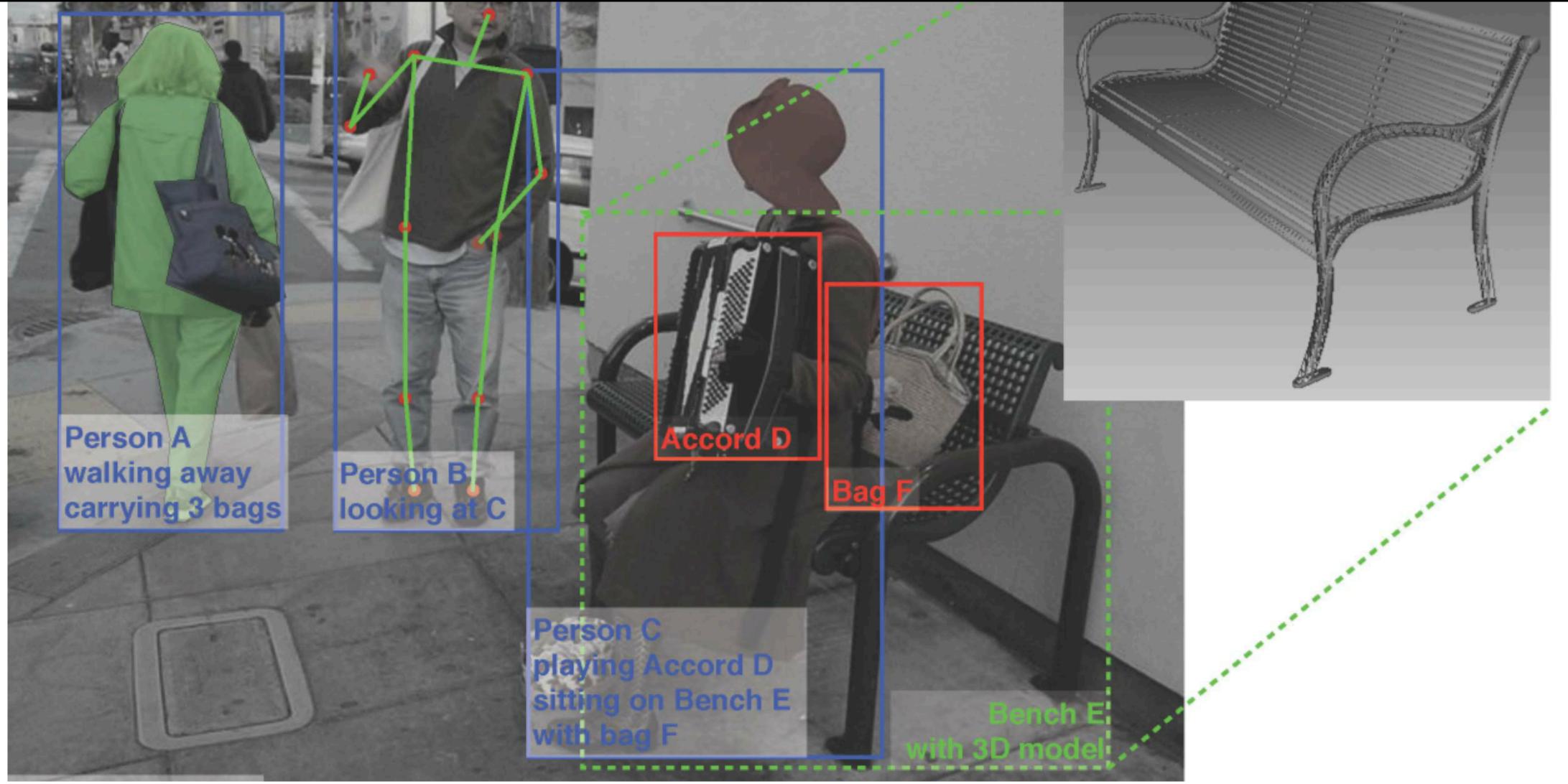
What we would like computer to infer?



Slide Credit: Jitendra Malik (UC Berkeley)

What we would like computer to infer?

Will person B put some money into person C's cup?



Slide Credit: Jitendra Malik (UC Berkeley)



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



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

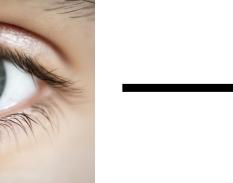
Sensing Device







Interpreting Device





Interpretation

lickr.com/photos/flamephoenix1991/8376271918



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



Interpretation

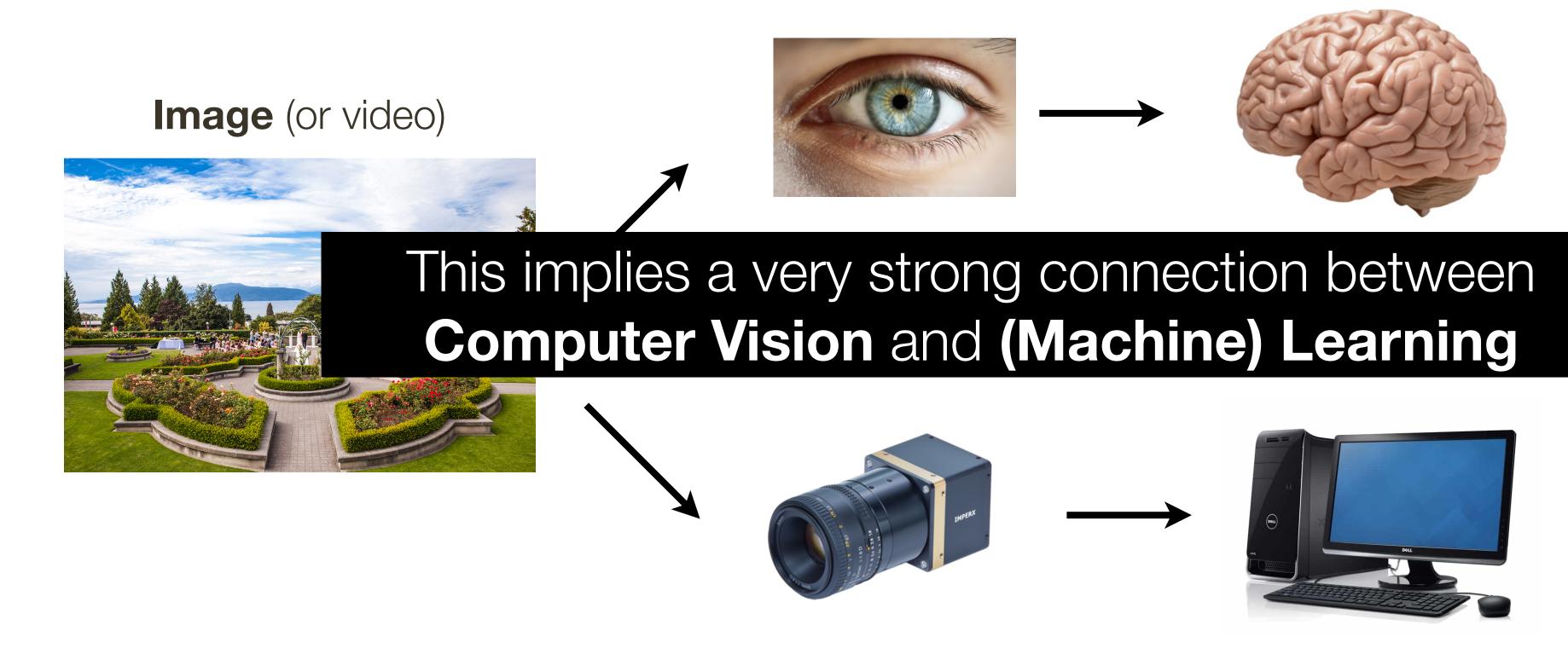
os/flamephoenix1991/8376271918





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

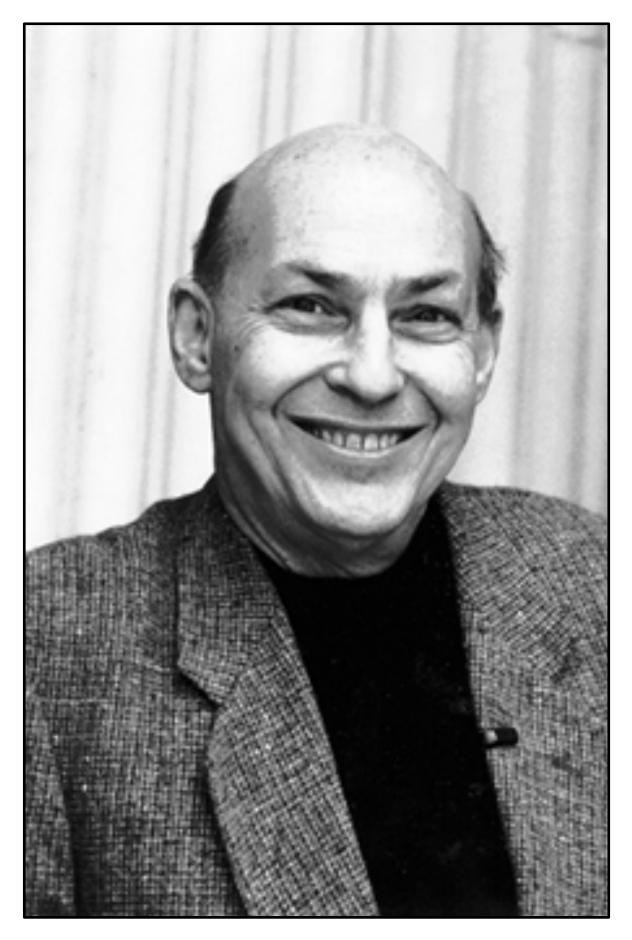
Sensing Device **Interpreting** Device



Interpretation



Computer vision ... the beginning ...



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)

Slide Credit: Devi Parikh (GA Tech)

MASSACHUSETTS INSTITUTE OF TECHNOLOGY PROJECT MA

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

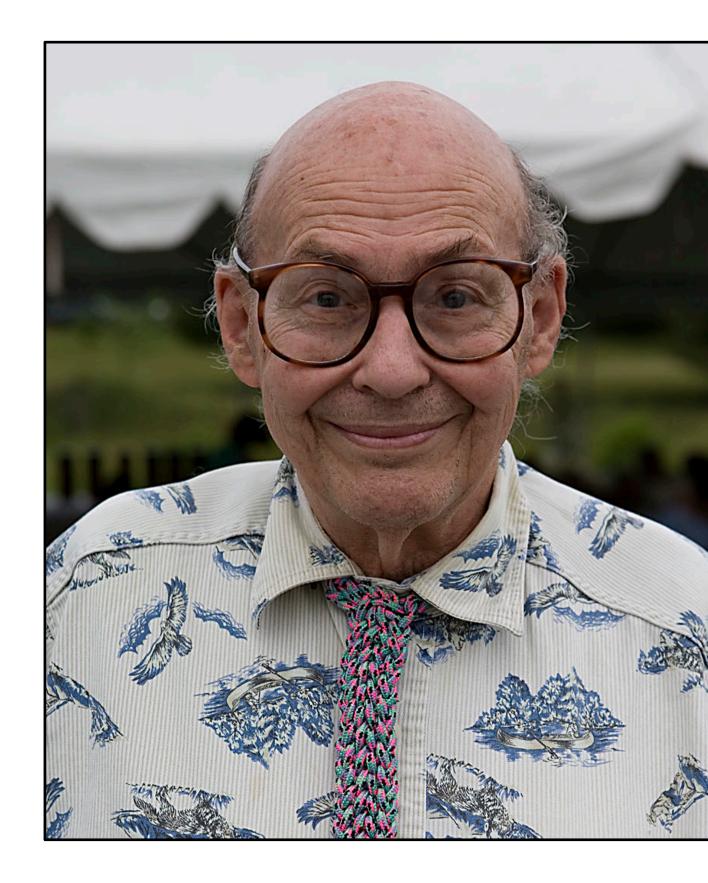
UMMER VISION PROJEC

an attempt to use our summer workers in the construction of a significant part of The particular task was chosen partly because sub-problems which will allow individuals to work independently and yes participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition"

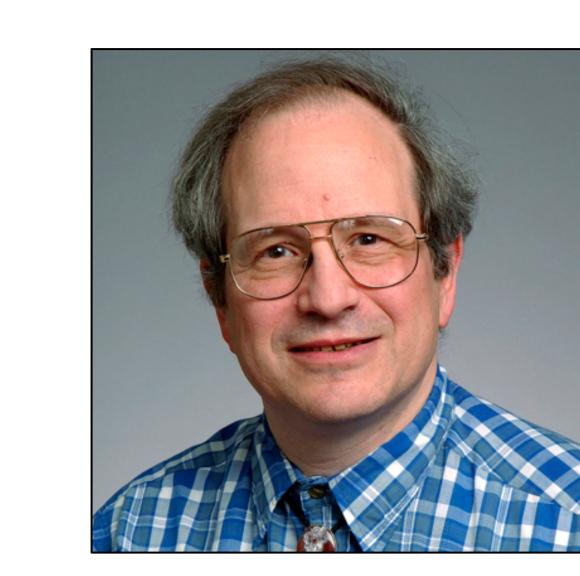
 $\dots >50$ years later



Computer vision ... the beginning ...



Slide Credit: Devi Parikh (GA Tech)



Gerald Sussman, MIT

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

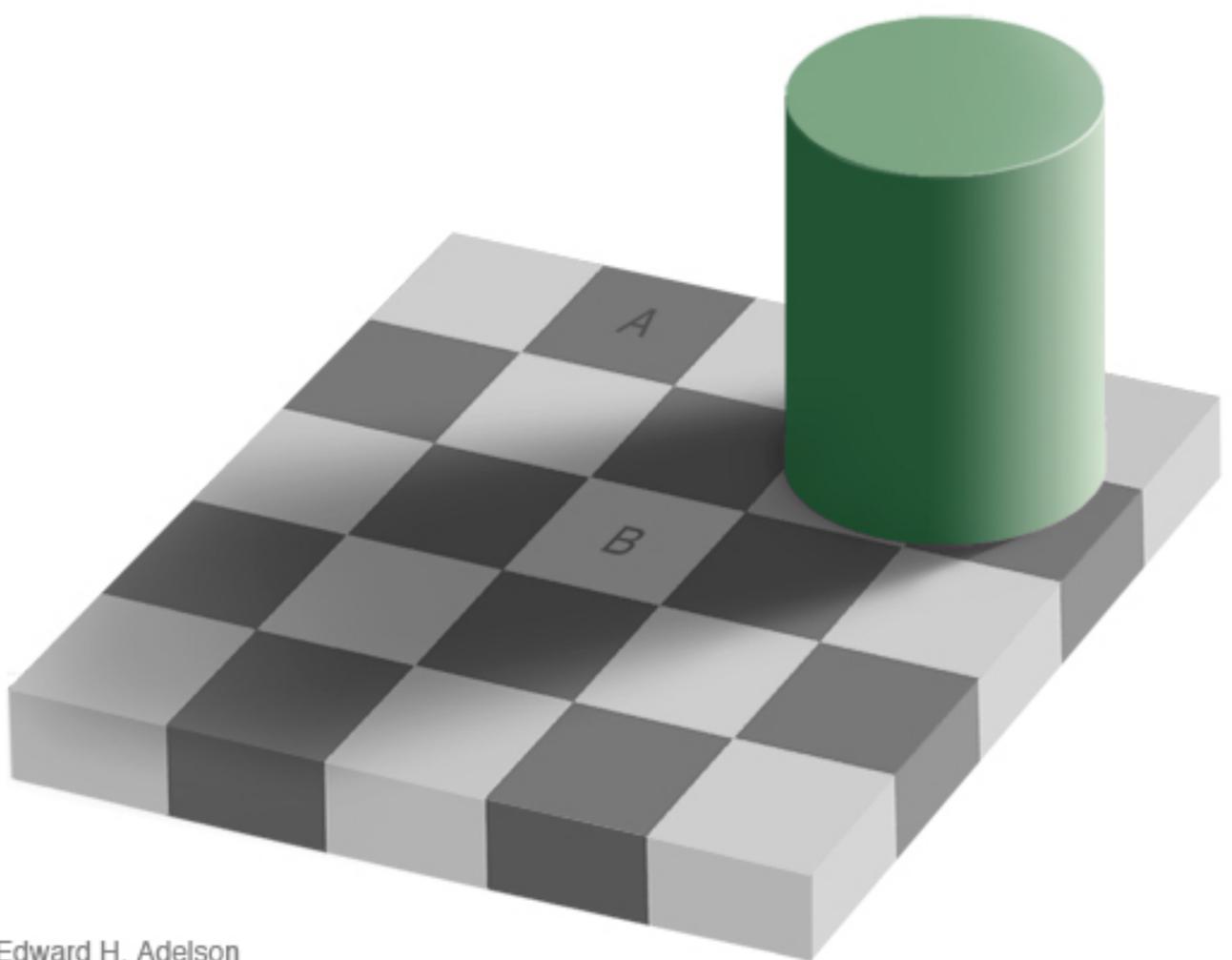
Can computers match (or beat) human vision?

• We've been at it for 50 years

Can computers match (or beat) human vision?

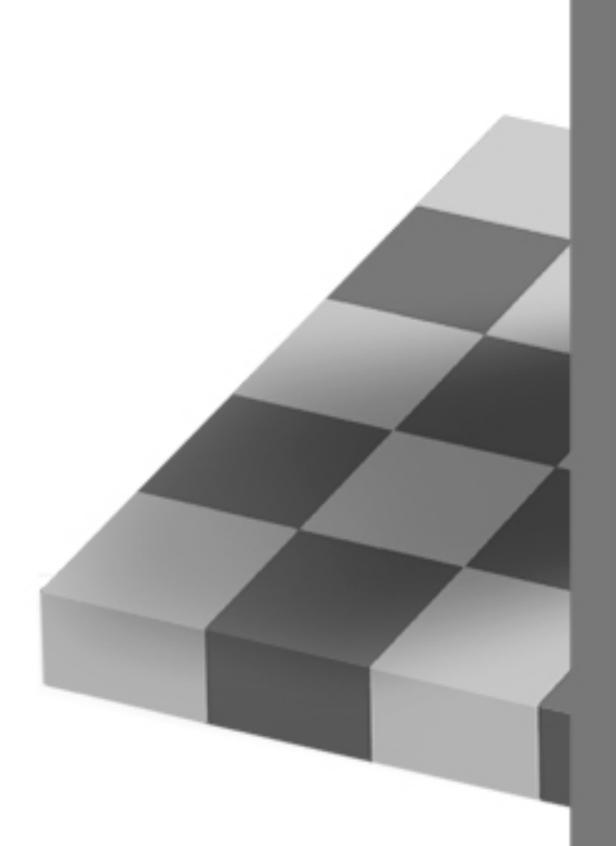
How good is human vision?

Can computers match (or beat) human vision?

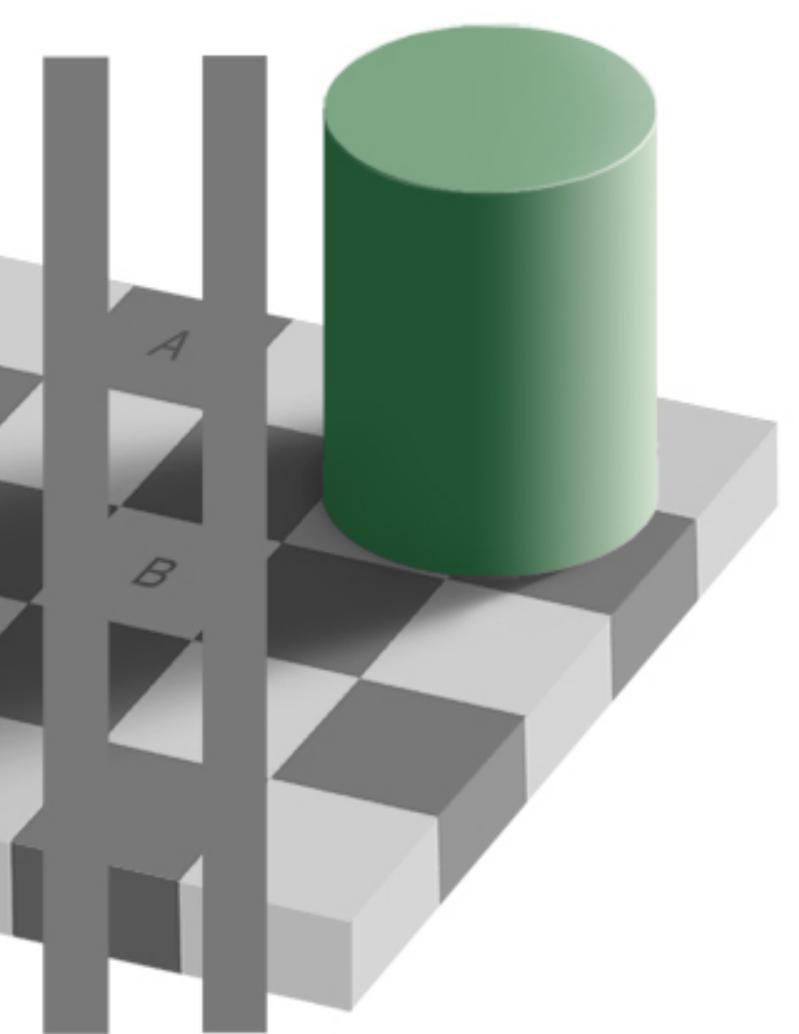


Edward H. Adelson

Can computers match (or beat) human vision?



Edward H. Adelson



Can computers match (or beat) human vision?

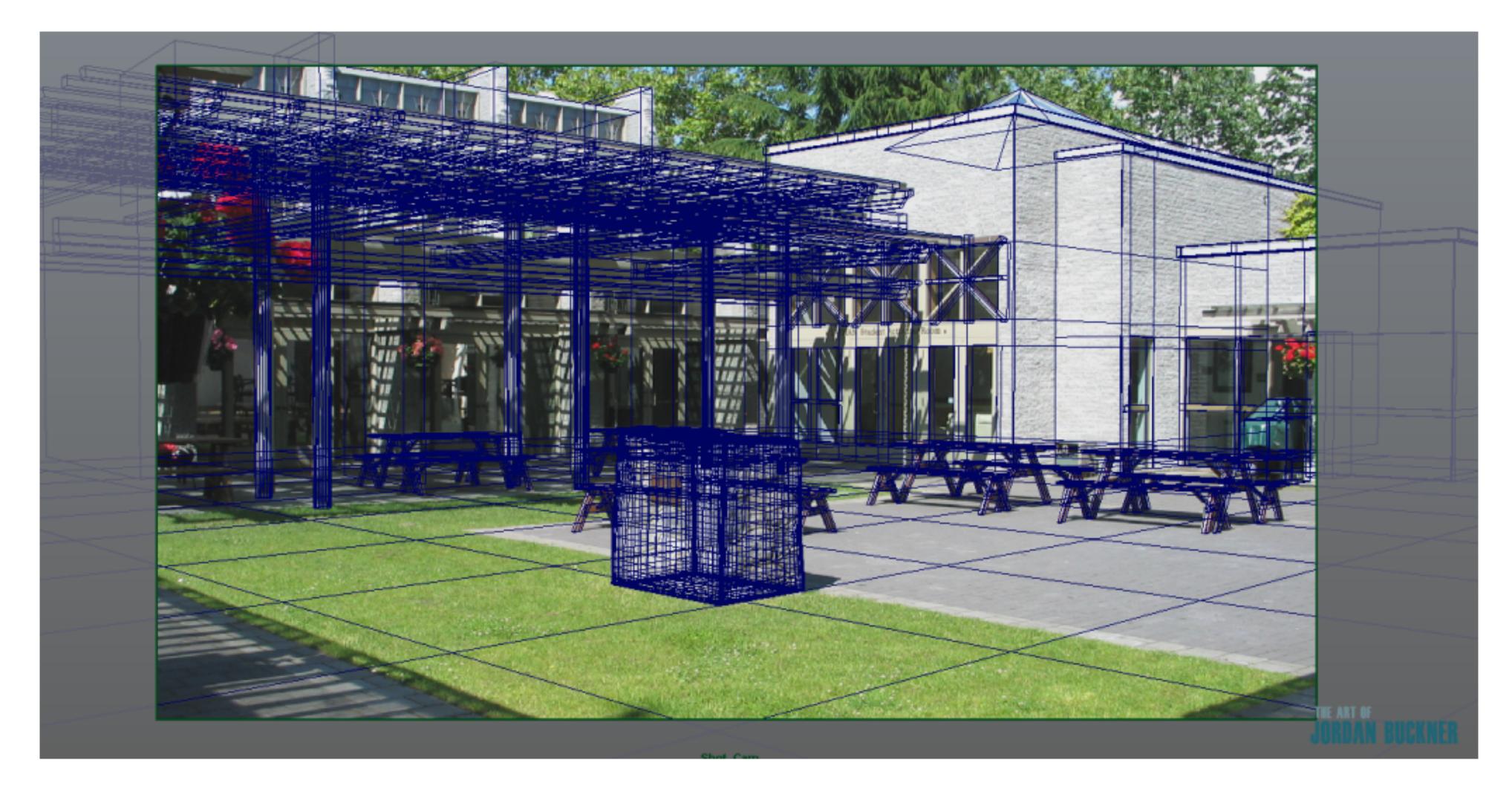
- How good is human vision?

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"





Alternative definition of computer vision "Inverse Computer Graphics"







Vision

Slide Credit: Kristen Grauman (UT Austin)



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

1. Vision for Measurement

Real-time stereo



Wang et al.

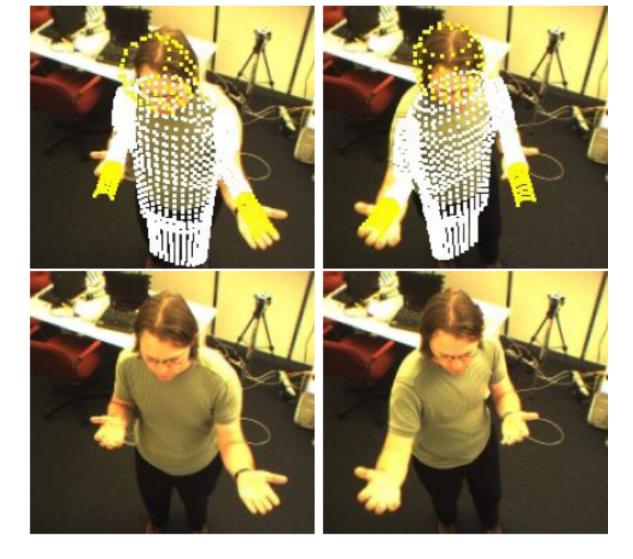
Structure from motion



Slide Credit: Kristen Grauman (UT Austin)



Snavely et al.



Demirdjian et al.

Slide Credit: Kristen Grauman (UT Austin)

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

III-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

people, scenes, and activities (*perception and interpretation*)

Slide Credit: Kristen Grauman (UT Austin)

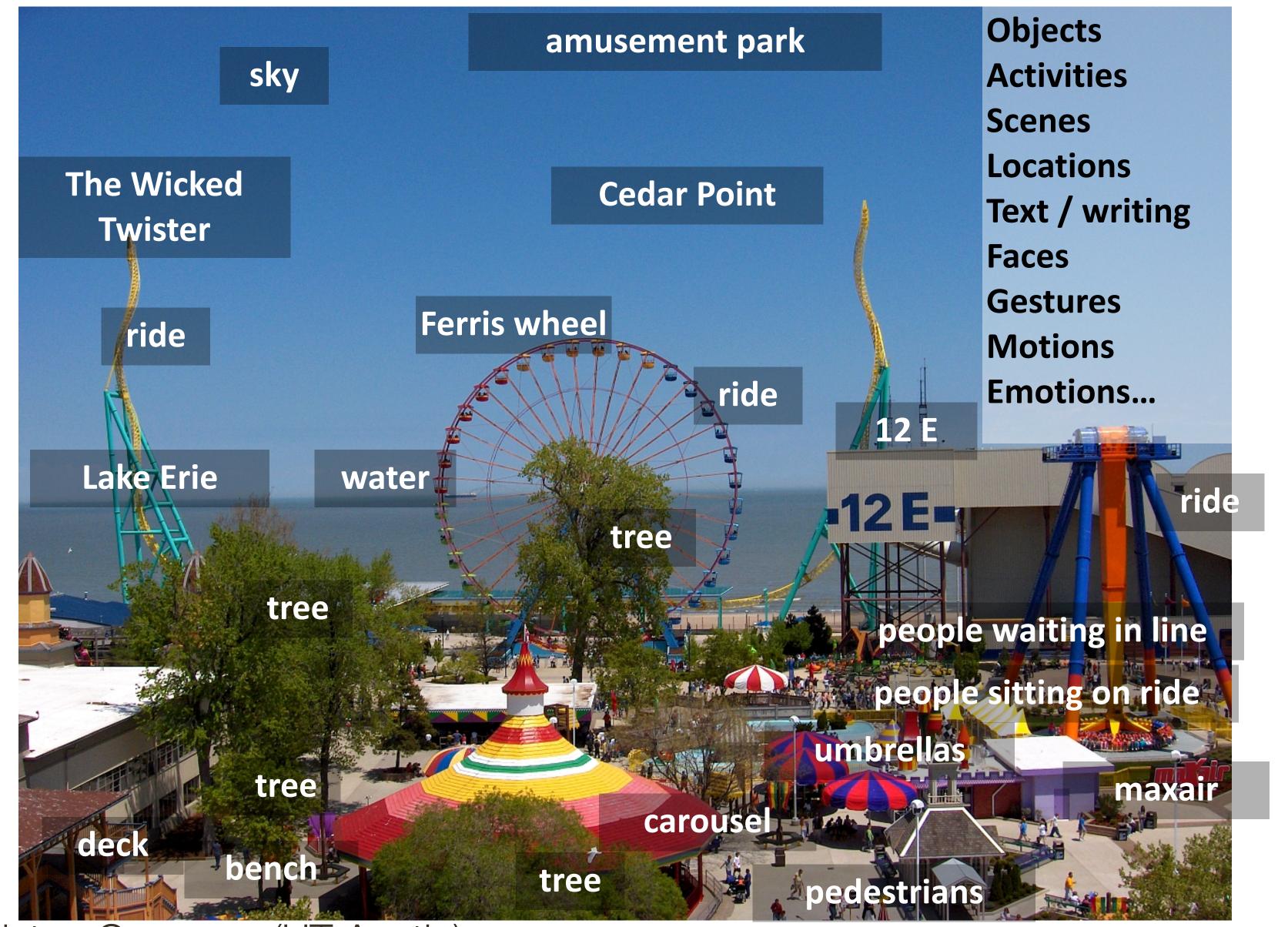
1. Computing properties of the 3D world from visual data (*measurement*)

2. Algorithms and representations to allow a machine to recognize objects,

2. Vision for Perception and Interpretation



2. Vision for Perception and Interpretation



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



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

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



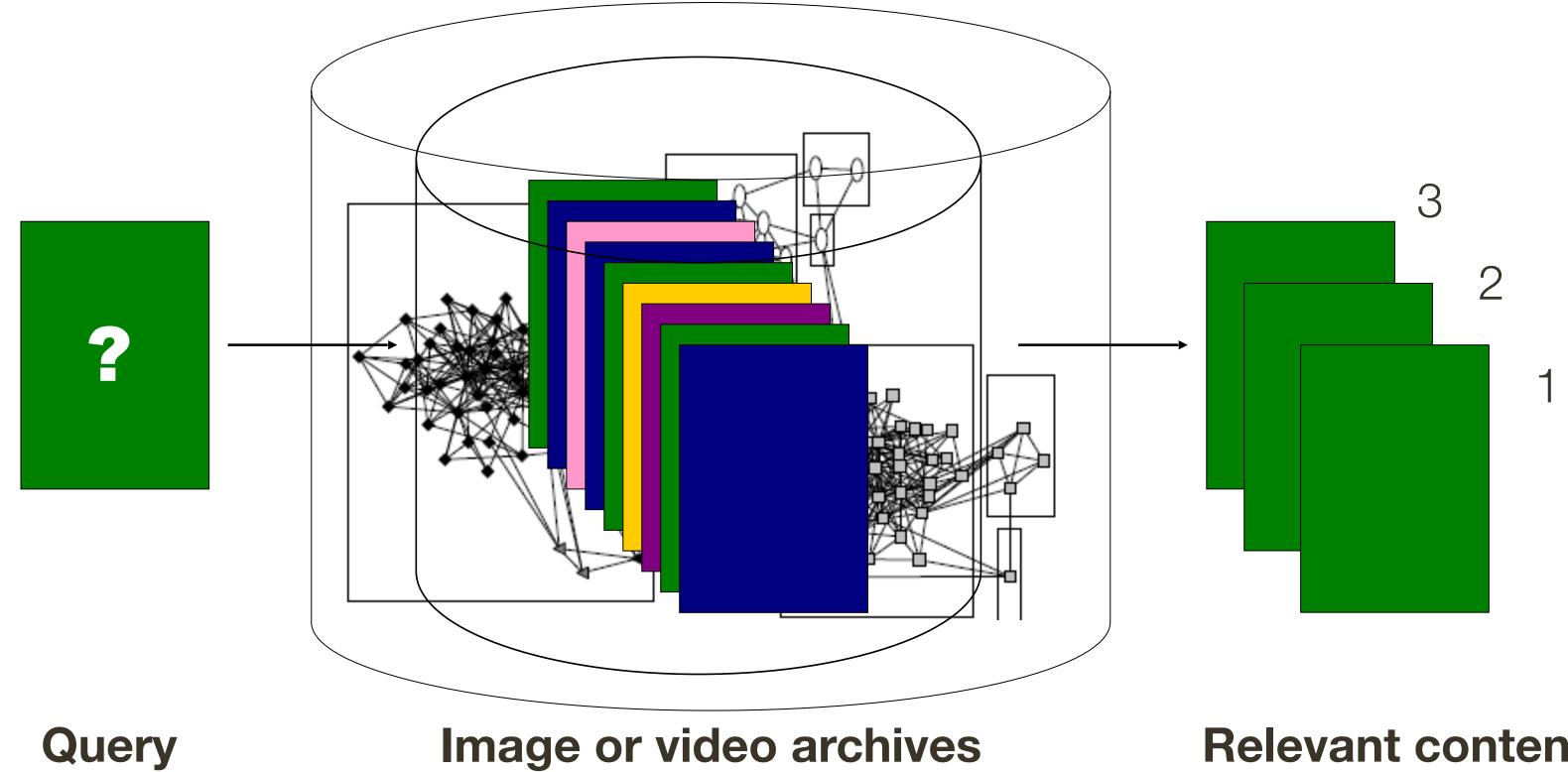
1. Computing properties of the 3D world from visual data (*measurement*)

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)

2. Algorithms and representations to allow a machine to recognize objects,



Relevant content

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)



Scale is enormous, explosion of visual content



*from iStock by Gettylmages



*from iStock by GettyImages





31.7 Million / hour

WhatsApp



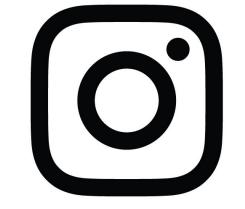
29.2 Million / hour





14.6 Million / hour

Instagram



2.9 Million / hour

Flickr



0.2 Million / hour

You Tube

18K hours / hour

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



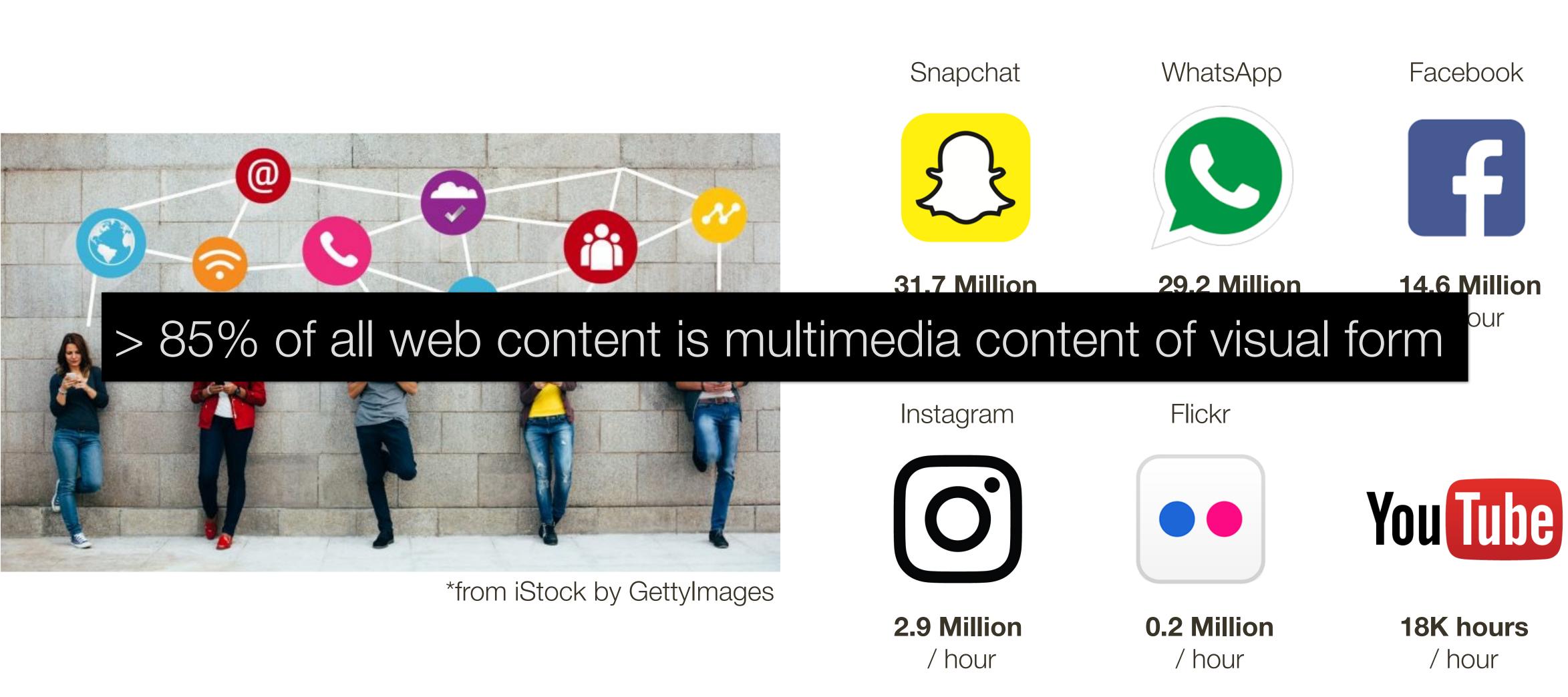












*based on article by Kimberlee Morrison in Social Times (2015)



- 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)

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

2. Algorithms and representations to allow a machine to recognize objects,

4. Visual Imagination



A brain riding a rocketship heading towards the moon.

A dragon fruit wearing karate belt in the snow.

A marble statue of a Koala DJ in front of a marble statue of a turntable. The Koala has wearing large marble headphones.



A Pomeranian is sitting on the Kings throne wearing a crown. Two tiger soldiers are standing next to the throne.

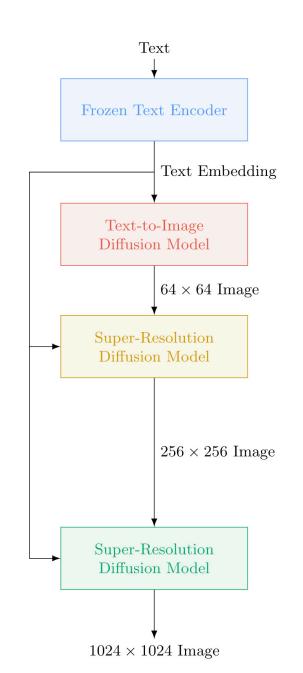
An extremely angry bird.

Android Mascot made from bamboo.

Three spheres made of glass falling into ocean. Water is splashing. Sun is setting.

A single beam of light enter the room from the ceiling. The beam of light is illuminating an easel. On the easel there is a Rembrandt painting of a raccoon.

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



"A Golden Retriever dog wearing a blue checkered beret and red dotted turtleneck."





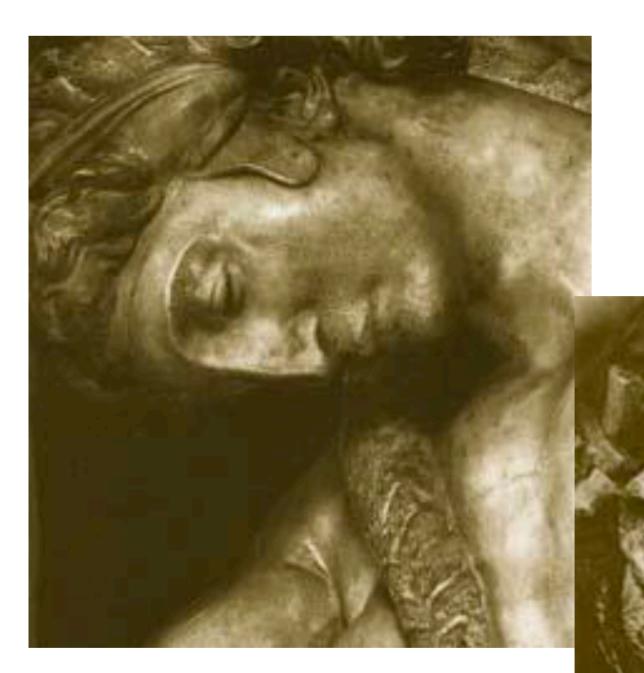
- people, scenes, and activities (perception and interpretation)
- 3. Algorithms to mine, search, and interact with visual data (search and organization)
- 4. Algorithms for manipulation or creation of image or video content (visual imagination)

Slide Credit: Kristen Grauman (UT Austin)

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

2. Algorithms and representations to allow a machine to recognize objects,

Challenges: Viewpoint invariance



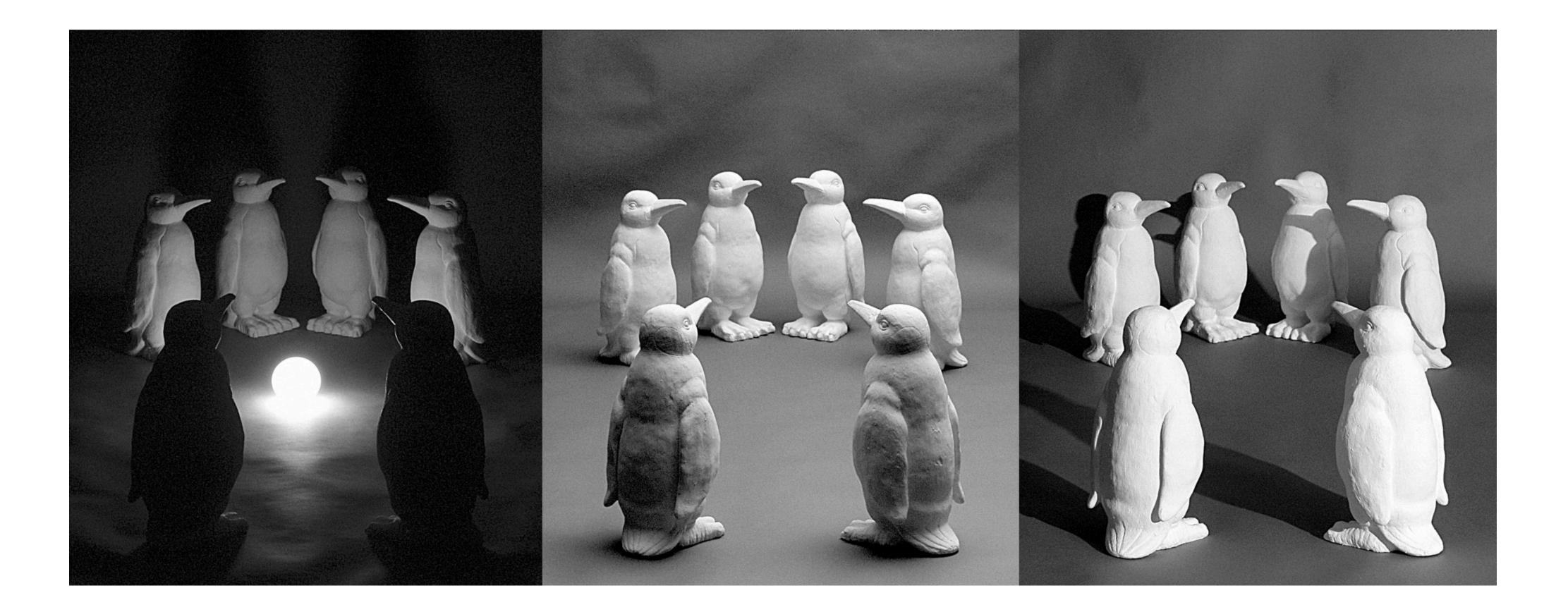
Michelangelo 1475-1564



*slide credit Fei-Fei, Fergus & Torralba



Challenges: Lighting



*image credit J. Koenderink

Challenges: Scale





*slide credit Fei-Fei, Fergus & Torralba



Challenges: Deformation



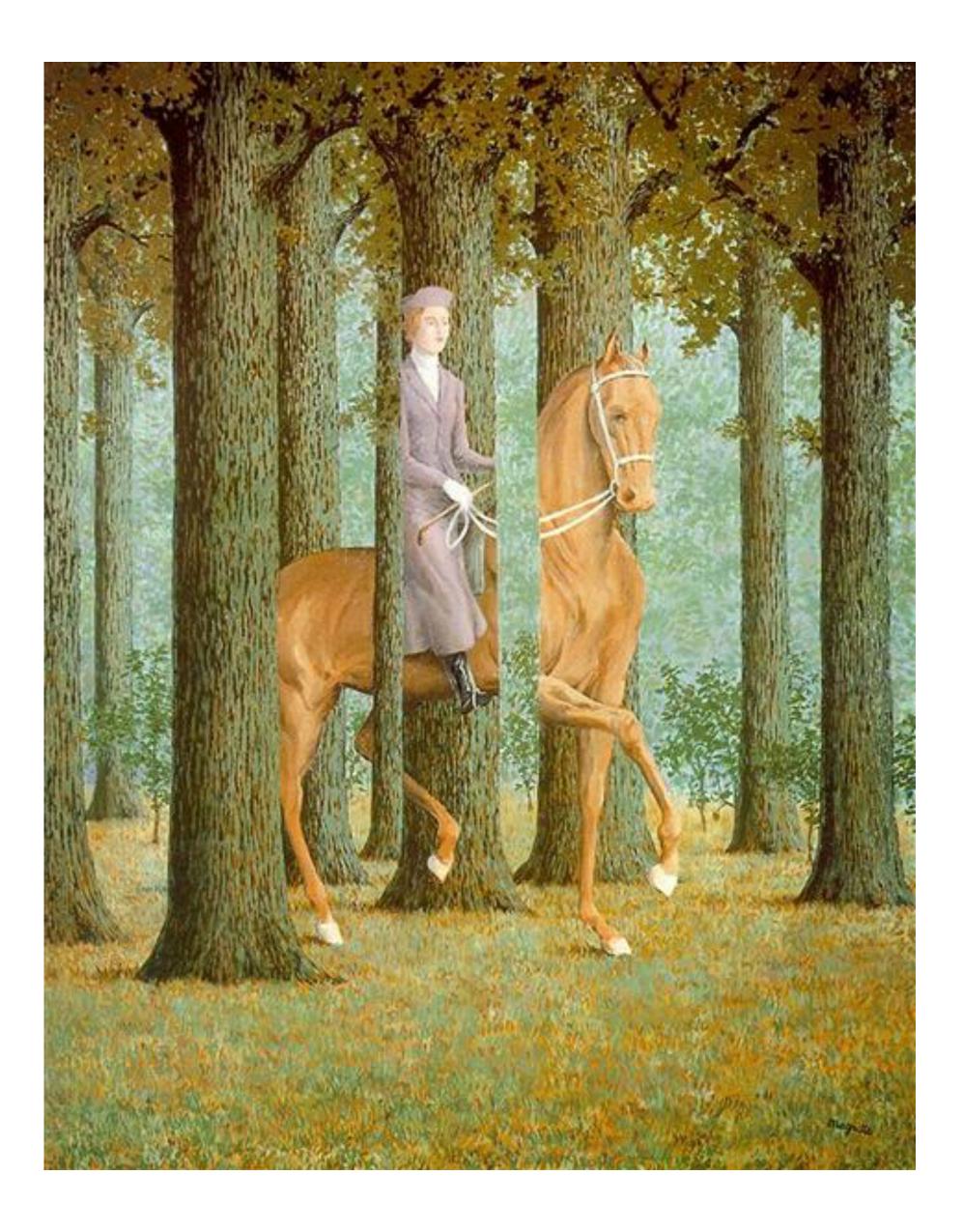


*image credit Peter Meer



Challenges: Occlusions

Rene Magritte 1965

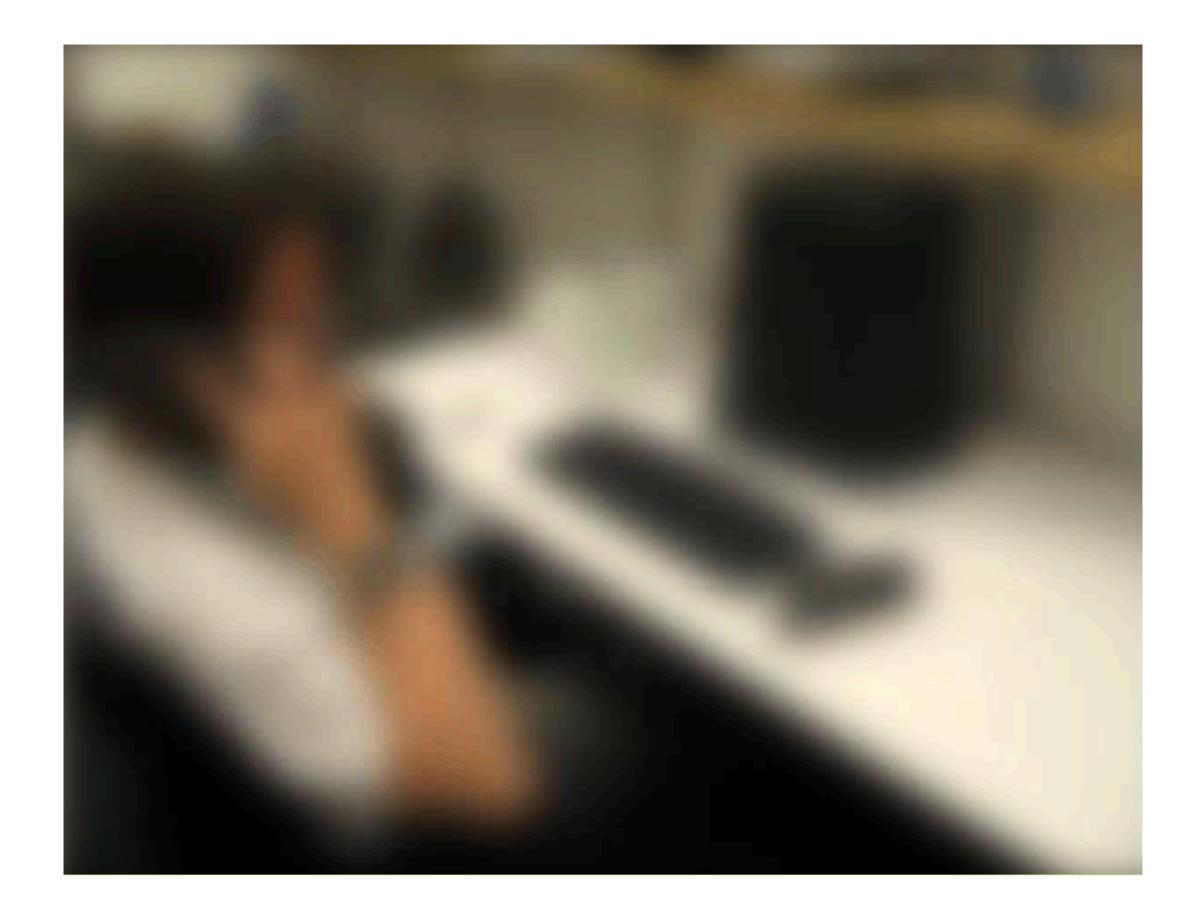


Challenges: Background clutter

Kilmeny Niland 1995



Challenges: Local ambiguity and context



*image credit Fergus & Torralba



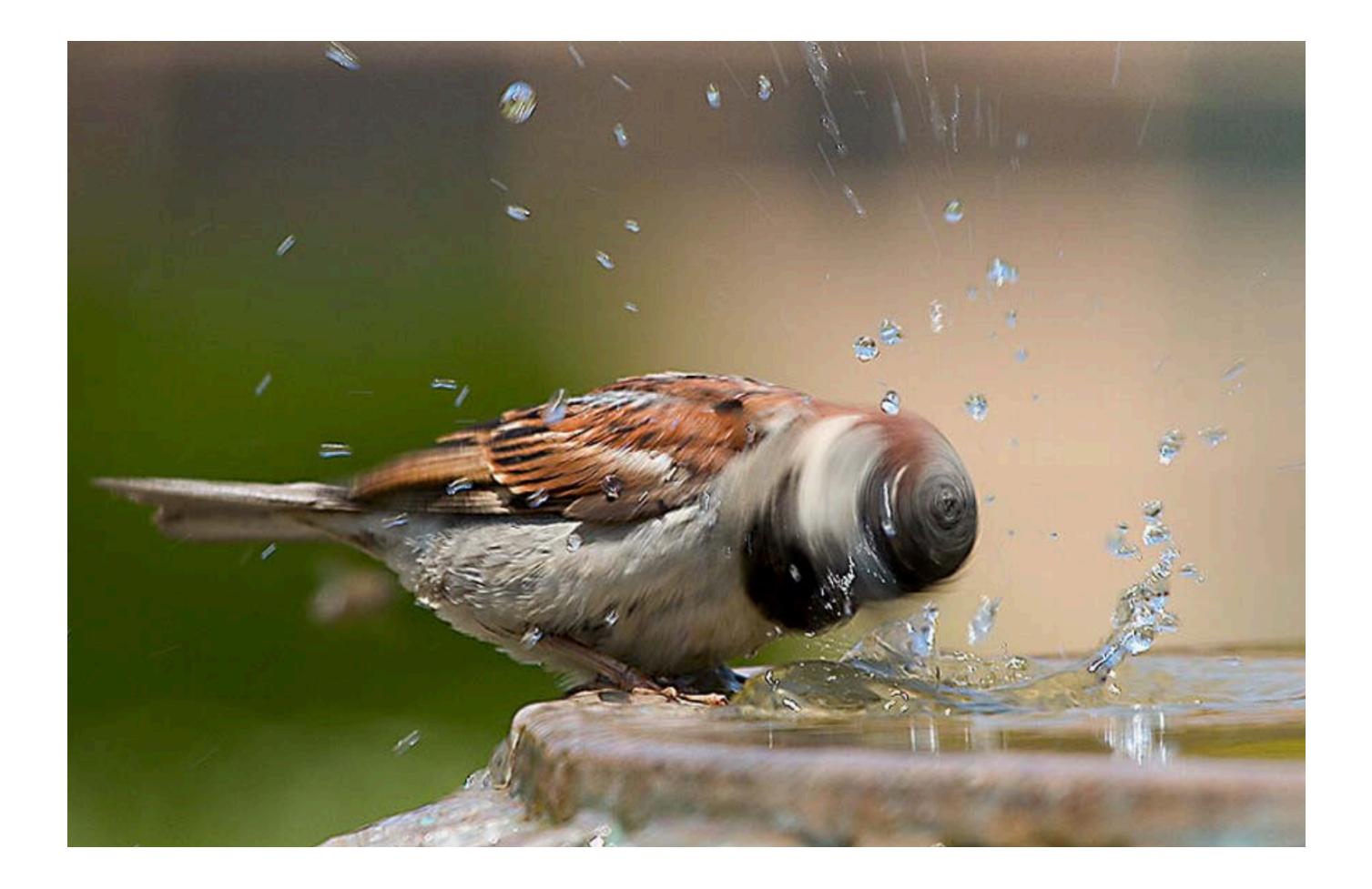
Challenges: Local ambiguity and context



*image credit Fergus & Torralba



Challenges: Motion



*image credit Peter Meer



Challenges: Object inter-class variation









*slide credit Fei-Fei, Fergus & Torralba

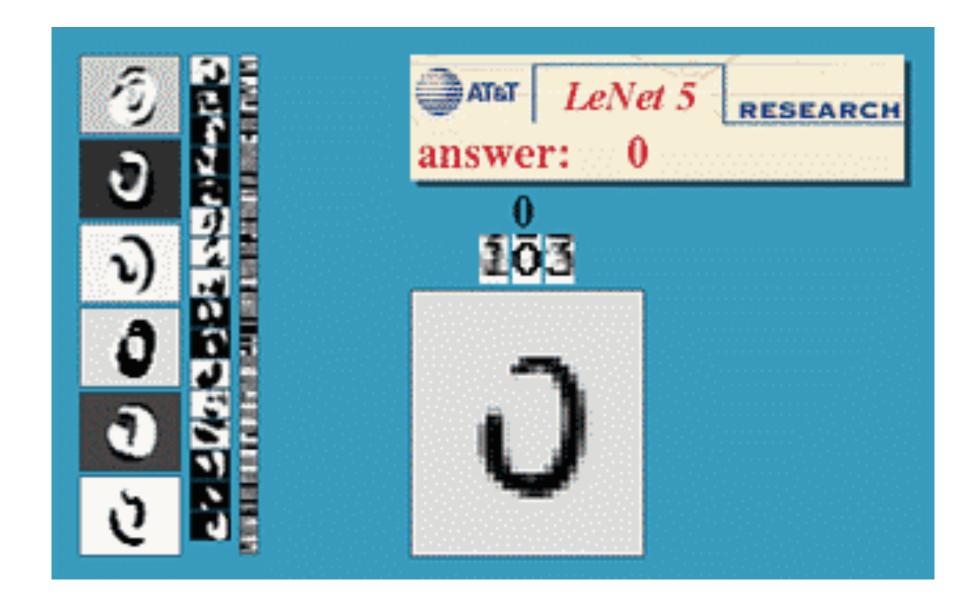


Computer Vision Applications

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)



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





Yann LeCun

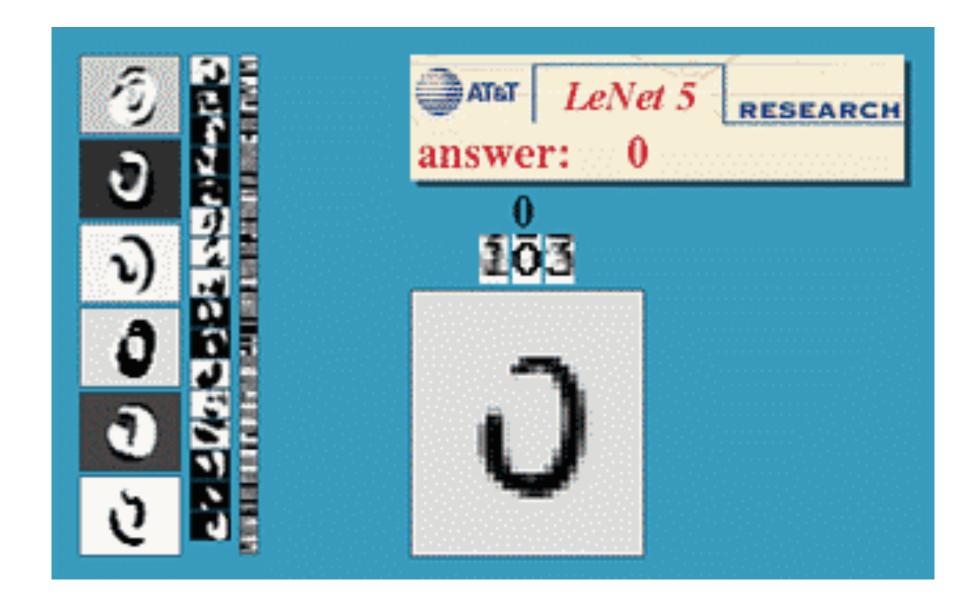


License plate readers http://en.wikipedia.org/wiki/Automatic_number_plate_recognition



Optical Character Recognition (OCR)

Technology to convert scanned documents to text (comes with any scanner now days)



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





Yann LeCun



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]

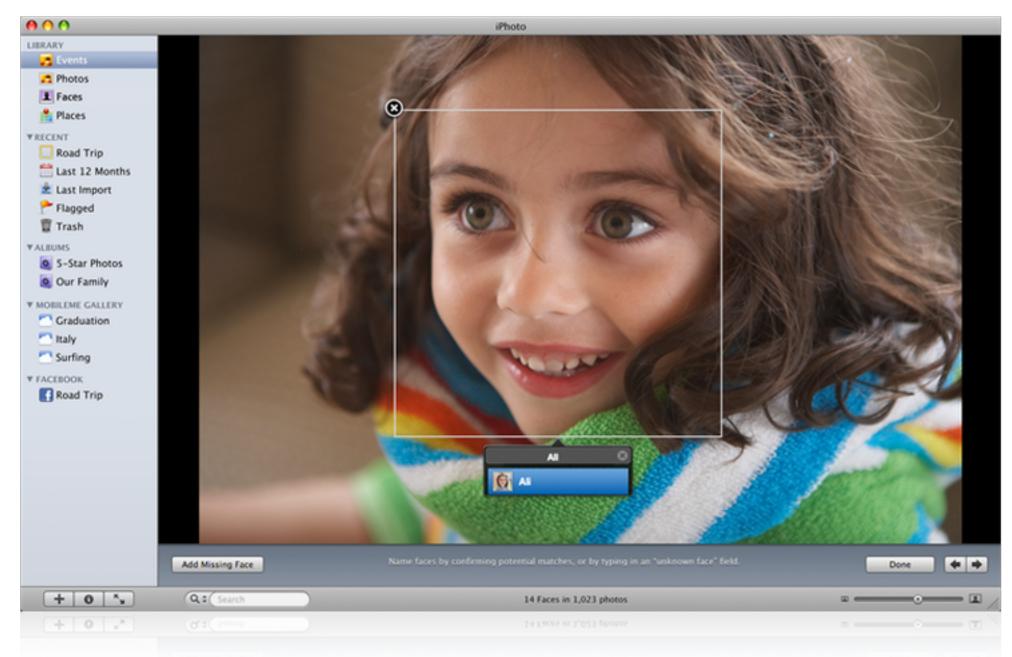
Face **Recognition**



Facebook

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

Apple's iPhoto



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

Vision for **Biometrics**

VOL. 167, NO. 6

NATIONAL GEOGRAPHIC

JUNE 1985

GREAT SALT LAKE: THE FLOODING DESERT 694

U.S.- MEXICAN BORDER: LIFE ON THE LINE 720

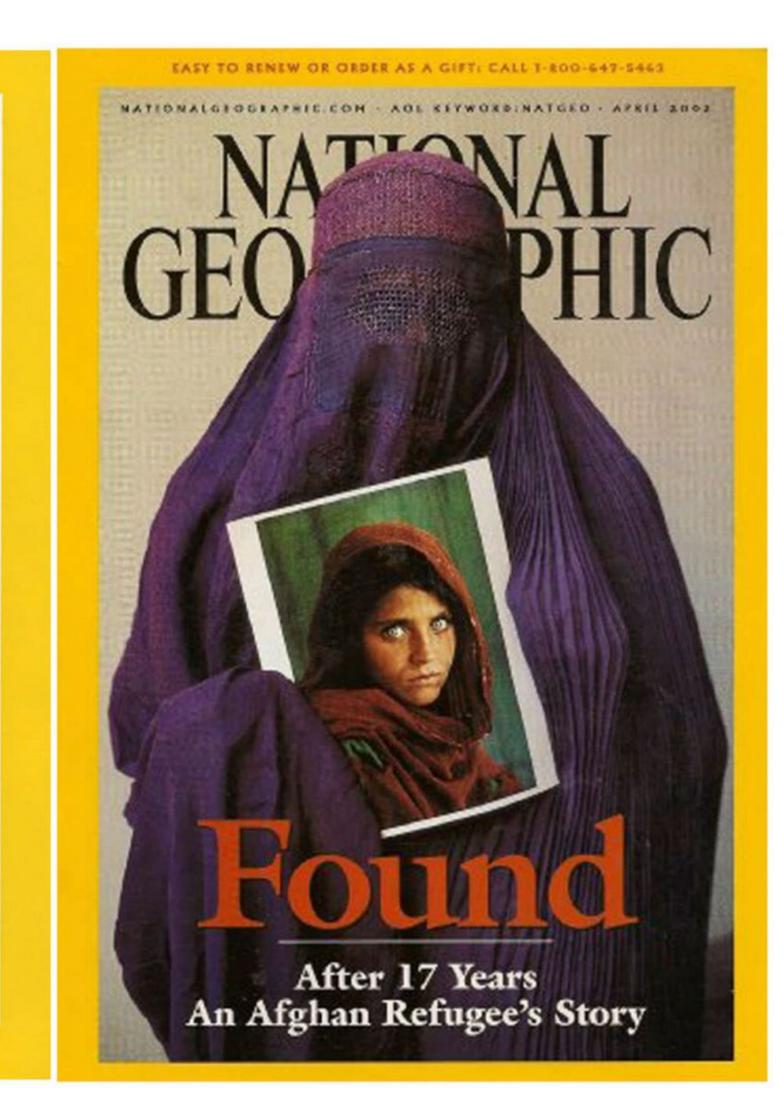
JAVA'S WILDLIFE RETURNS 750

Along Afghanistan's War-torn Frontier 772

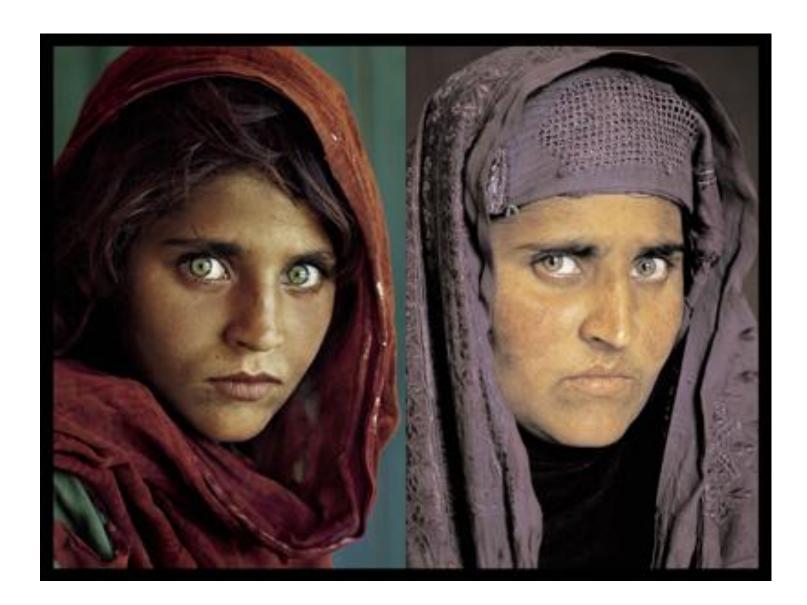
Haunted eyes tell of an Afghan refugee's fears

FAIR SKIES FOR THE CAYMAN ISLANDS 798

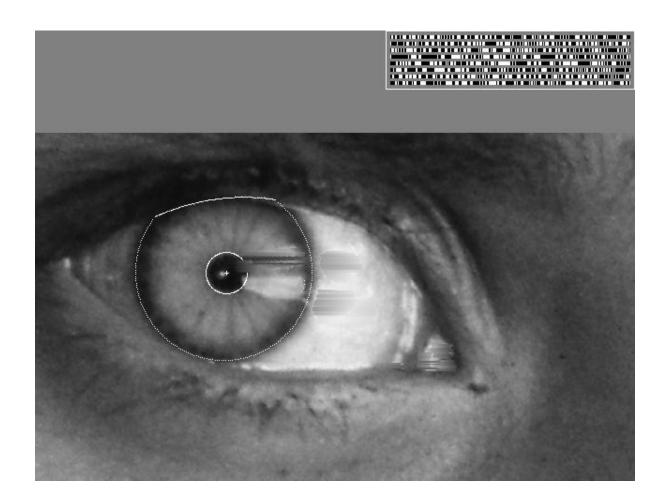
SEE NATIONAL GEOGRAPHIC EXPLORER EVERY SUNDAY ON NICKELODEON CABLE TV



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

Image Credit: James Hays (GA Tech)

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

Camera Tracking



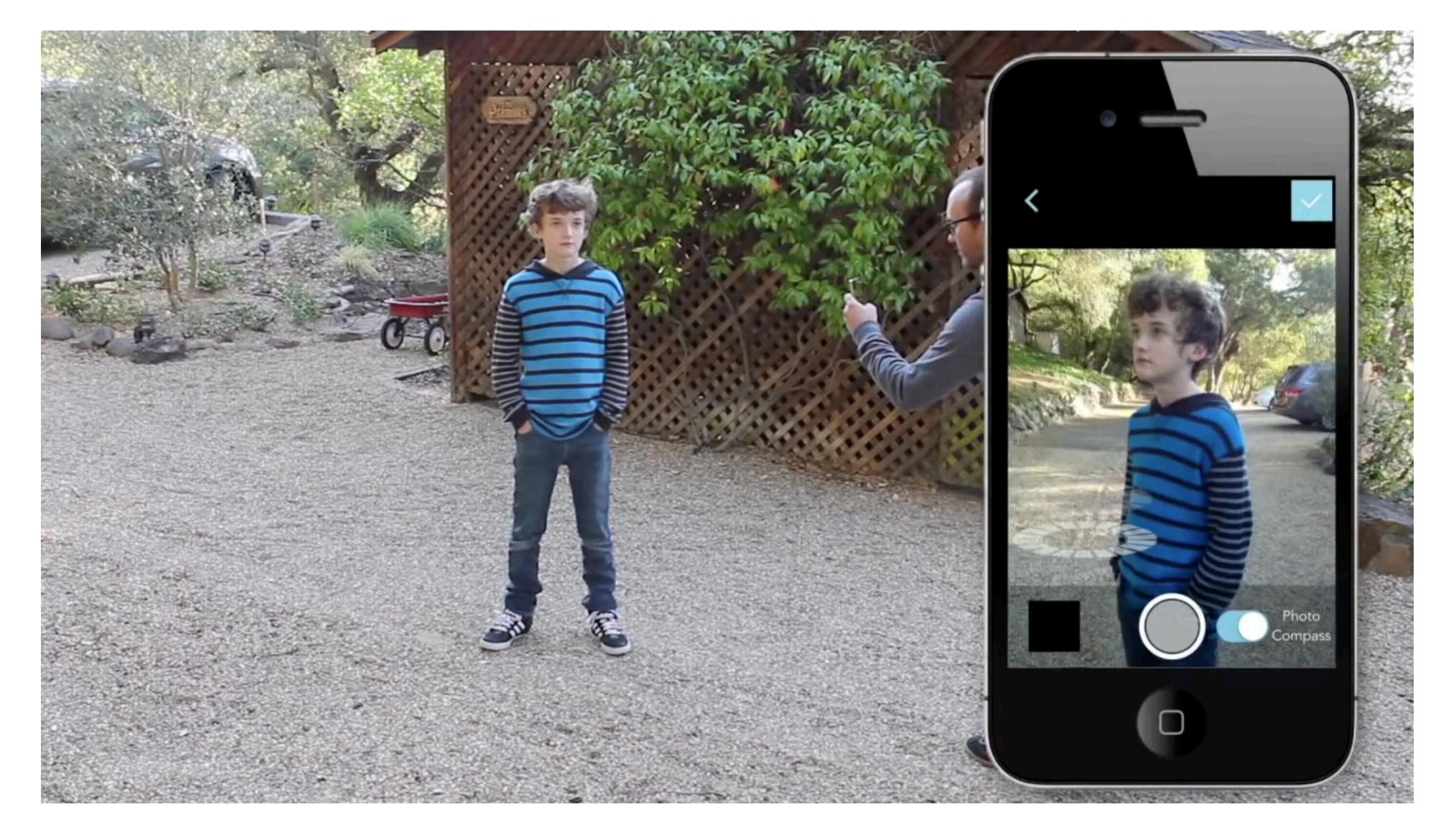
[Boujou — Vicon / OMG]

Camera Tracking



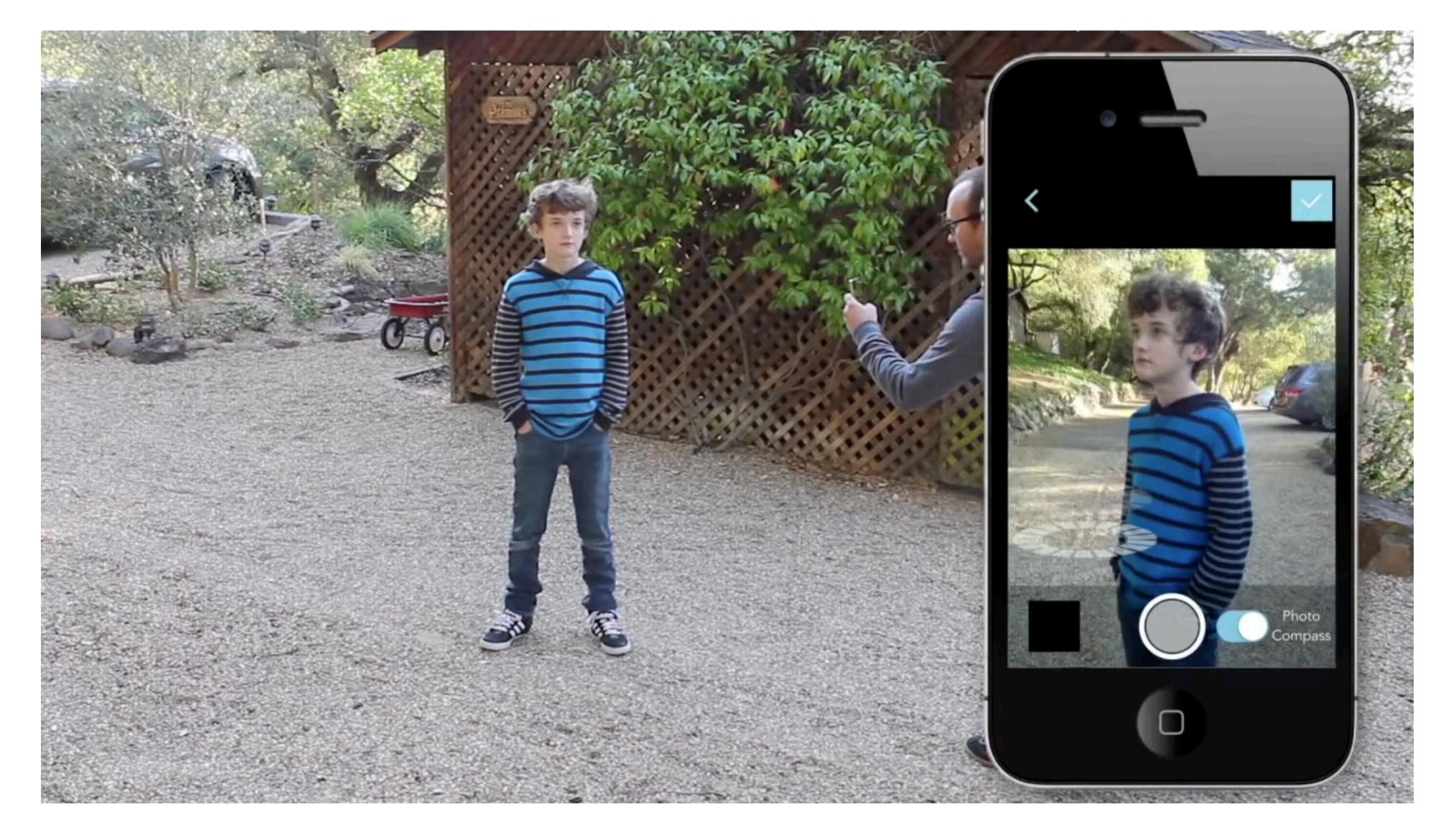
[Boujou — Vicon / OMG]

3D Reconstruction



[Autodesk 123D Catch]

3D Reconstruction



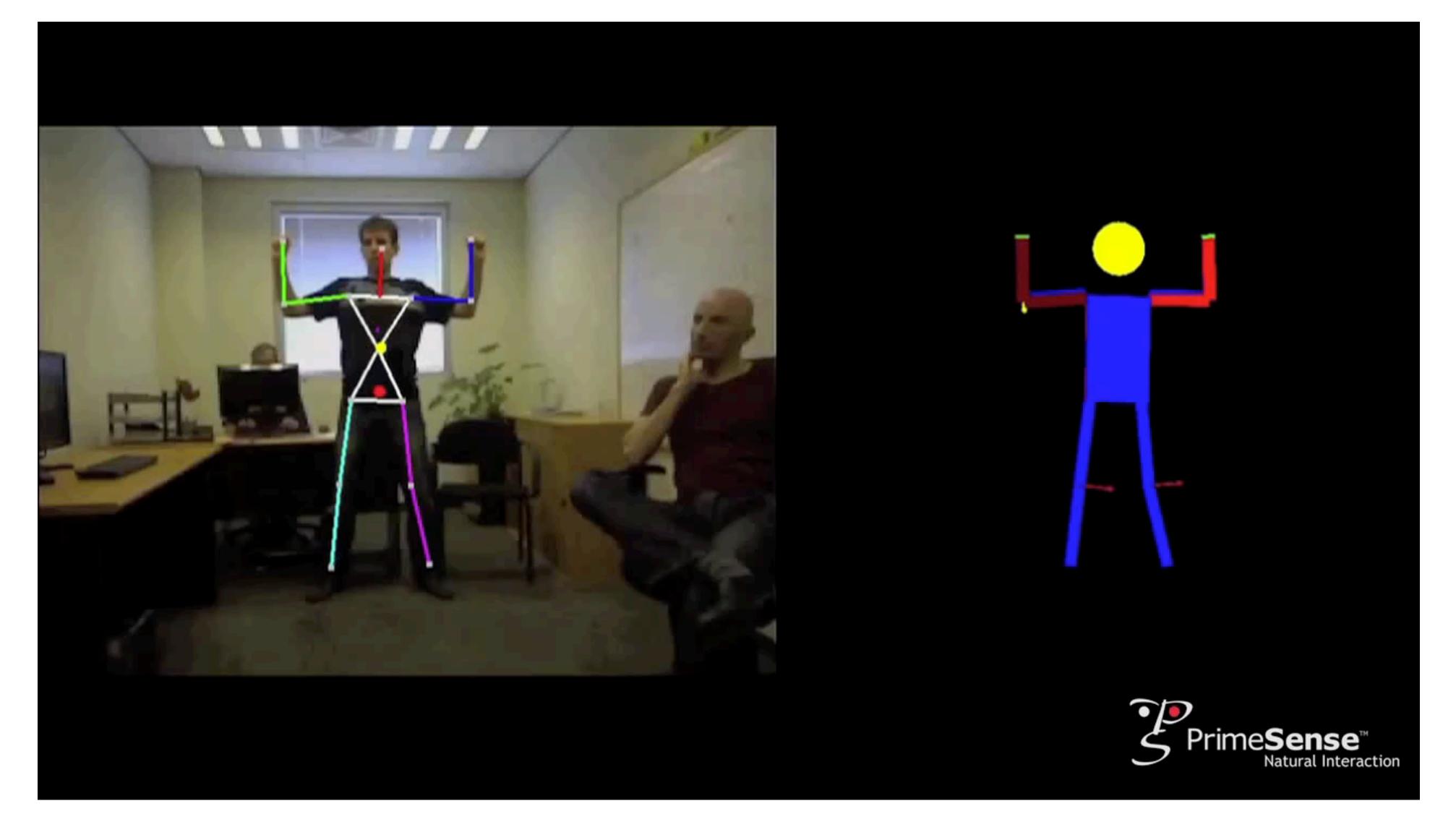
[Autodesk 123D Catch]

Body Pose Tracking



[Microsoft Xbox Kinect]

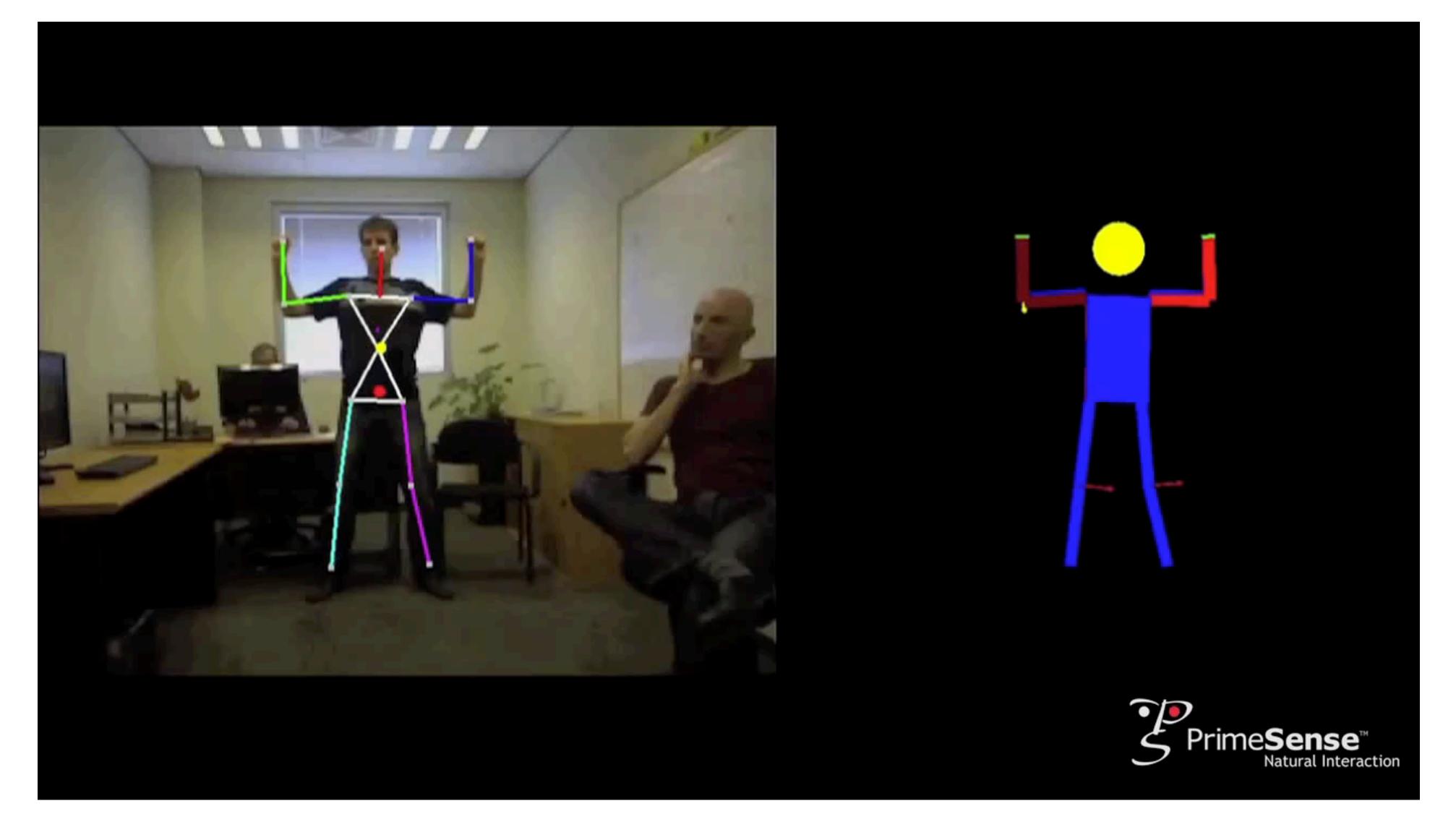
Body Pose Tracking





[PrimeSense]

Body Pose Tracking

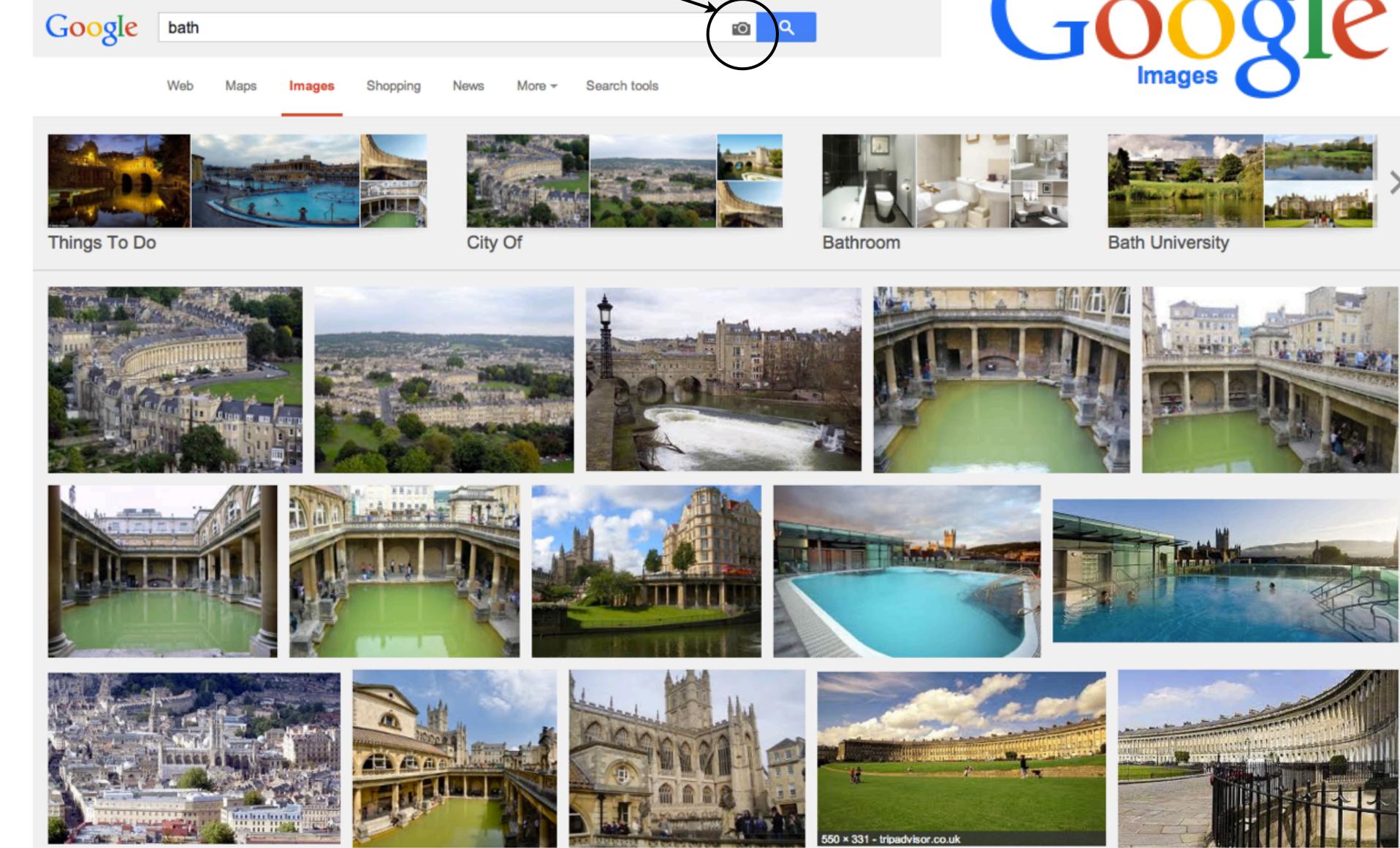




[PrimeSense]

Image Recognition and Search

Search by image









Self-Driving Cars



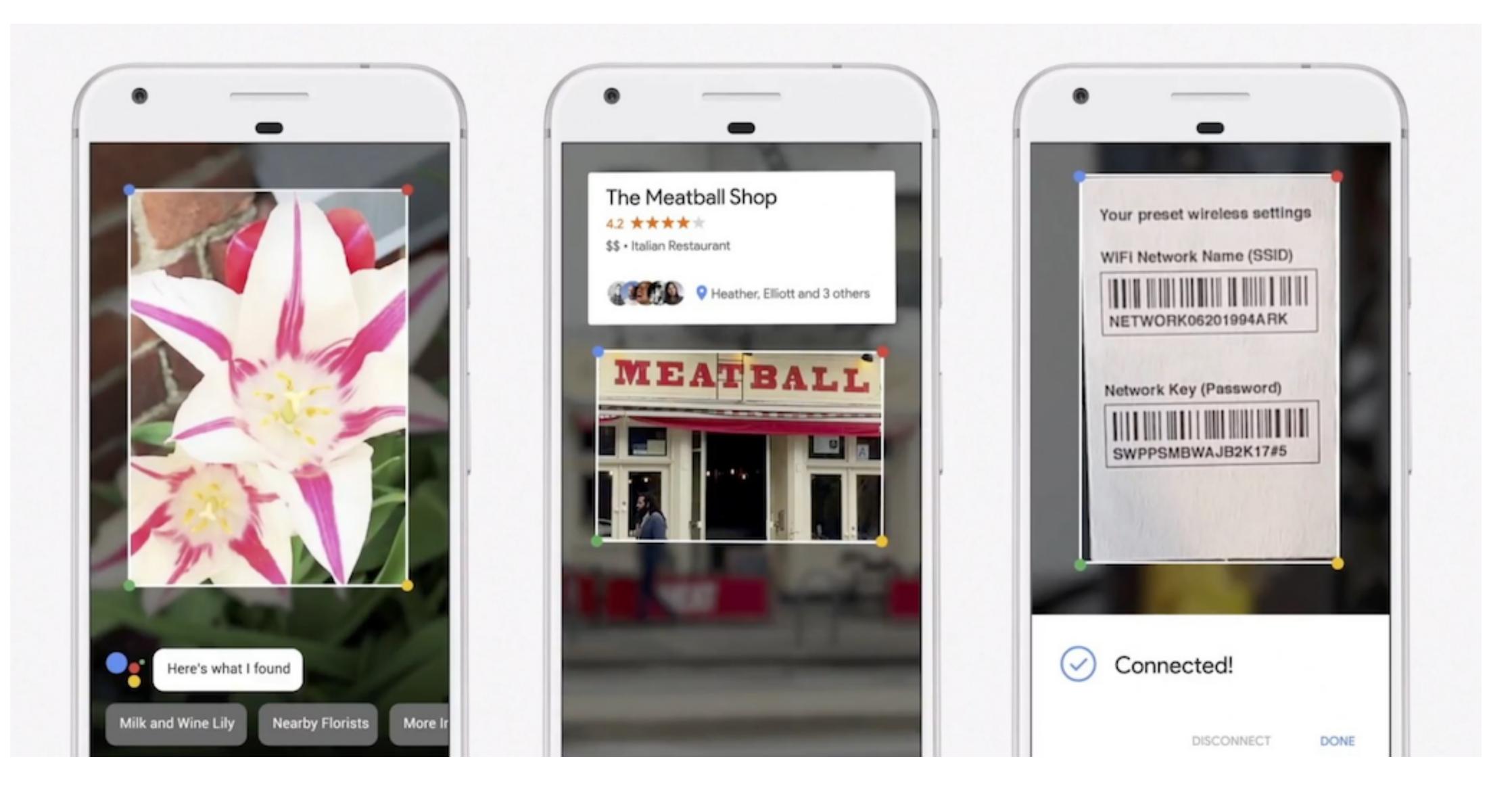
[Google]

AR / VR



[Microsoft HoloLens]

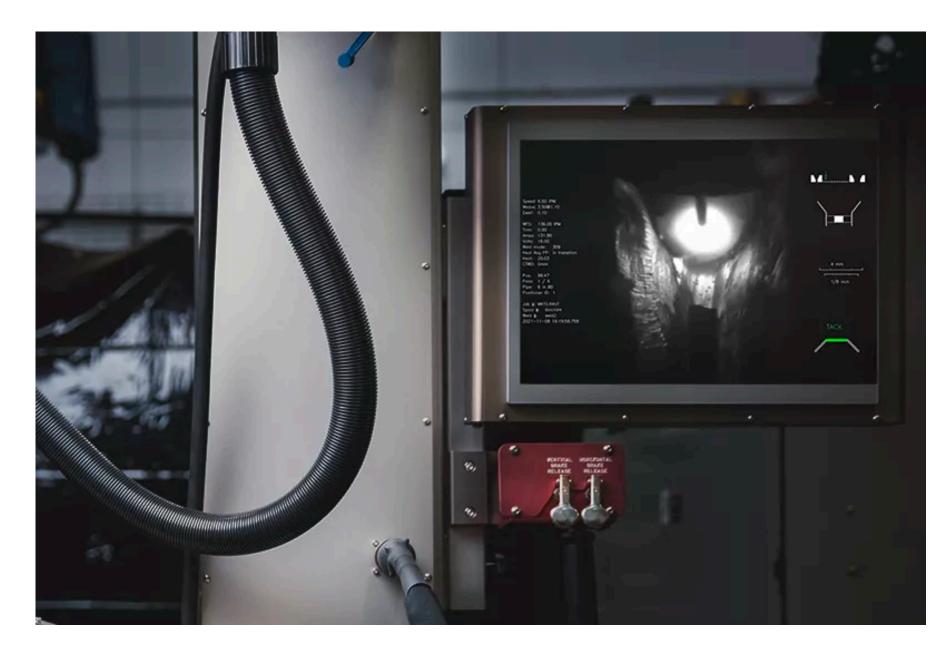
Mobile Apps



[Google Lens]

Industrial







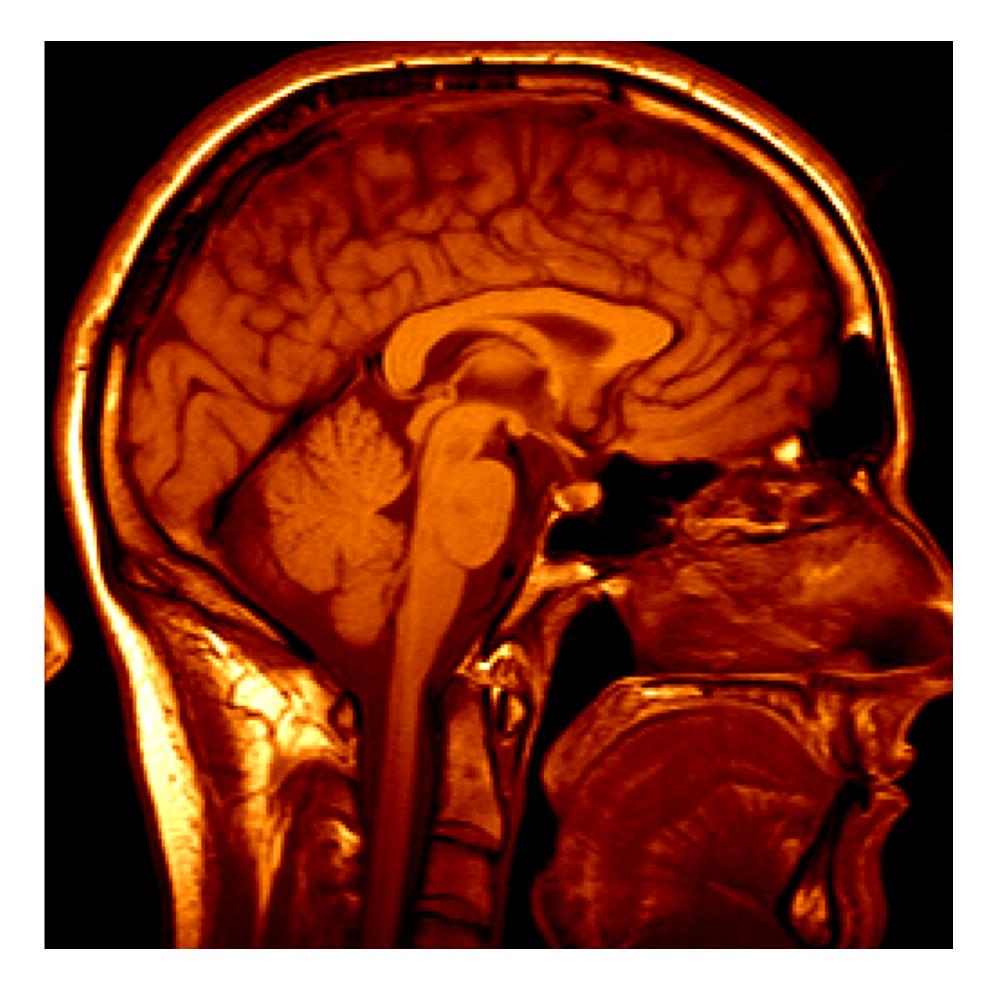
Machine Vision controlled welding robotics







Medicine



3D imaging MRI, CT

Slide Credit: James Hays (GA Tech)



Image guided surgery <u>Grimson et al., MIT</u>

Art

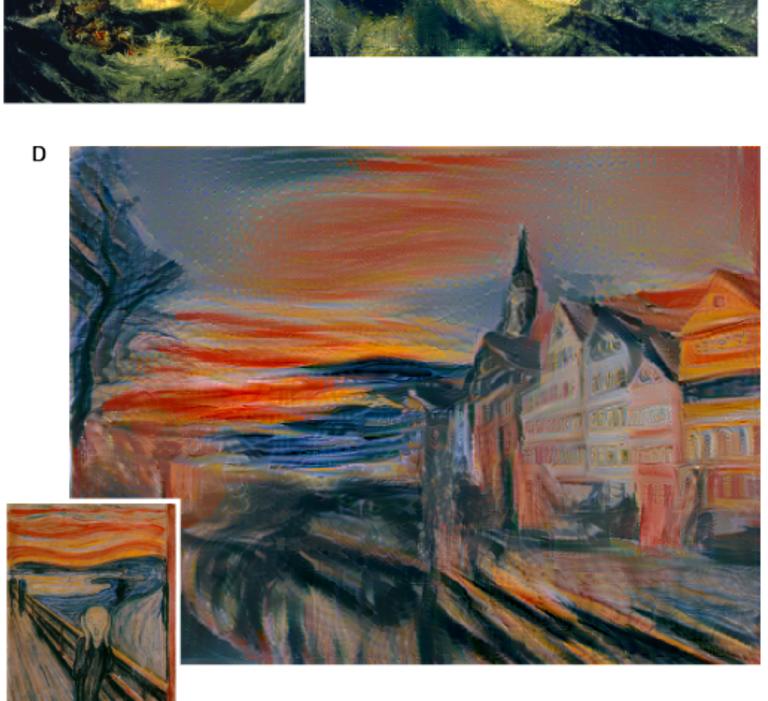
Α



С



[Gatys, Ecker, Bethge 2015]





B





Art

TEXT DESCRIPTION

An astronaut Teddy bears A bowl of soup

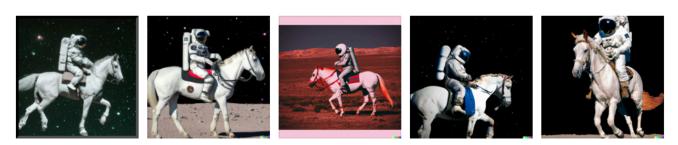
riding a horse lounging in a tropical resort in space playing basketball with cats in space

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

 \rightarrow







[Dall-Ev2]

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'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

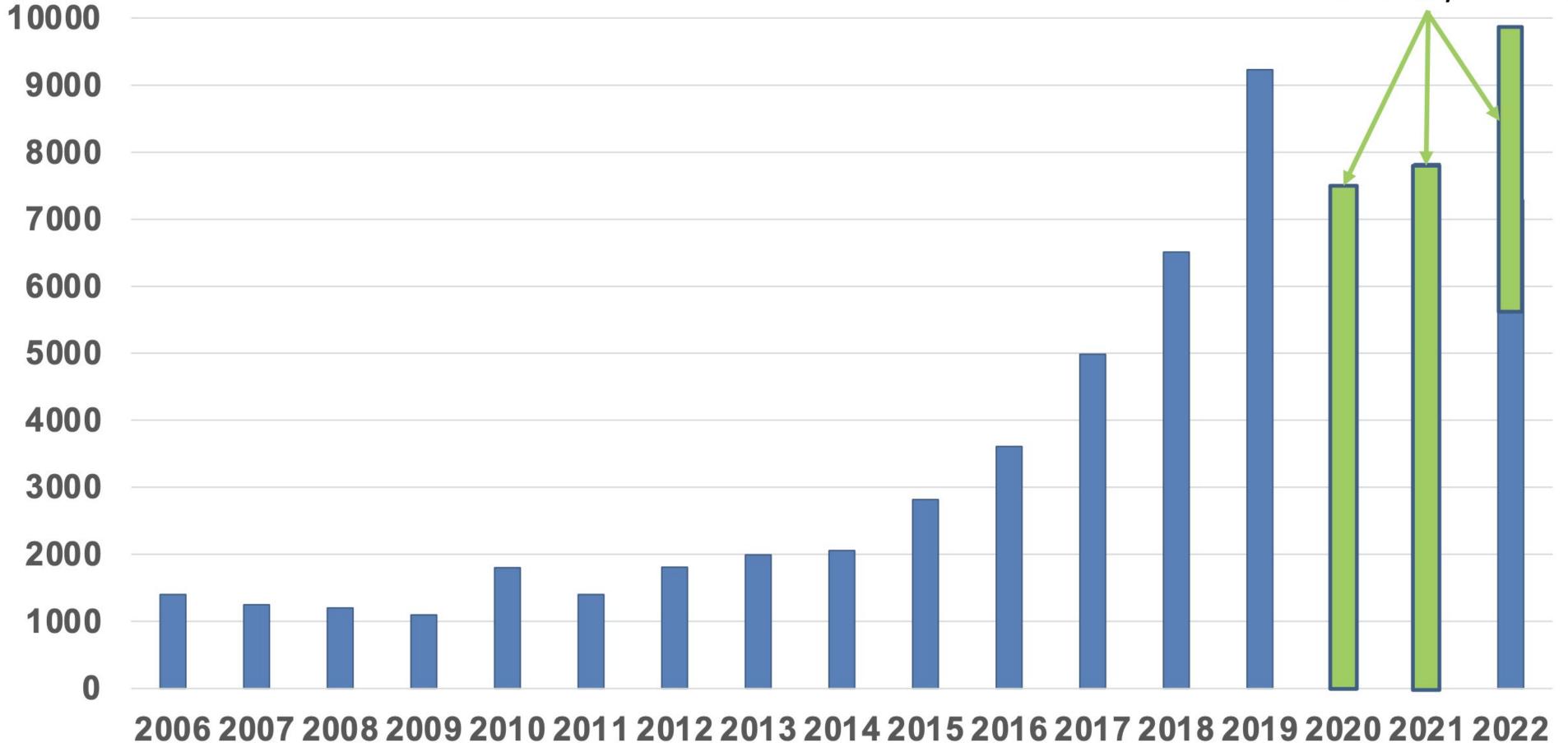




PR JUNE NEW ORLEANS 19-24 2022 LOUISIANA



CVPR Attendance



Remote/Virtual attendees

A VANCOUVER V

Welling History 12

100 A 10

CVPR 2023 June 17-23, 2023



Course Schedule

Leonid Sigal

Associate Professor, University of British Columbia

Menu

HOME
ABOUT
BIOGRAPHY
cv
STUDENTS AND COLLABORATORS
RESEARCH
TEACHING
CPSC 532S WINTER 1, 2022
CPSC 532S
WINTER 2, 2020
CPSC 425
WINTER 1, 2020
CPSC 425
WINTER 2, 2019
CPSC 425
WINTER 2, 2018
CPSC 532S
WINTER 2, 2018
CPSC 425
WINTER 1, 2018

Topics in Artificial Intelligence (CPSC 532S):

Multimodal Learning with Vision, Language and Sound

Winter Term 1, 2022

Course Information

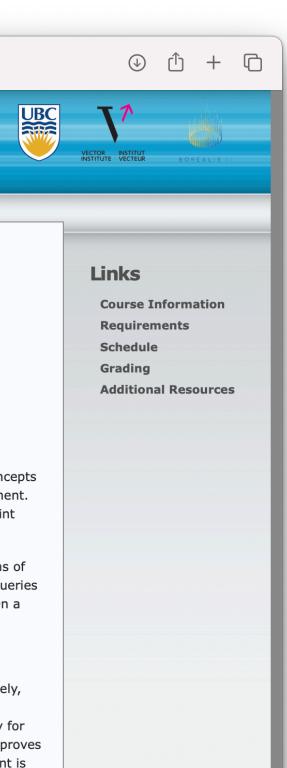
Multimodal machine learning is a multi-disciplinary research field which addresses some of the core goals of artificial intelligence by integrating and modeling two or more data modalities (e.g., visual, linguistic, acoustic, etc.). This course will teach fundamental concepts related to multimodal machine learning, including (1) representation learning, (2) translation and mapping, and (3) modality alignment. While the fundamental techniques covered in this course are applicable broadly, the focus will on studying them in the context of joint reasoning and understanding of images/videos and language (text).

🔒 cs.ubc.ca

In addition to fundamentals, we will study recent rich body of research at the intersection of vision and language, including problems of (i) generating image descriptions using natural language, (ii) visual question answering, (iii) retrieval of images based on textural queries (and vice versa), (iv) generating images/videos from textual descriptions, (v) language grounding and many other related topics. On a technical side, we will be studying neural network architectures of various forms, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), memory networks, attention models, neural language models, structures prediction models.

Content Delivery and Covid Precautions: The lectures will be offered *in-person only* and no recordings will be made. Unfortunately, for this reason, a hybrid delivery of material will not be available. We will experiment with hybrid office hours, as we believe this will benefit the students. Students are strongly **encouraged and expected** (but not required) **to wear masks in class**. This is largely for the benefit of your fellow students with whom you will sit in close proximity. Instructor will not wear a mask when lecturing (this improves delivery of the material) but will put on the mask in close interaction setting or when requested by students. If at any point a student is

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



C

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



;)

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

Previously taught by:

- 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: <u>45%</u>







6 graded and 1 ungraded (optional) assignment

Midterm Exam (October 19th): <u>15%</u>

Clicker Questions **Bring** your i>Clickers to class

Register your remote: <u>https://canvas.ubc.ca/</u> before the next class

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 <u>10%</u> 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.

(we will test them next week)

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)



Scikit-Learn

Assignments contribute <u>45%</u> to your final score

- You will use the **Python**, with the following libraries: Python Imaging Library (PIL), NumPy, Matplotlib, SciPy,
- Assignment 0 (which is ungraded) will introduce you to this.



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 <u>15%</u> to your final score



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 30% to your final score

Grading issues & Academic Misconduct

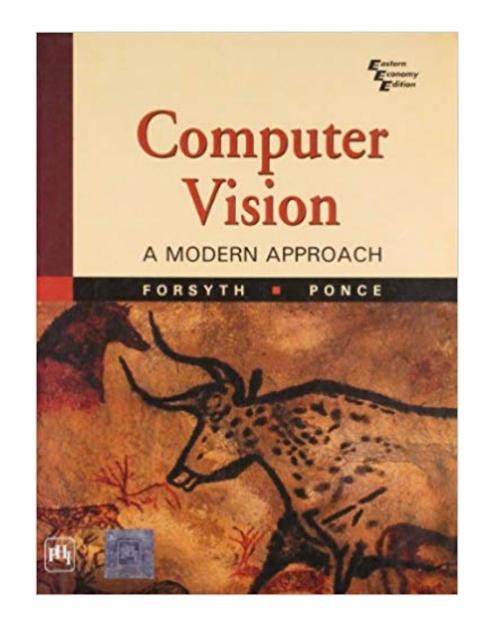
Strict policy: Grading mistakes happen, it's just a nature of life. If you see an issue with your grade, you have 1 week from the release of any assignment grade to bring a <u>specific issue</u> to our attention.

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



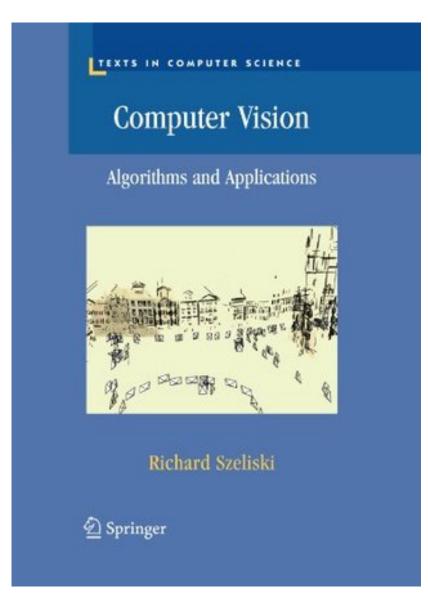
Textbooks

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:

Reminders:

- **[optional]** Watch TED talk by Prof. Fei-Fei Li https://www.youtube.com/watch?v=40riCqvRoMs



- Next Lecture: Szeliski Chapter 2, Forsyth & Ponce (2nd ed.) 1.1.1 - 1.1.3

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

