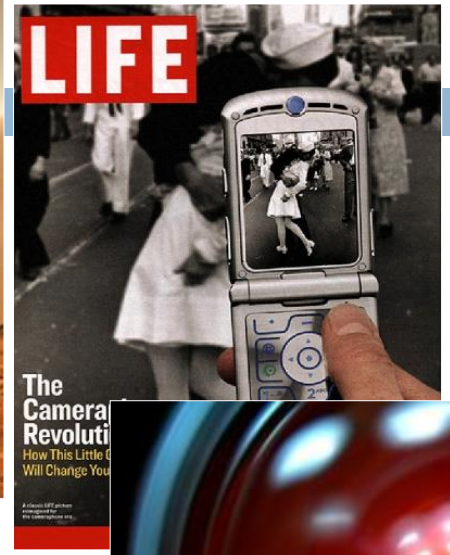
Computer Vision:
Foundations and Applications

Deepraj shukla

deeprajshukla@gmail.com

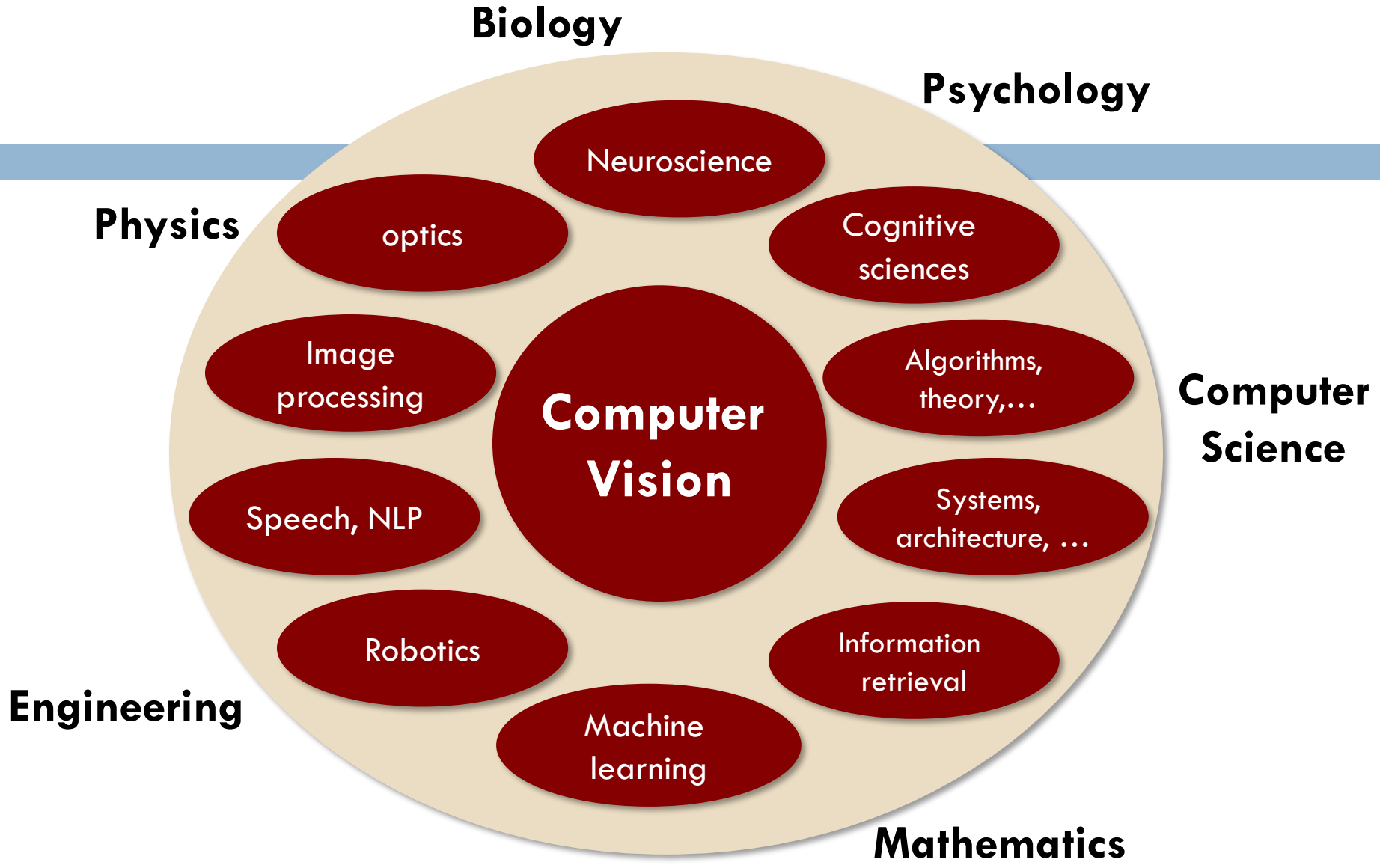
BEL , Bangalore



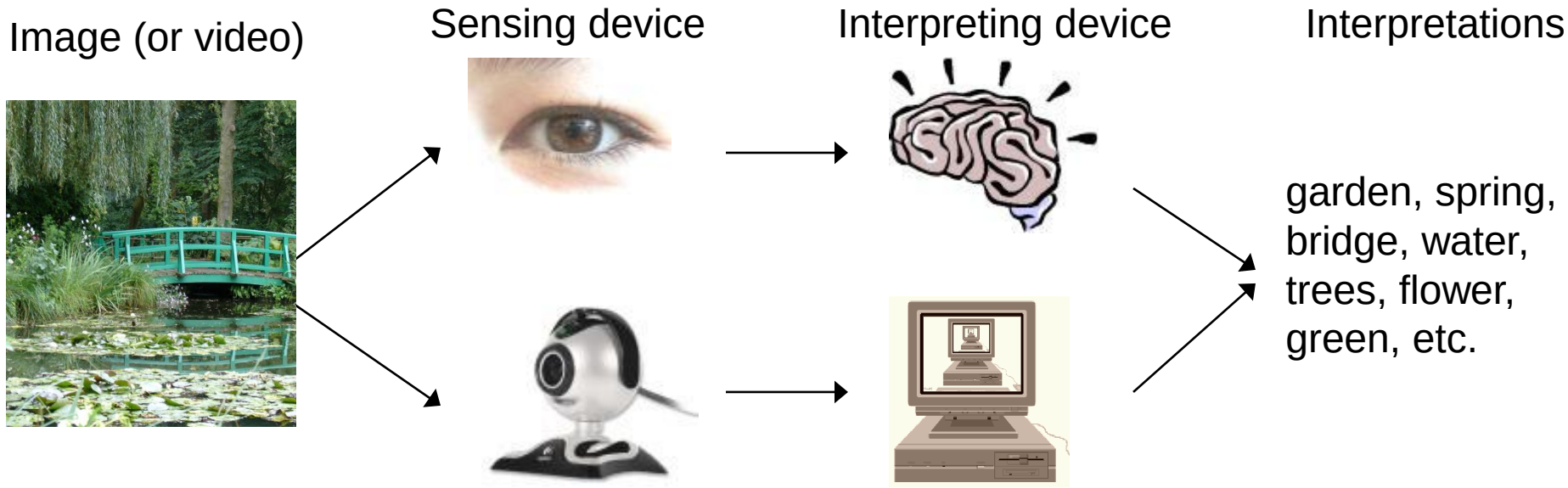
Introduction

Jason Salavon
GAN experiment on Twitter

Mario Klingemann, GAN experiment on Twitter


Biology
Psychology
Physics
Computer Science
Engineering
Mathematics
Neuroscience
optics
Cognitive sciences
Image processing
Computer Vision
Algorithms, theory,…
Speech, NLP
Systems, architecture, …
Robotics
Information retrieval
Machine learning

What is (computer) vision?



The goal of computer vision

- To bridge the gap between pixels and “meaning”

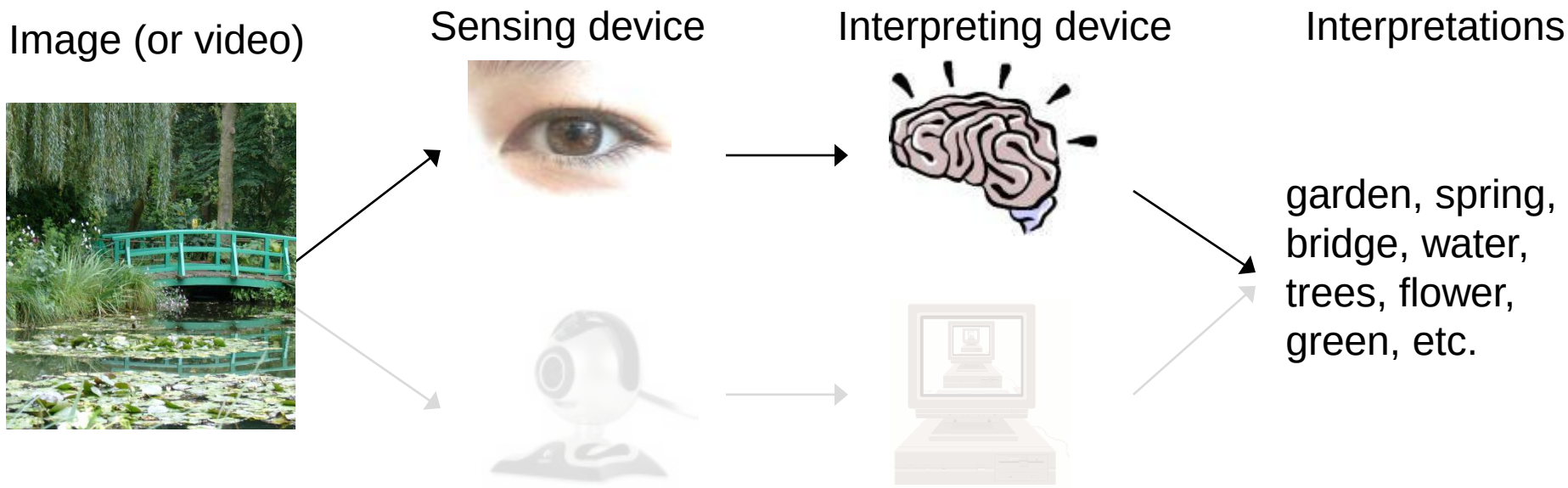
What we see

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

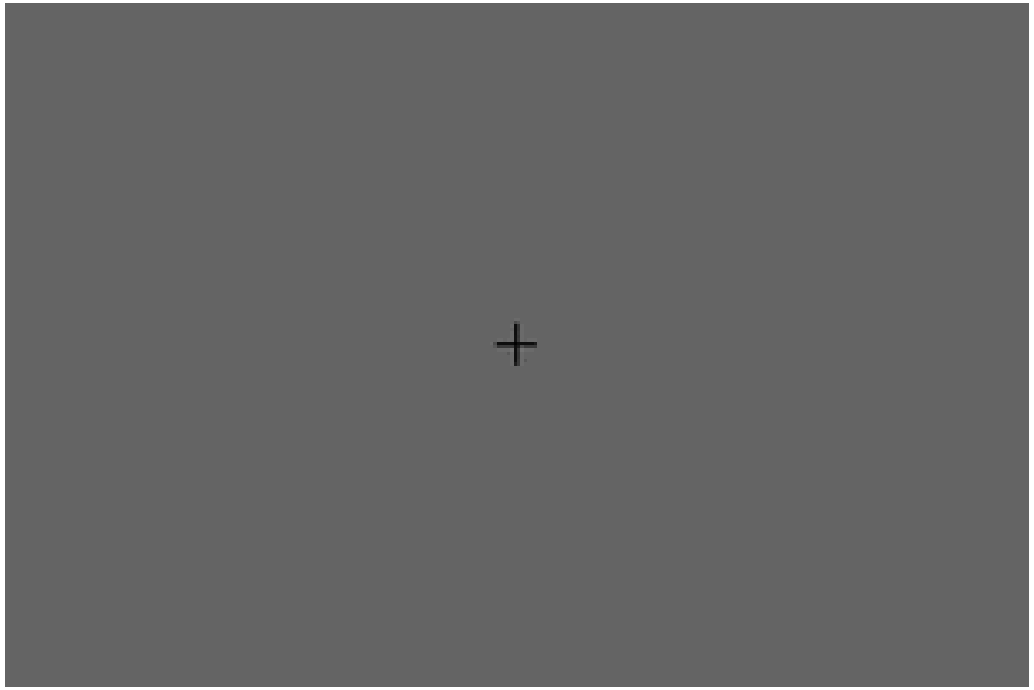
What is ~~(computer)~~ vision?



1981: Nobel Prize in medicine

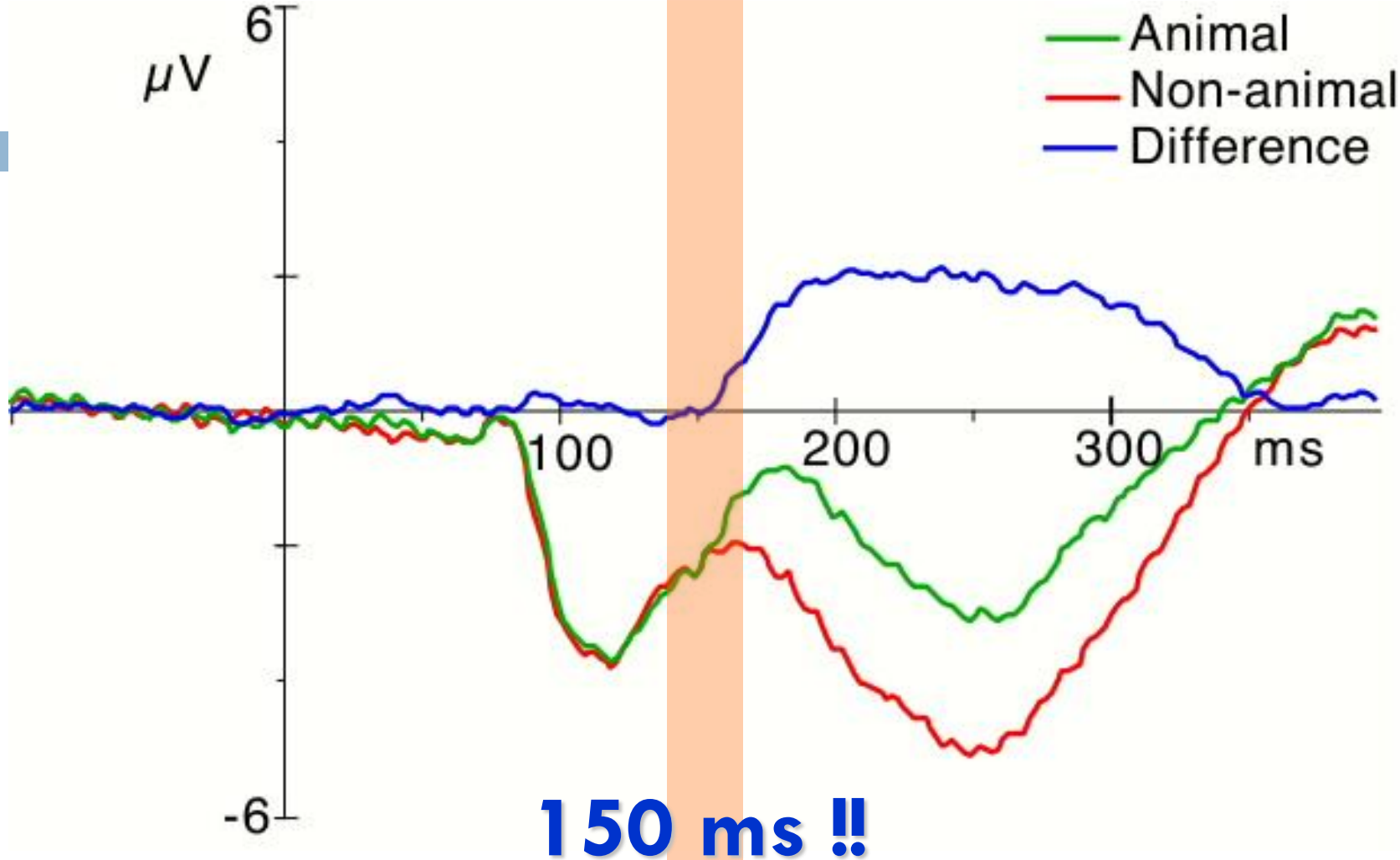
Hubel & Wiesel

Human vision is superbly efficient



Potter, Biederman, etc. 1970s

Thorpe, et al. *Nature,* 1996



150 ms !!

Thorpe, et al. *Nature,* 1996

Change Blindness

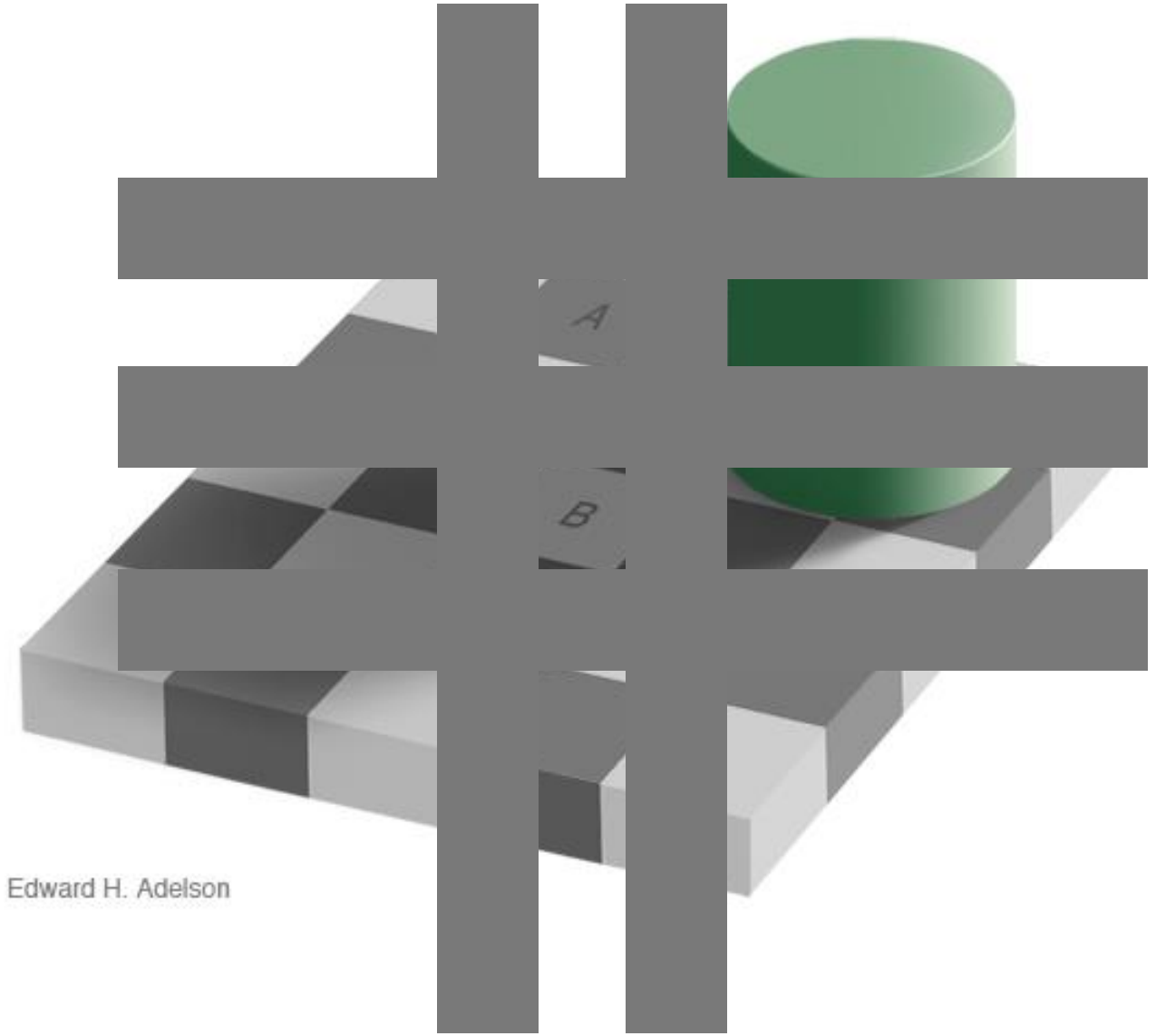
Rensink, O’regan, Simon, etc.

Change Blindness

Rensink, O’regan, Simon, etc.

Segmentation

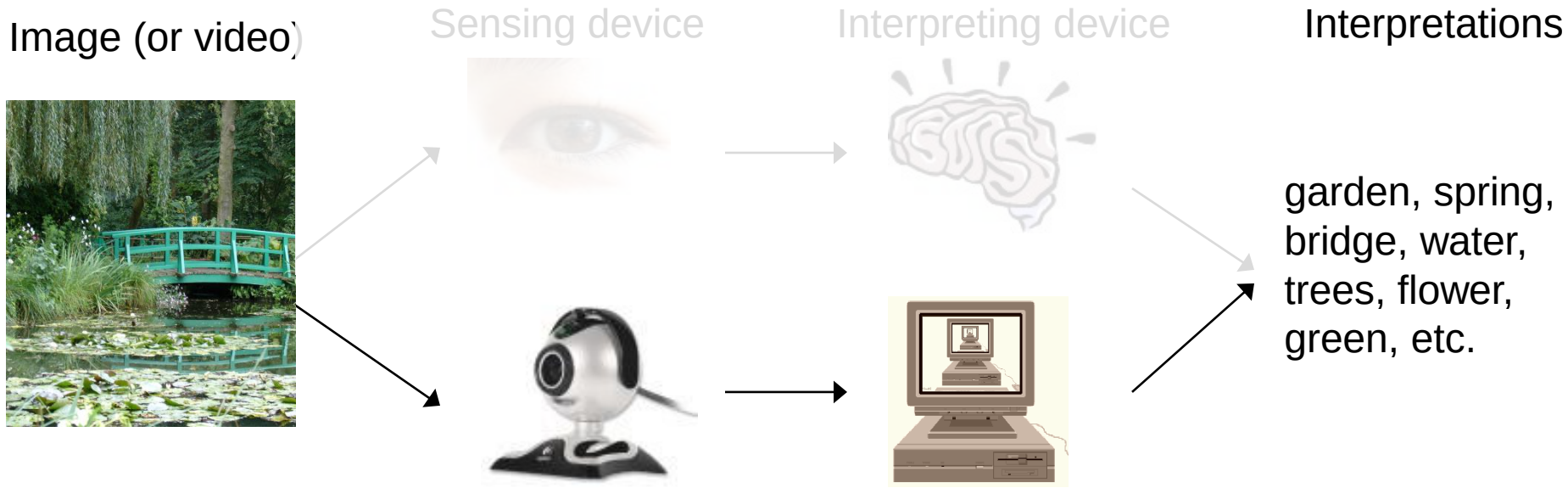
Perception



Edward H. Adelson

Motion without movement

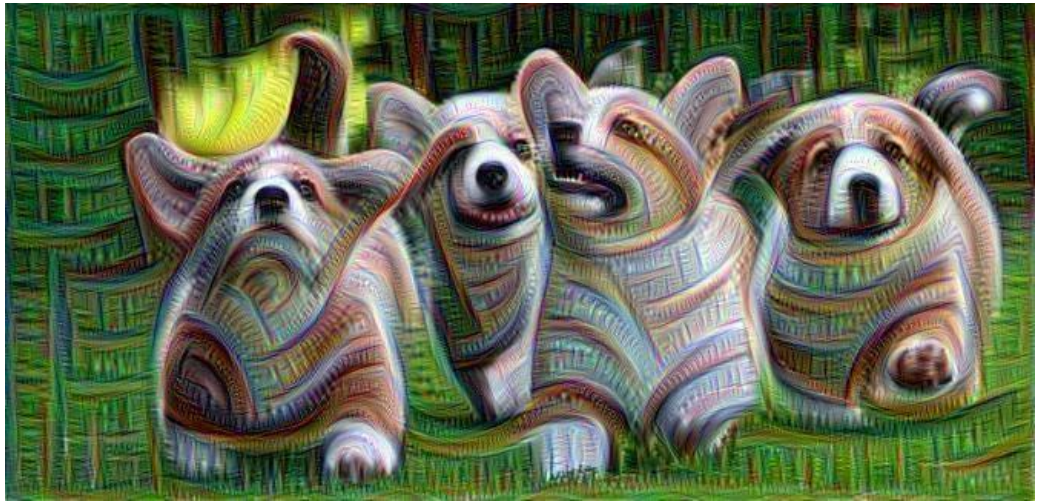
What is (computer) vision?

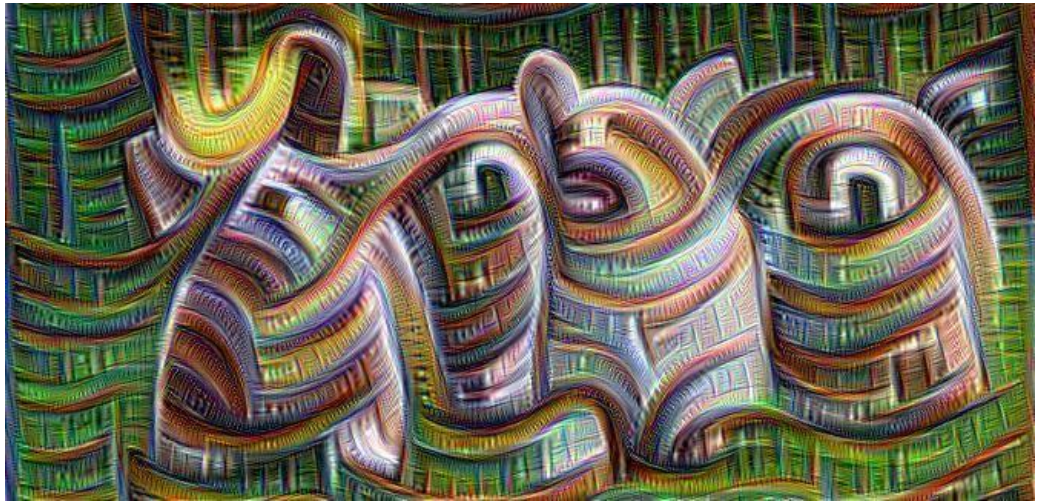


1837: Niépce, First photo of one's meal

1838: Boulevard du Temple, Daguerre

1838: First selfie, Robert Cornelius

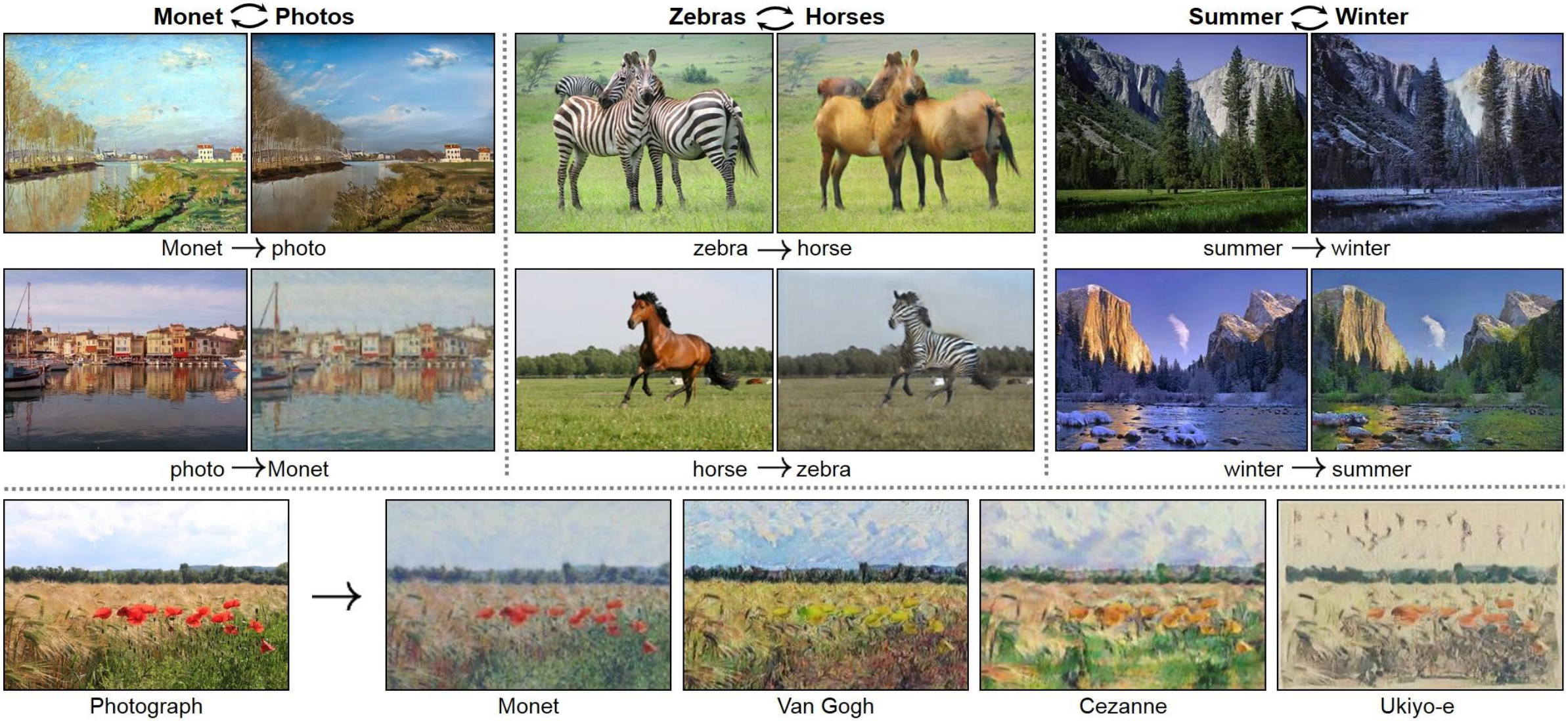




DeepDreams
[Mordvintsev et al. 2015]

Neural Style Transfer
[Gatys et al. 2015]

Neural Style Transfer
[Gatys et al. 2015]



CycleGAN [Zhu et al. 2017]

The goal of computer vision

- To bridge the gap between pixels and “meaning”

What we see

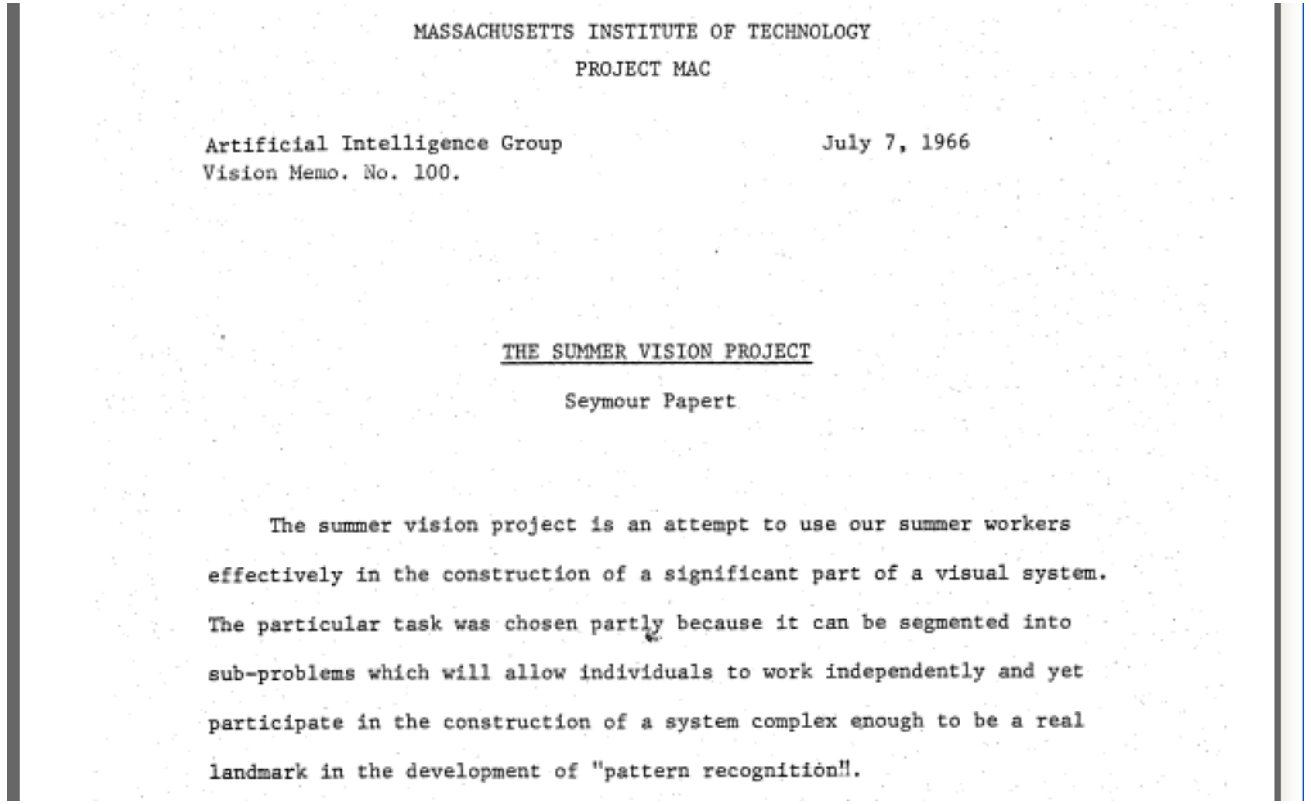
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

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

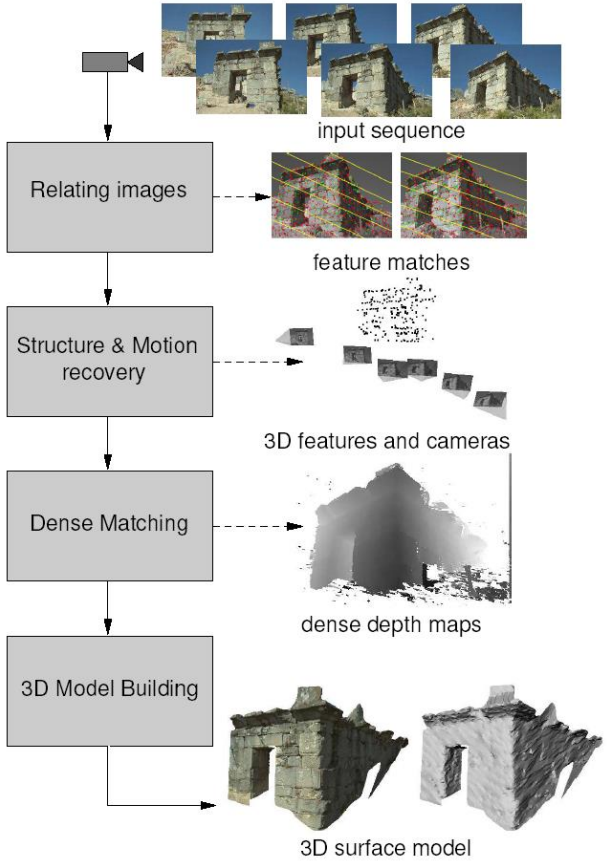
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

What kind of information can we extract from an image?

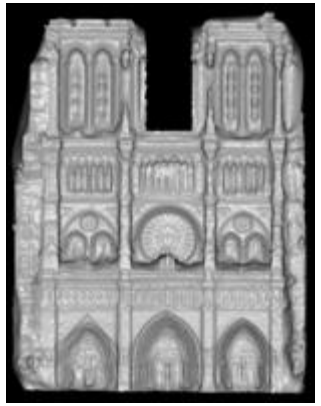
- Metric 3D information
- Semantic information

Vision as measurement device

Real-time stereo | Structure from motion | Reconstruction from



Pollefeys et al.



Goesele et al.

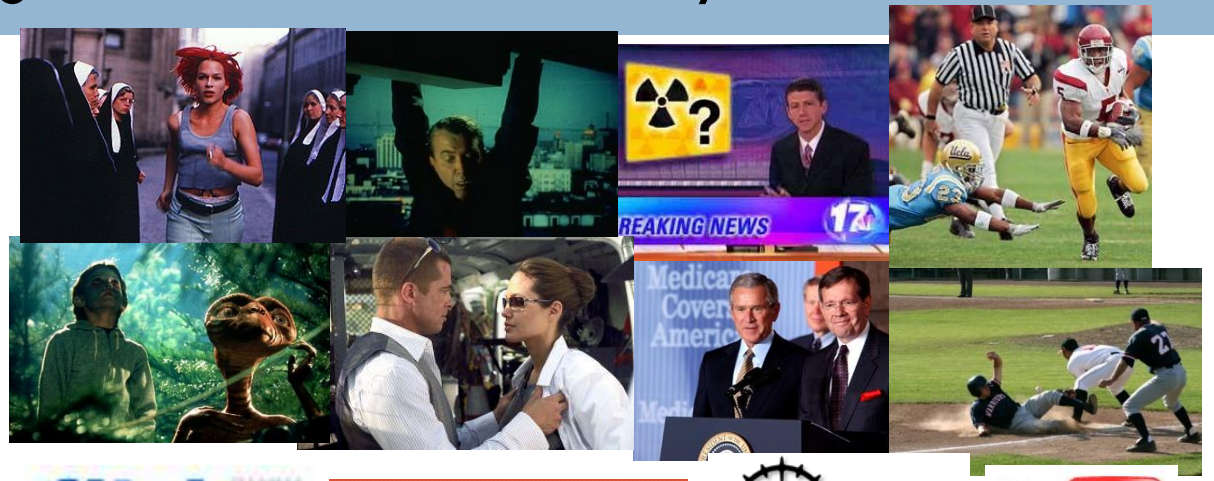
Objects
Activities
Scenes
Locations
Text / writing
Faces
Gestures
Motions
Emotions…

Vision as a source of semantic information

Slide credit: Kristen Grauman

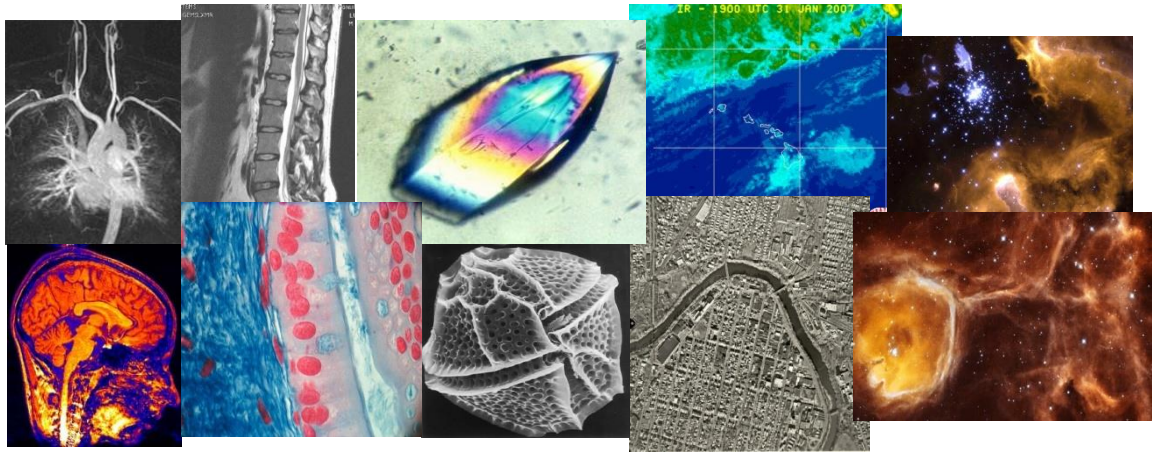
Why study computer vision?

- 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

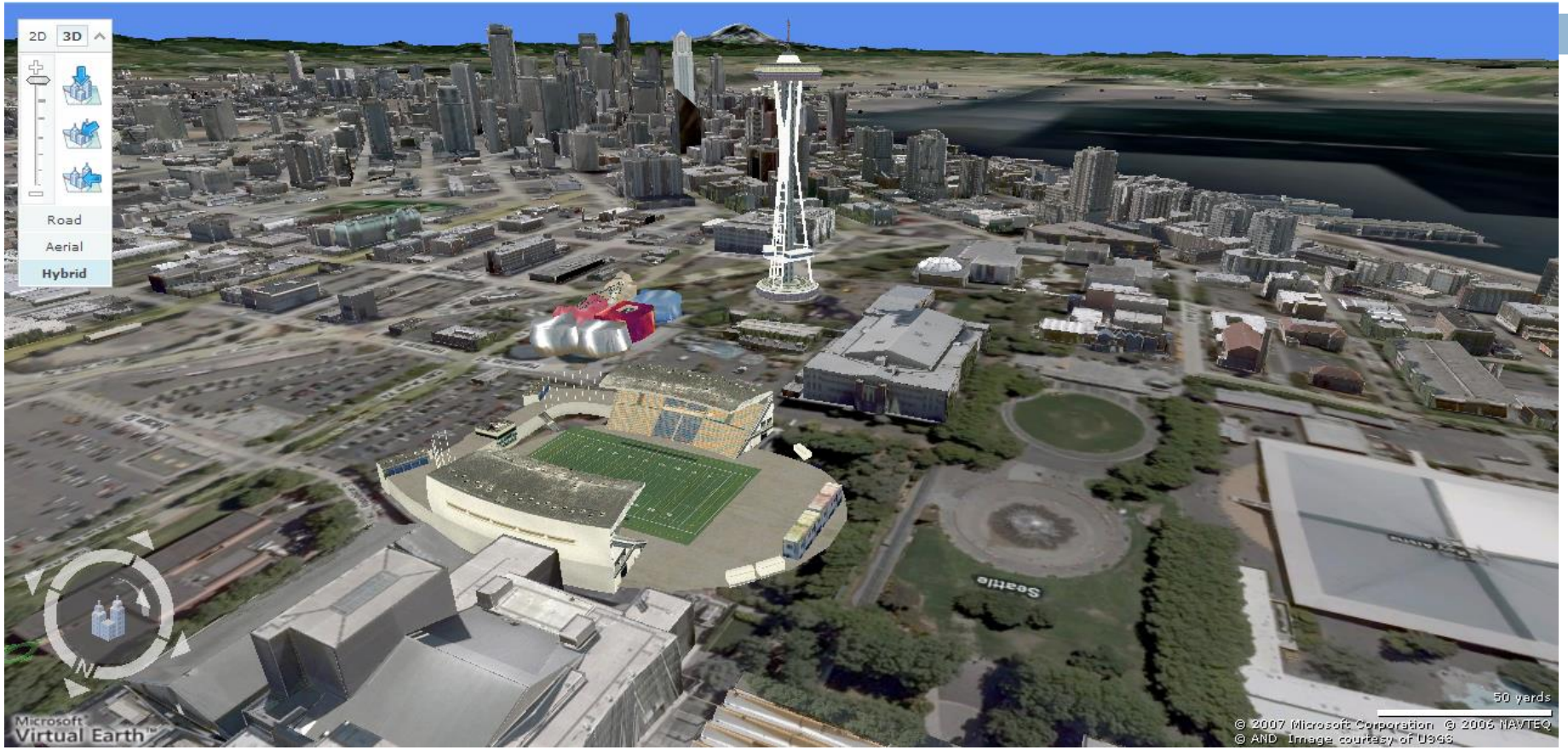
90% of all web traffic is images and videos

Majority of the internet is dark matter without computer vision

Special effects: shape and motion capture

Source: S. Seitz

3D urban modeling



Bing maps, Google Streetview

Source: S. Seitz

3D urban modeling: Microsoft Photosynth

http://photosynth.net

Source: S. Seitz

Face detection

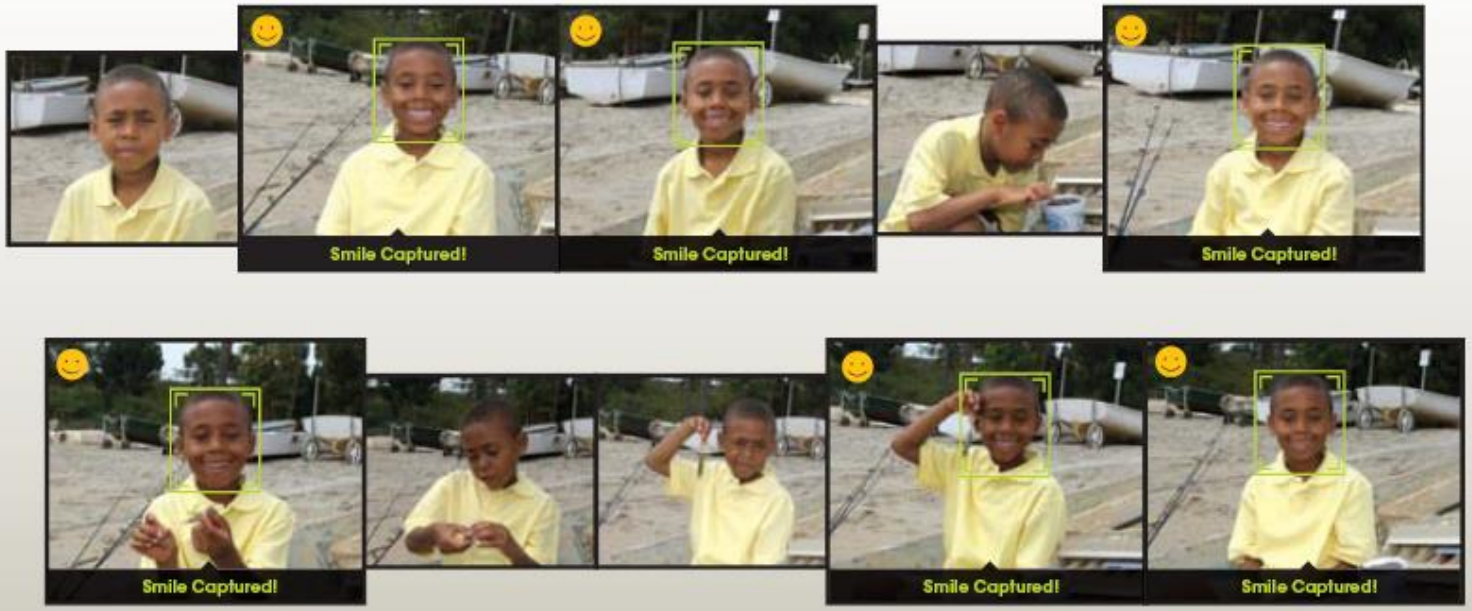
- Many digital cameras now detect faces
  - Canon, Sony, Fuji, …

Source: S. Seitz

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



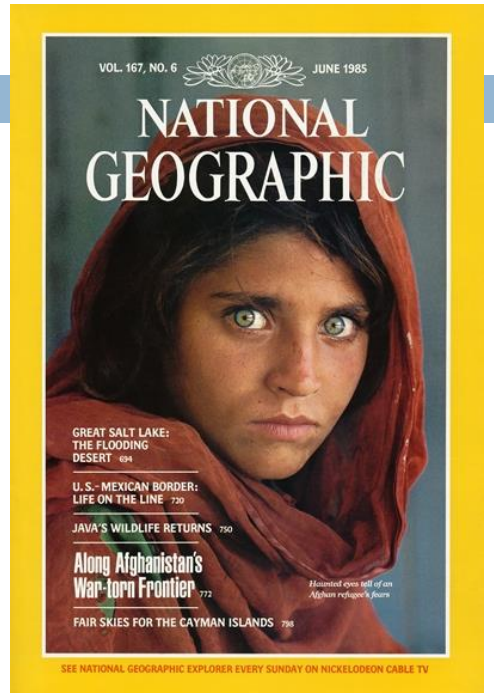
Sony Cyber-shot® T70 Digital Still Camera

Source: S. Seitz

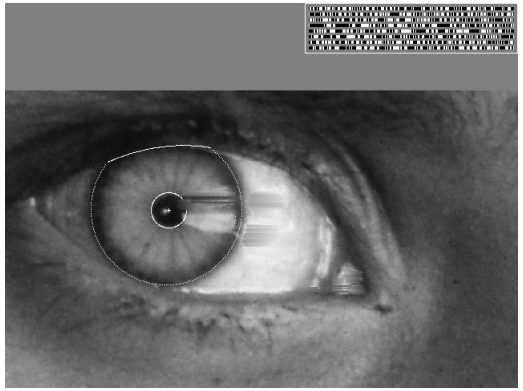
Face recognition: Apple iPhoto software

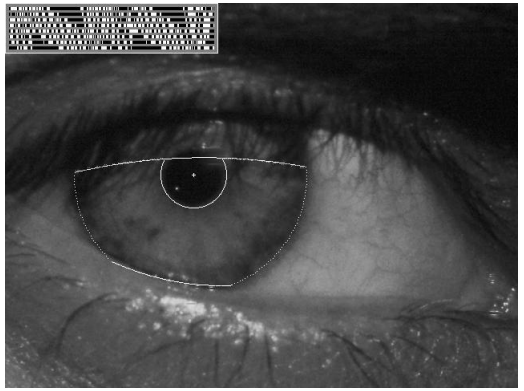
http://www.apple.com/ilife/iphoto/

Biometrics



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

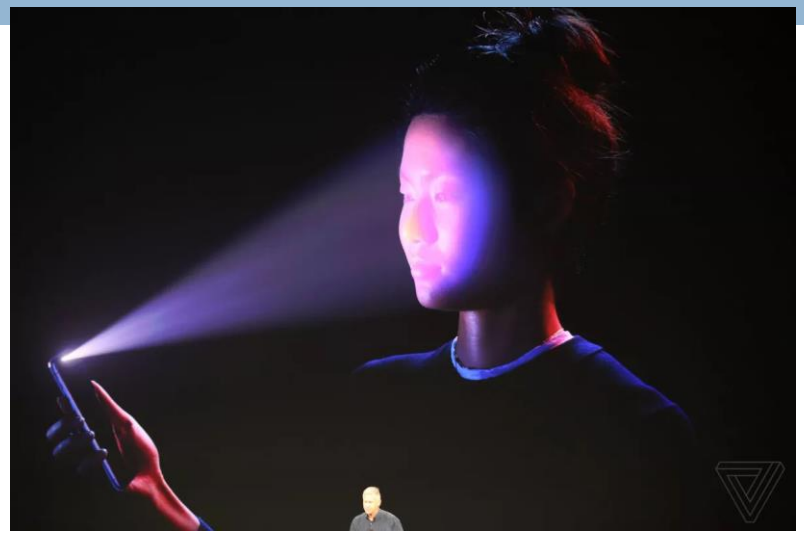




Source: S. Seitz

Biometrics

Fingerprint scanners on many new laptops, other devices



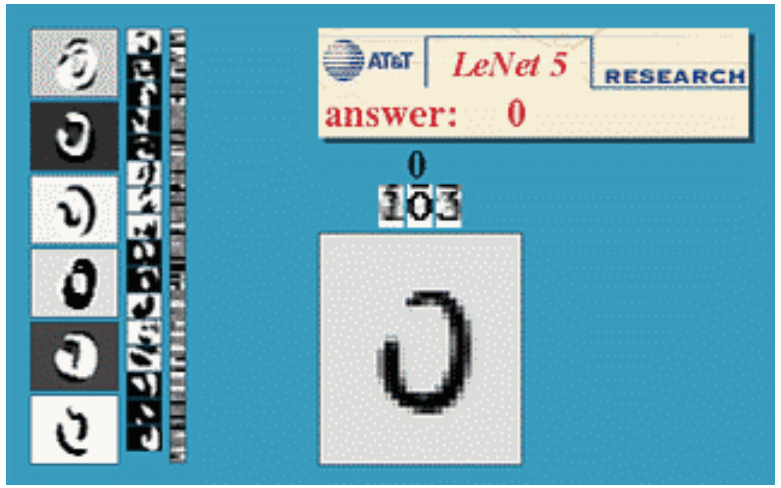
Face recognition systems now beginning to appear more widely

iphone X just introduced face recognition

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

Source: S. Seitz

Google maps: Annotate all houses and streets

Avenue des Sapins

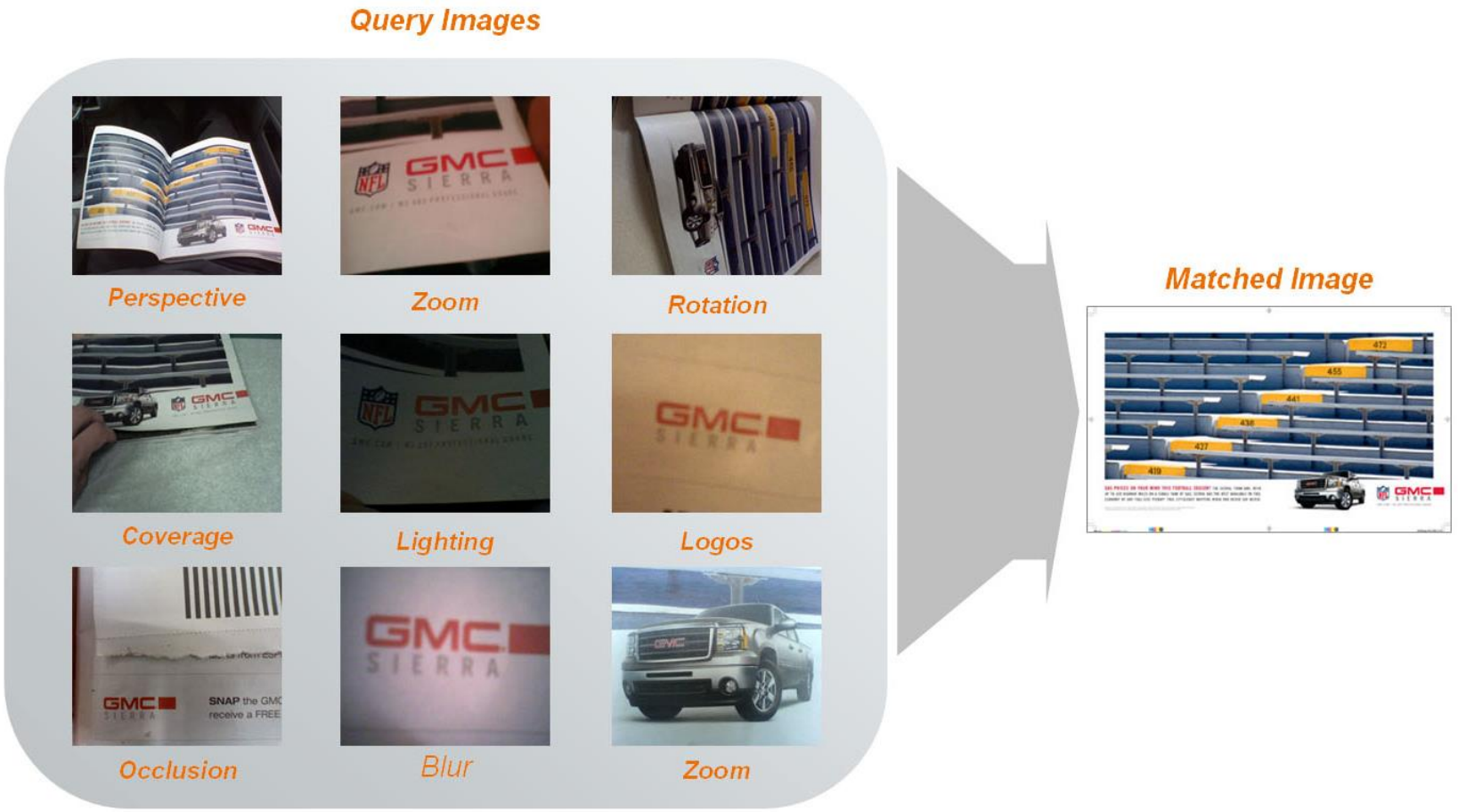
Goodfellow et al. 2014

Toys and Robots

Mobile visual search: iPhone Apps

snaptell

kooaba

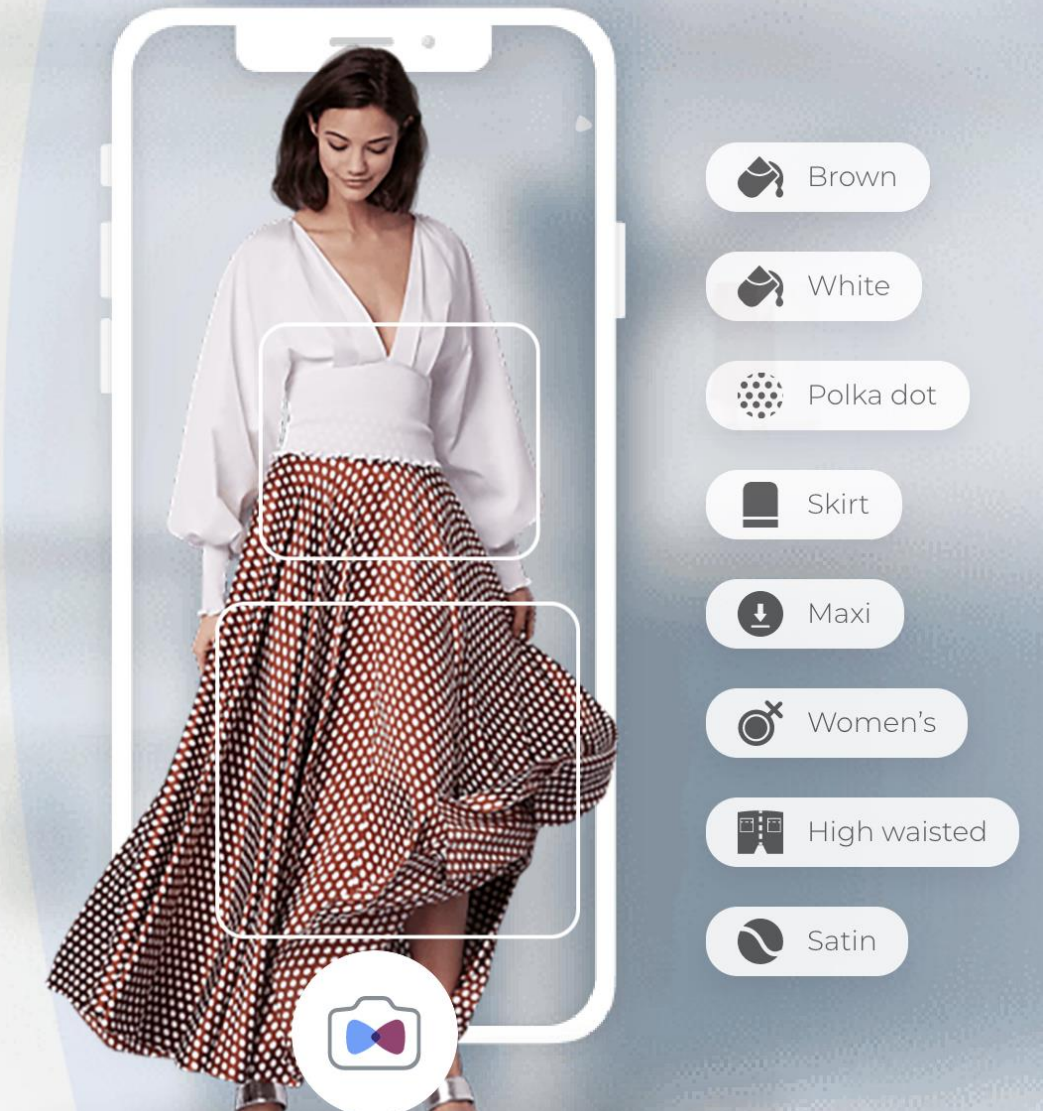


The Leader in Visual AI for Retail

Syte changes the way retailers connect shoppers with the products that inspire them by delivering the best Visual AI technology for retail. Discover our solutions that empower retailers to increase customer engagement, and boost conversion and sales.

Enter Business Email See Syte in Action!

Watch Video



TOMMY HILFIGER FARFETCH bonprix SHOPSTYLE

Snapstacles and Google glasses

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."

Source: A. Shashua, S. Seitz

Vision in supermarkets

LaneHawk by EvolutionRobotics

"A smart camera is flush-mounted in the checkout 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

Vision-based interaction (and games)

Microsoft's Kinect

Sony EyeToy

Assistive technologies

Source: S. Seitz

Augmented Reality

Virtual Reality

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

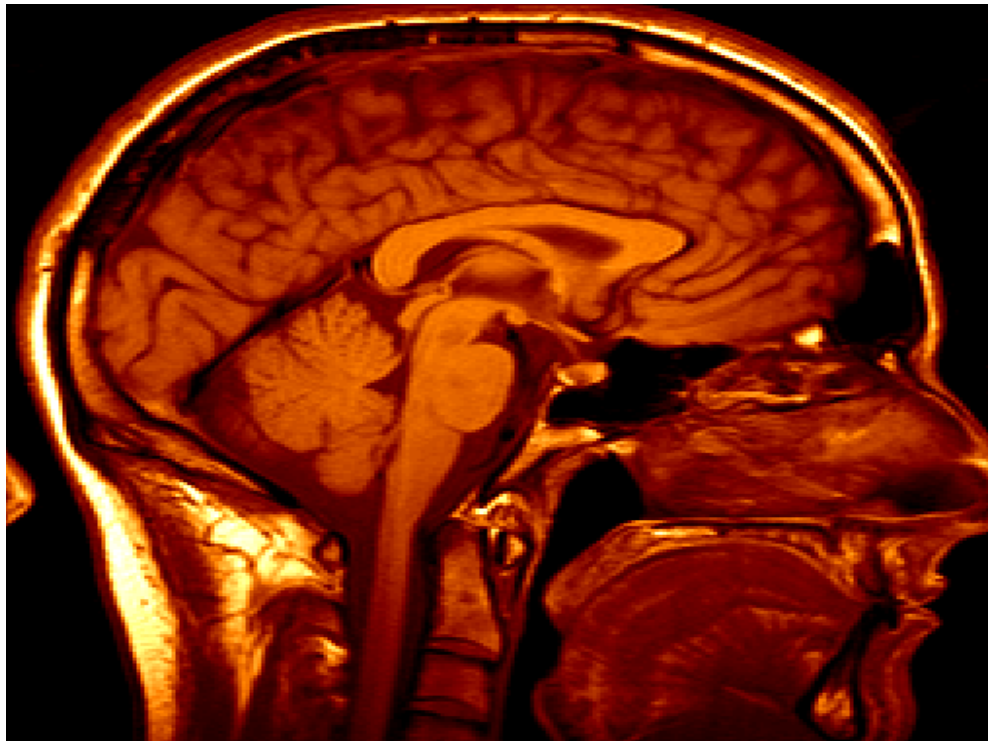
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Source: S. Seitz

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

Medical imaging



3D imaging
MRI, CT

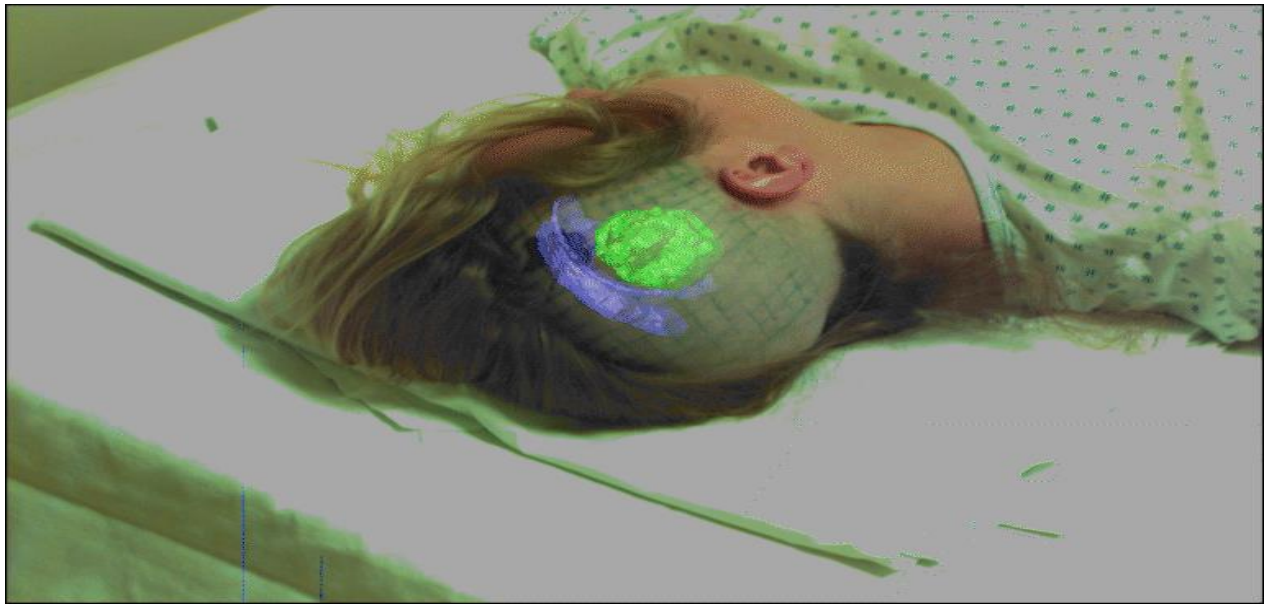


Image guided surgery
Grimson et al., MIT

Computer vision in sports

Hawk-Eye: helping/improving referee decisions

Computer vision in sports

SportVision: improving viewer experiences

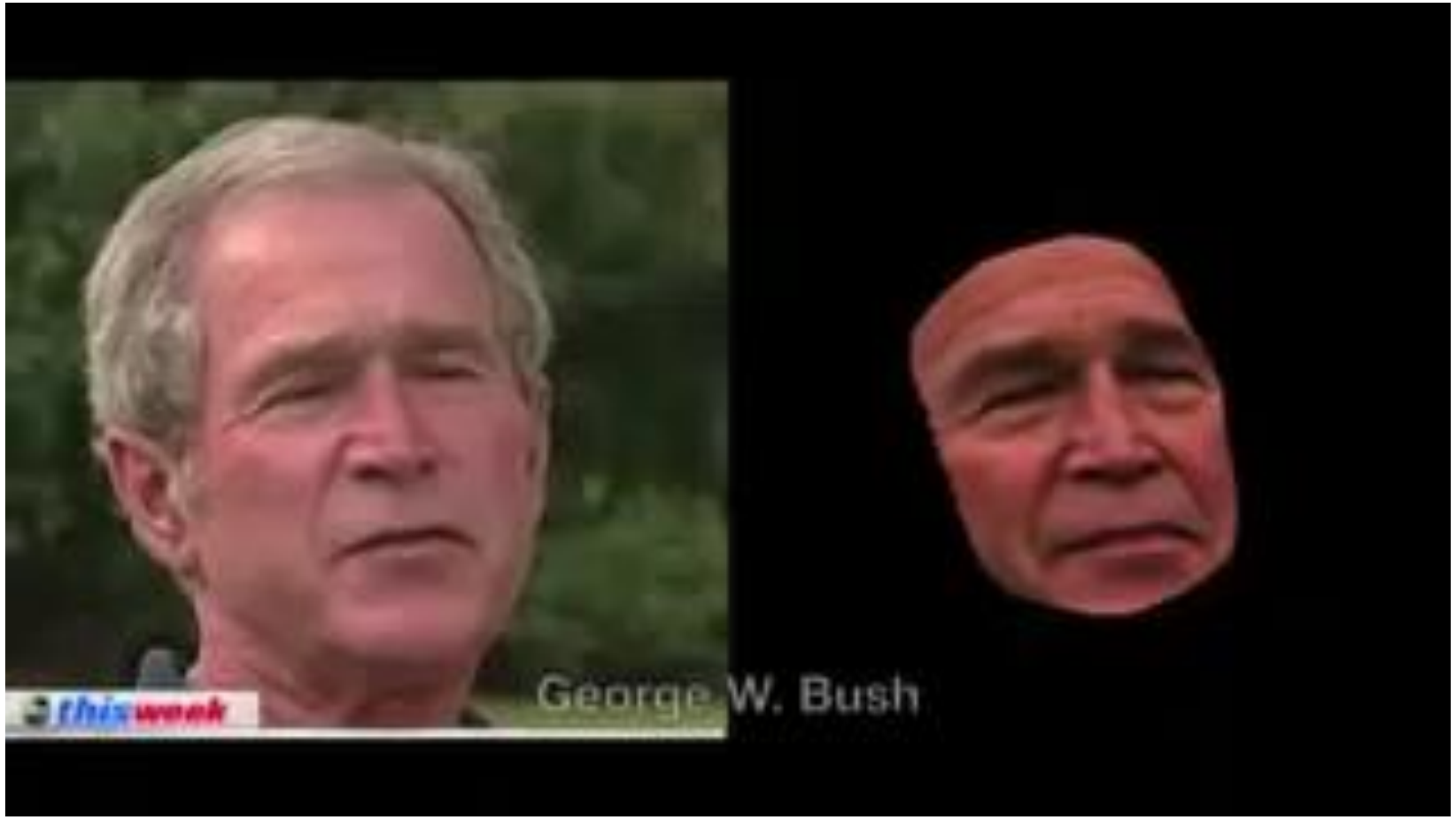
Computer vision in sports

Replay Technologies: improving viewer experiences

Computer vision in sports

Play tracking

Face Landmark Alignment – 3D Persona



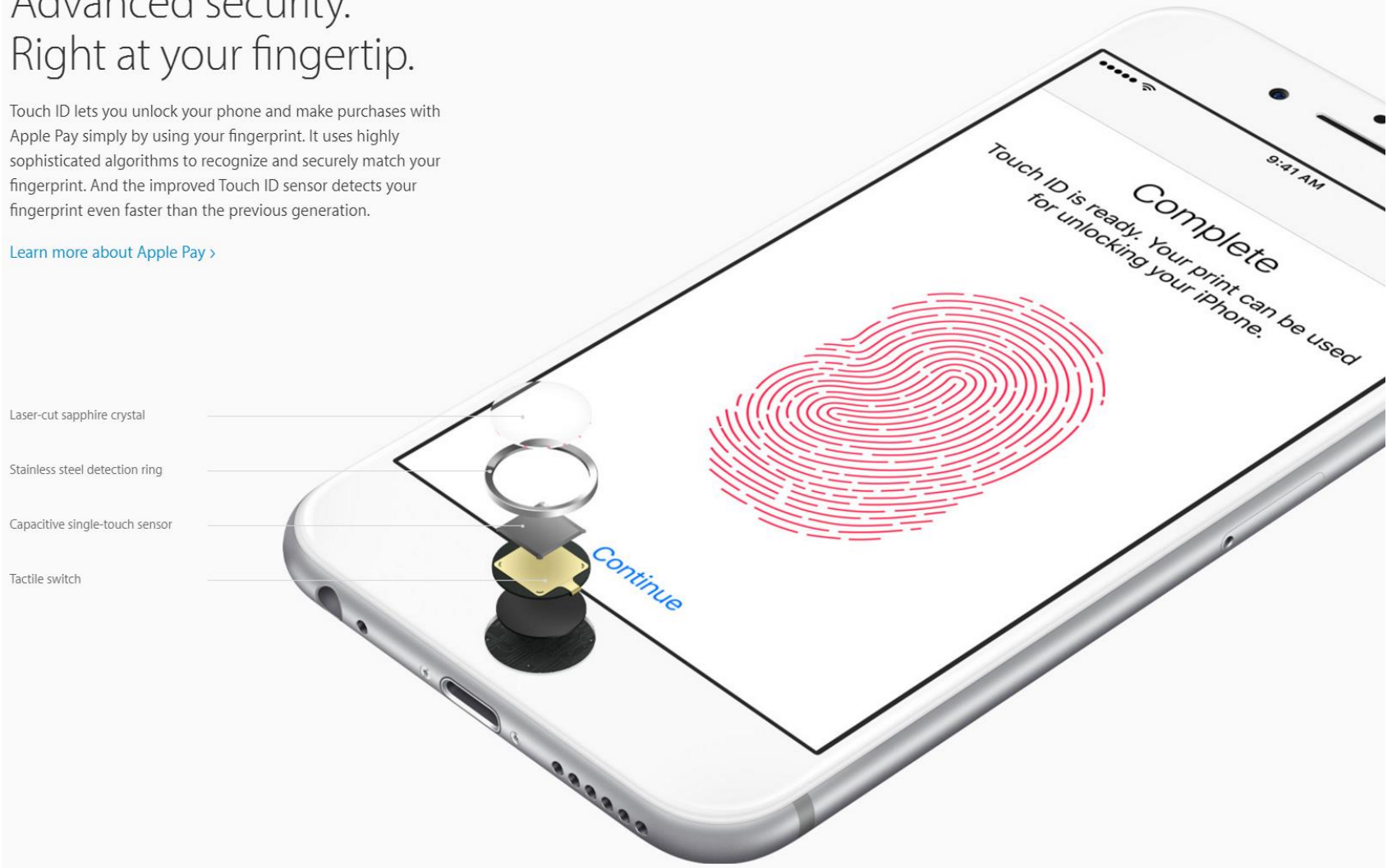
What Makes Tom Hanks Look Like Tom Hanks ICCV 2015

Vision-based Biometrics

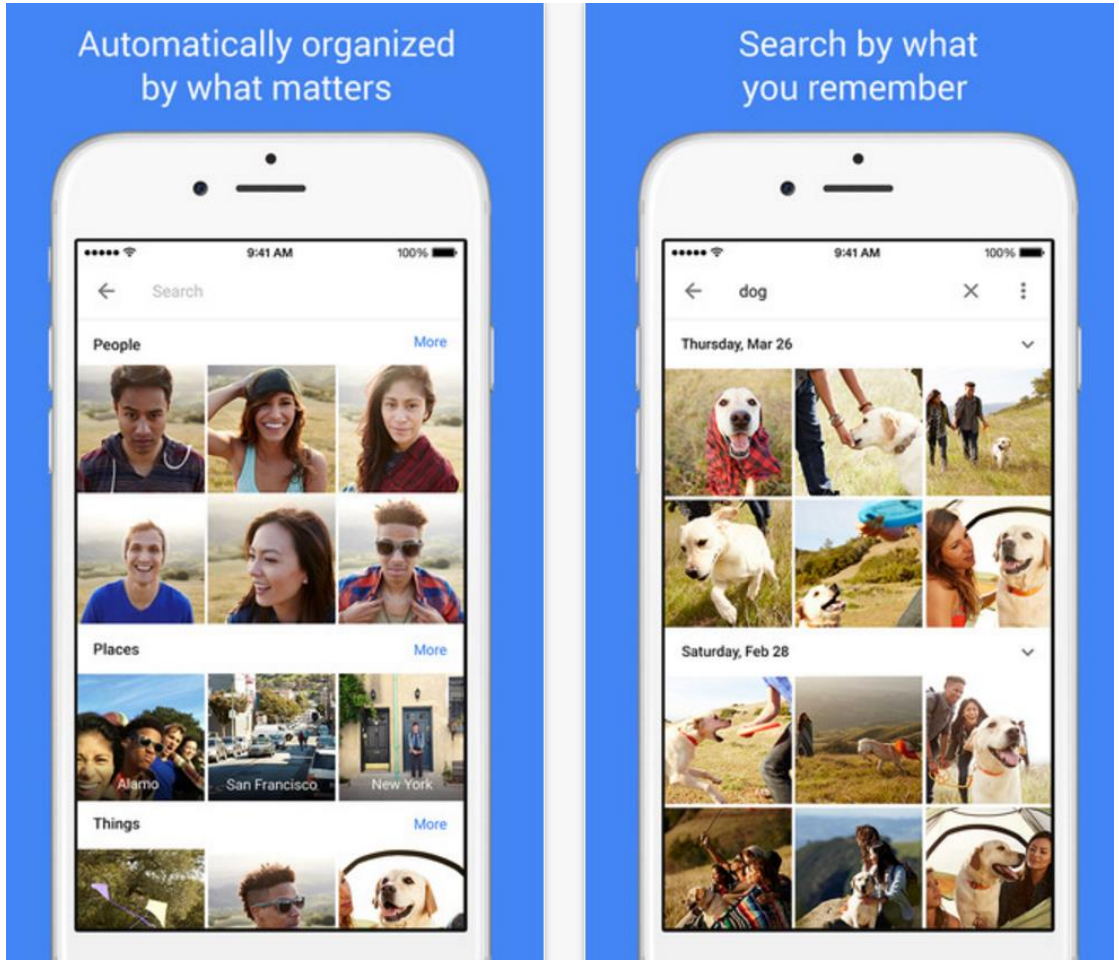
Touch ID.
Advanced security.
Right at your fingertip.

Touch ID lets you unlock your phone and make purchases with Apple Pay simply by using your fingerprint. It uses highly sophisticated algorithms to recognize and securely match your fingerprint. And the improved Touch ID sensor detects your fingerprint even faster than the previous generation.

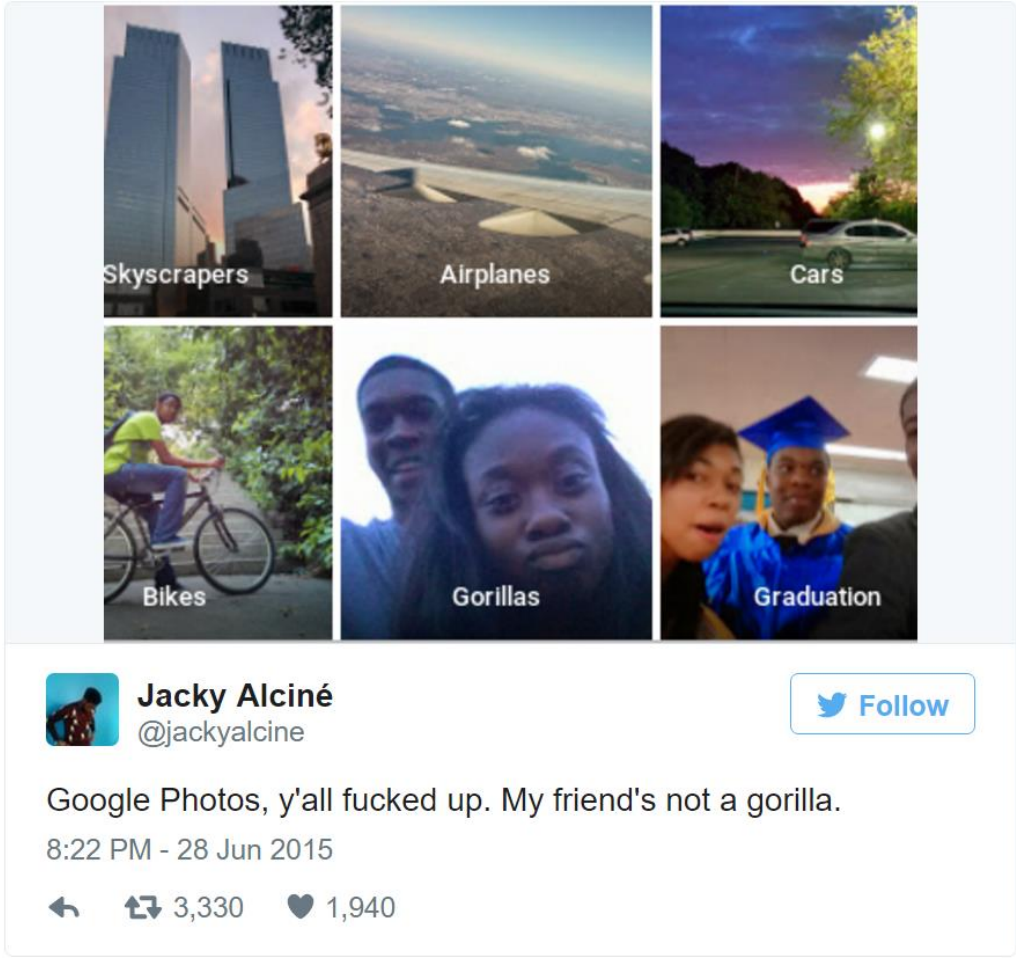
Learn more about Apple Pay ›



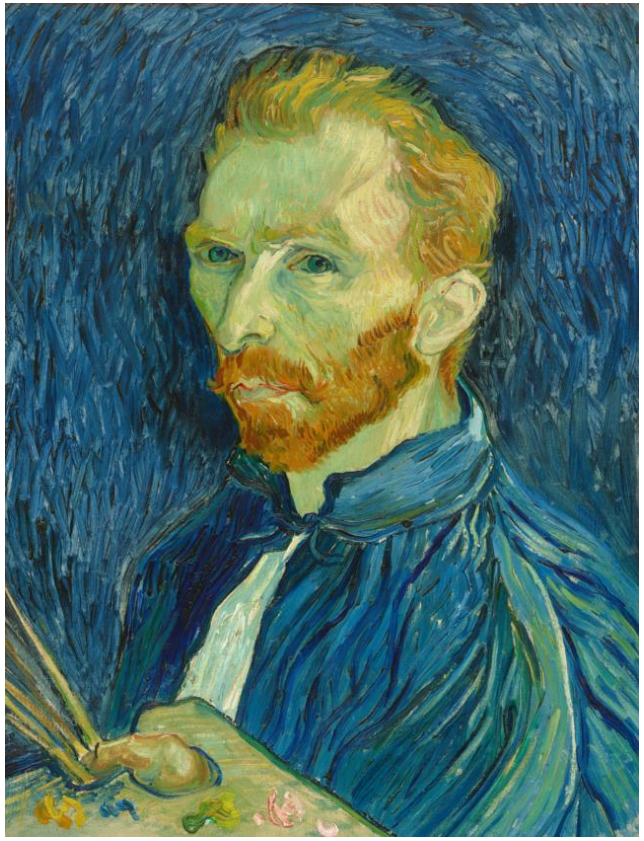
Visual recognition for photo organization



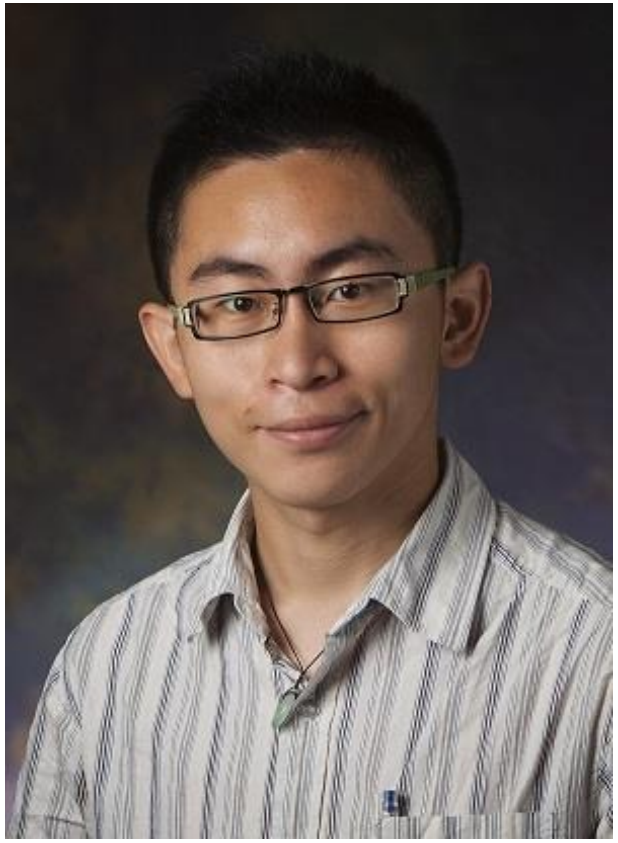
Google photo



Style transfer



Source image (**Style**)



Target image (**Content**)

Output (deepart)

A Neural Algorithm of Artistic Style [Gatys et al. 2015]

Special effects: Matting and composition

Kylie Minogue - Come Into My World

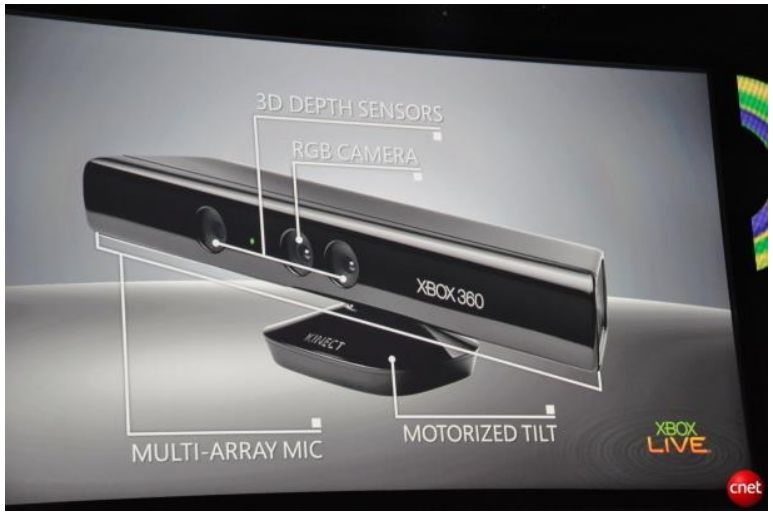
Google cars

Google in talks with Ford, Toyota and Volkswagen to realise driverless cars

http://www.theatlantic.com/technology/archive/2014/05/all-the-world-a-track-the-trick-that-makes-googles-self-driving-cars-work/370871/

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



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

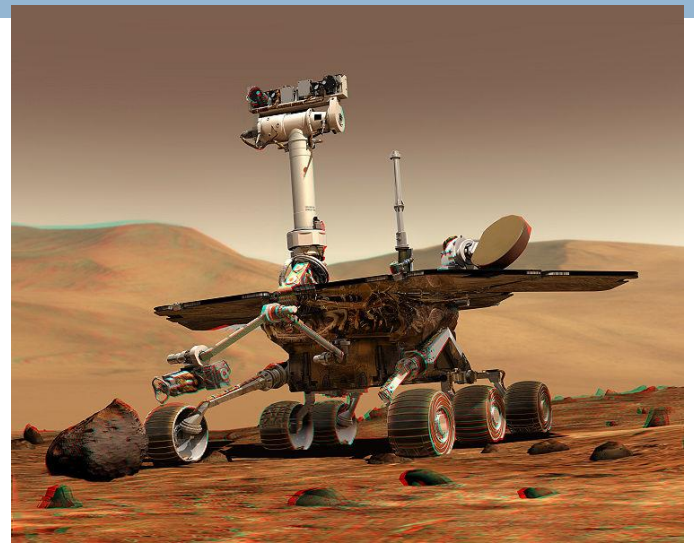
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Industrial robots

Vision-guided robots position nut runners on wheels

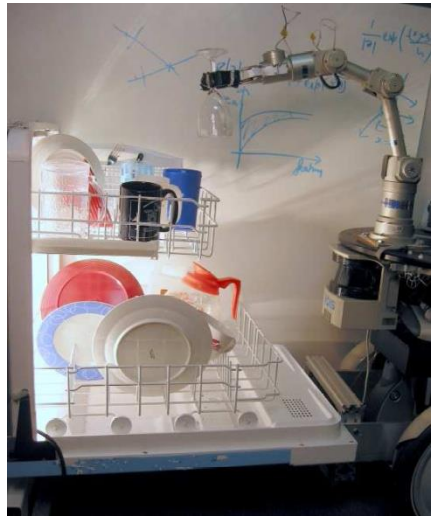
http://www.automationworld.com/computer-vision-opportunity-or-threat

Mobile robots



NASA's Mars Spirit Rover

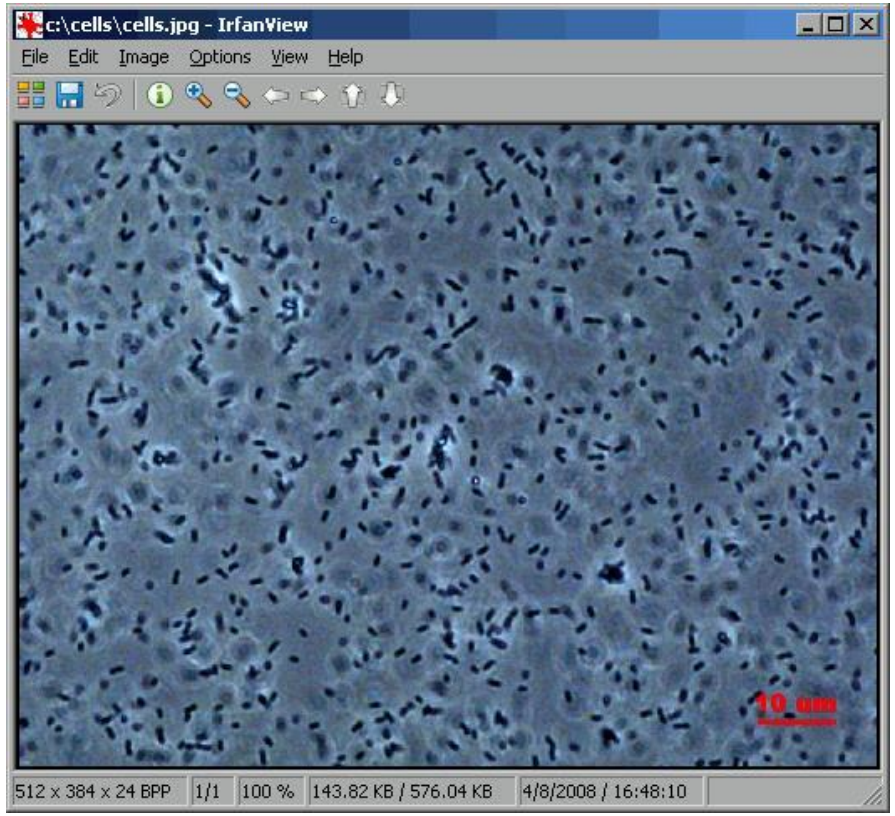
http://www.robocup.org/



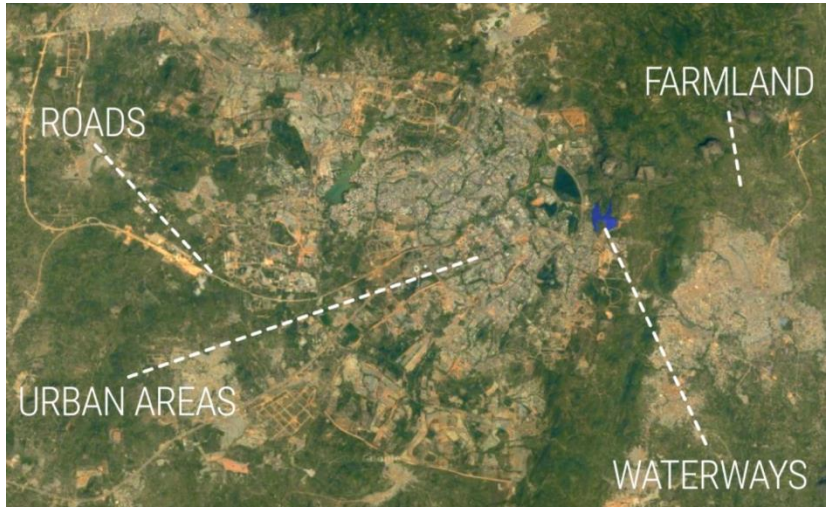
Saxena et al. 2008
STAIR at Stanford

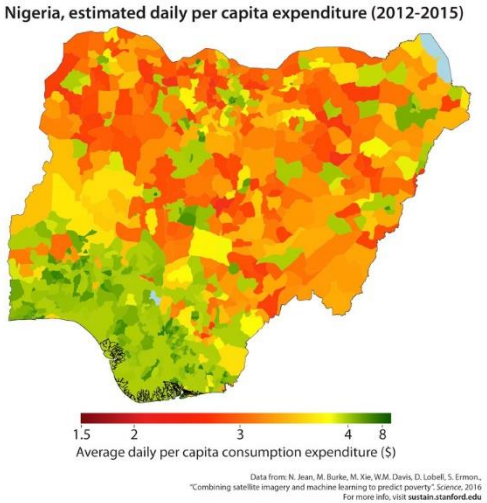
http://www.youtube.com/watch?v=DF39Ygp53mQ

Computer vision for the mass



Counting cells





Predicting poverty

Current state of the art

- Many of these are less than 5 years old
- Very active and exciting research area!
- To learn more about vision applications and companies
  - David Lowe maintains an excellent overview of vision companies
    - http://www.cs.ubc.ca/spider/lowe/vision.html

MGMT "When You Die"