COMPUTER VISION

Pascal Fua
EPFL CVLab
CH-1015 Lausanne
Switzerland
http://cvlab.epfl.ch/~fua

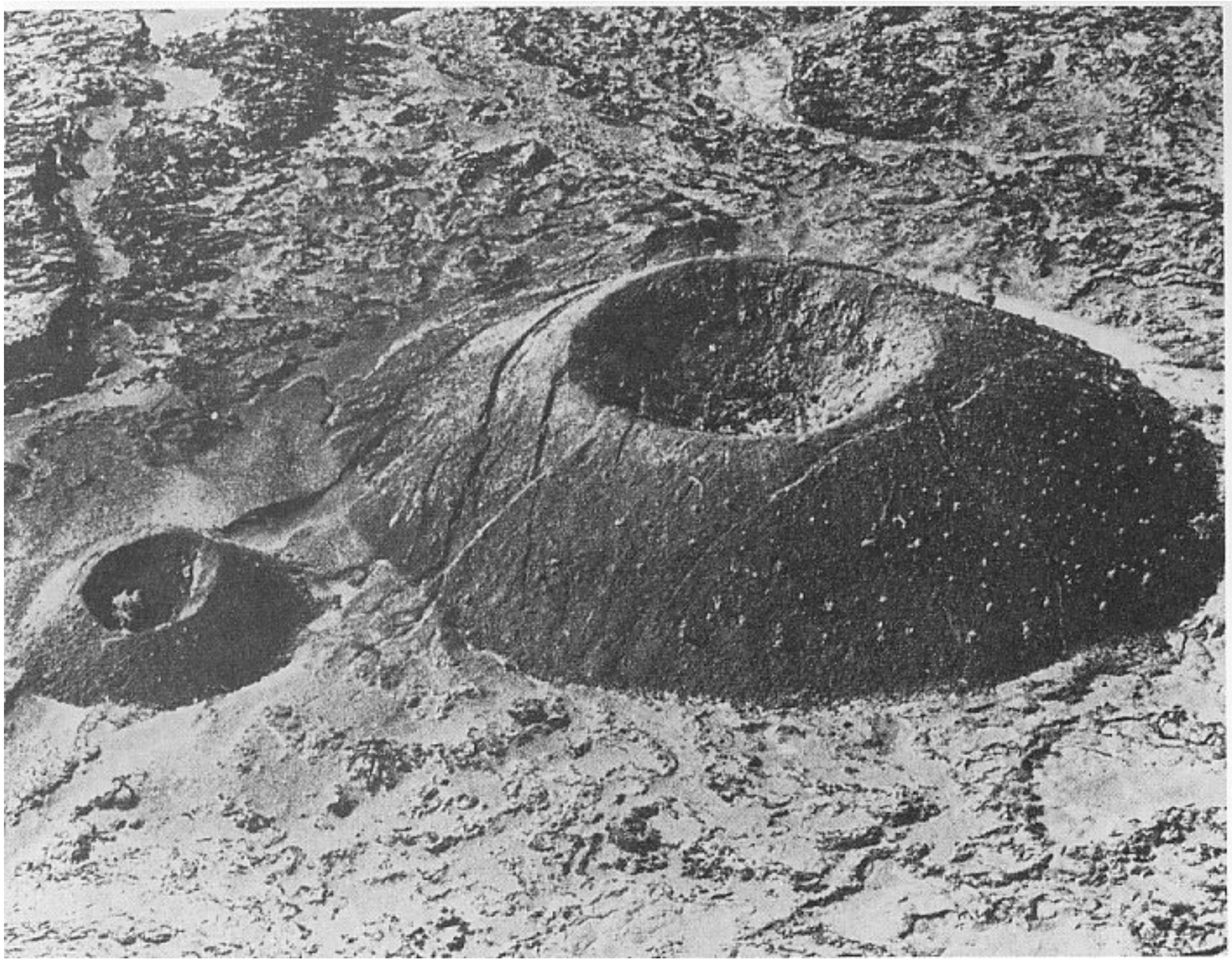
COMPUTER VISION

Goal: Inferring the properties of the world from one or more images

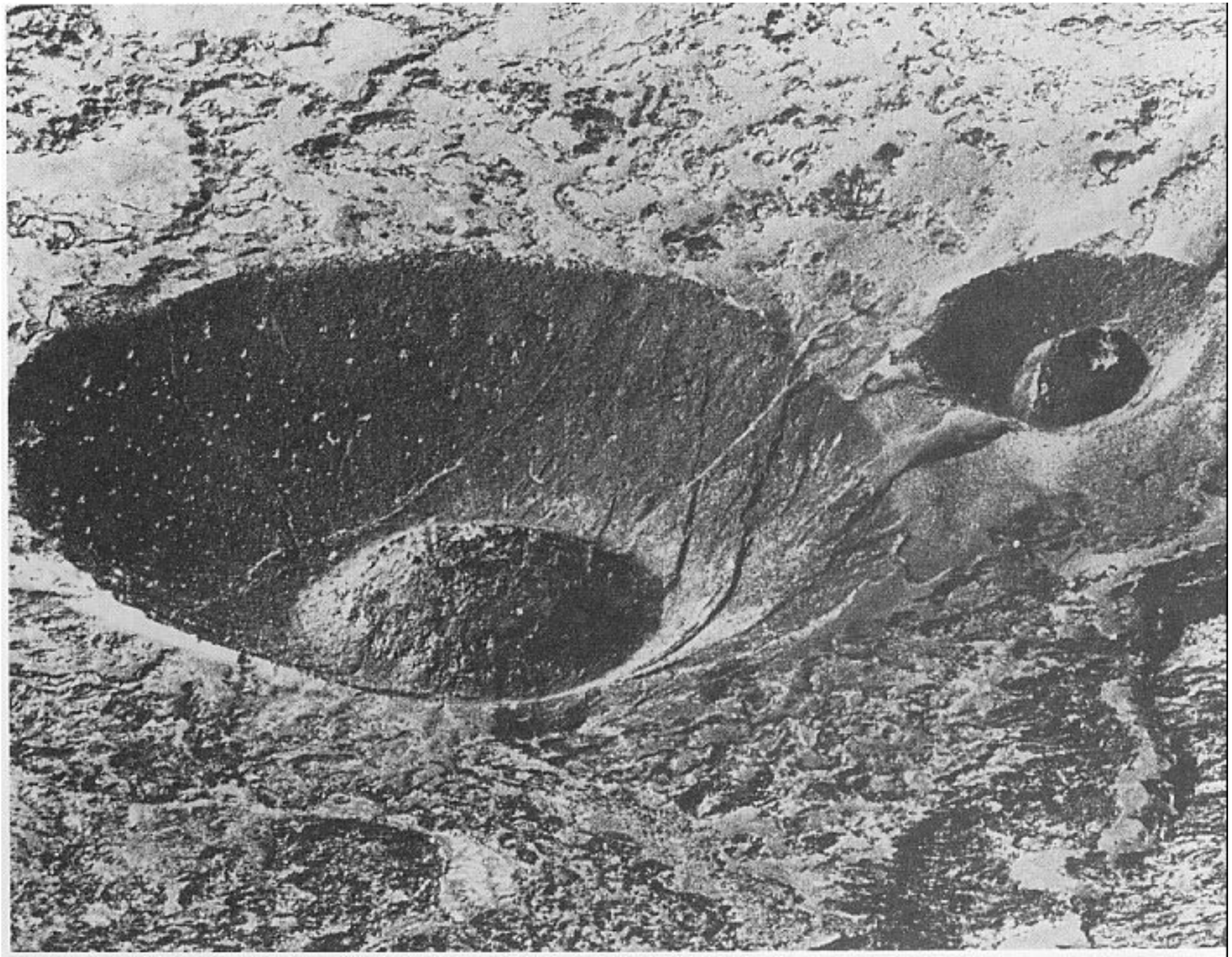
- Photographs
- Video Sequences
- Medical images
- Microscopy

→ **Image Understanding**

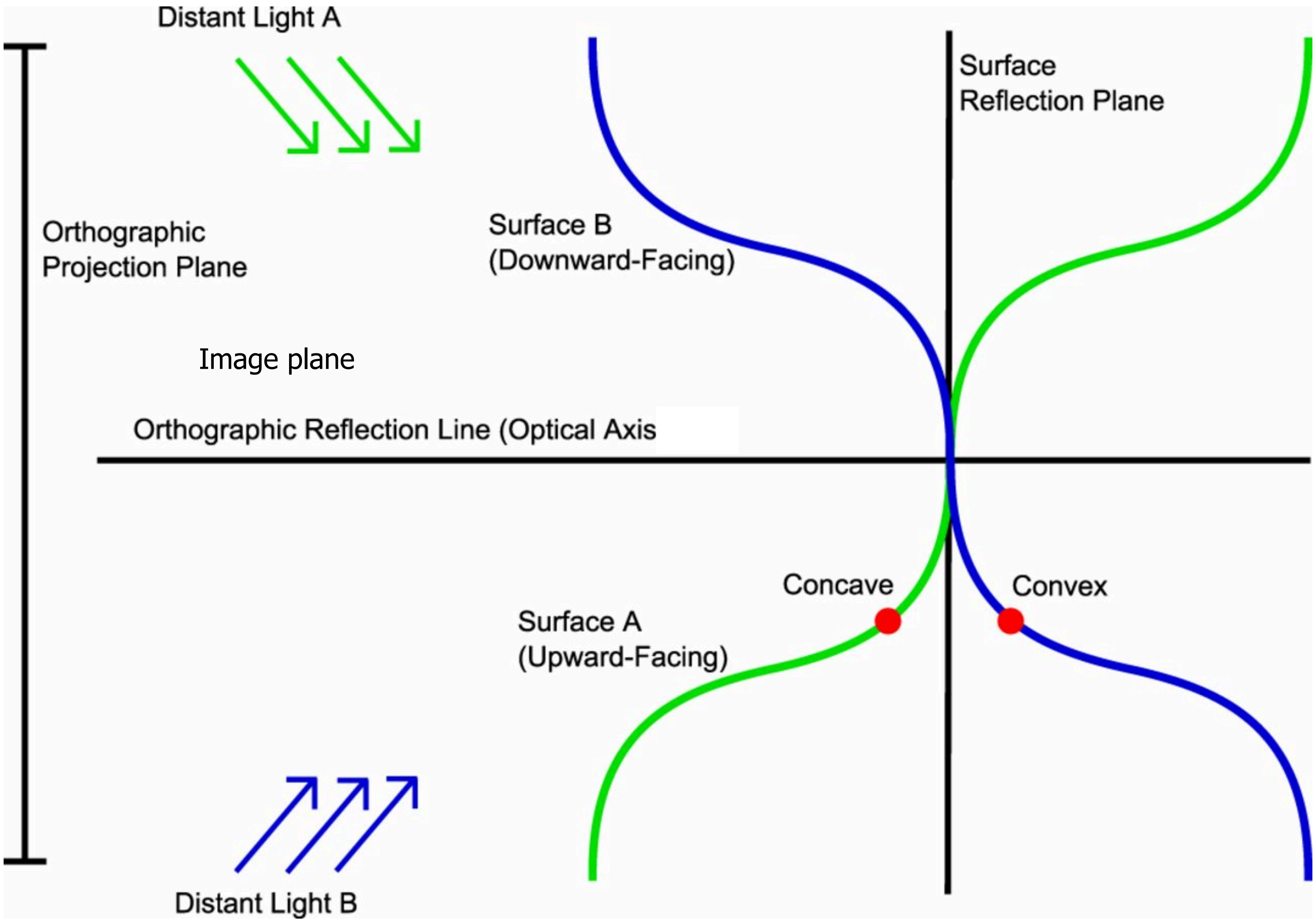
WHAT DO YOU SEE?



AND NOW?



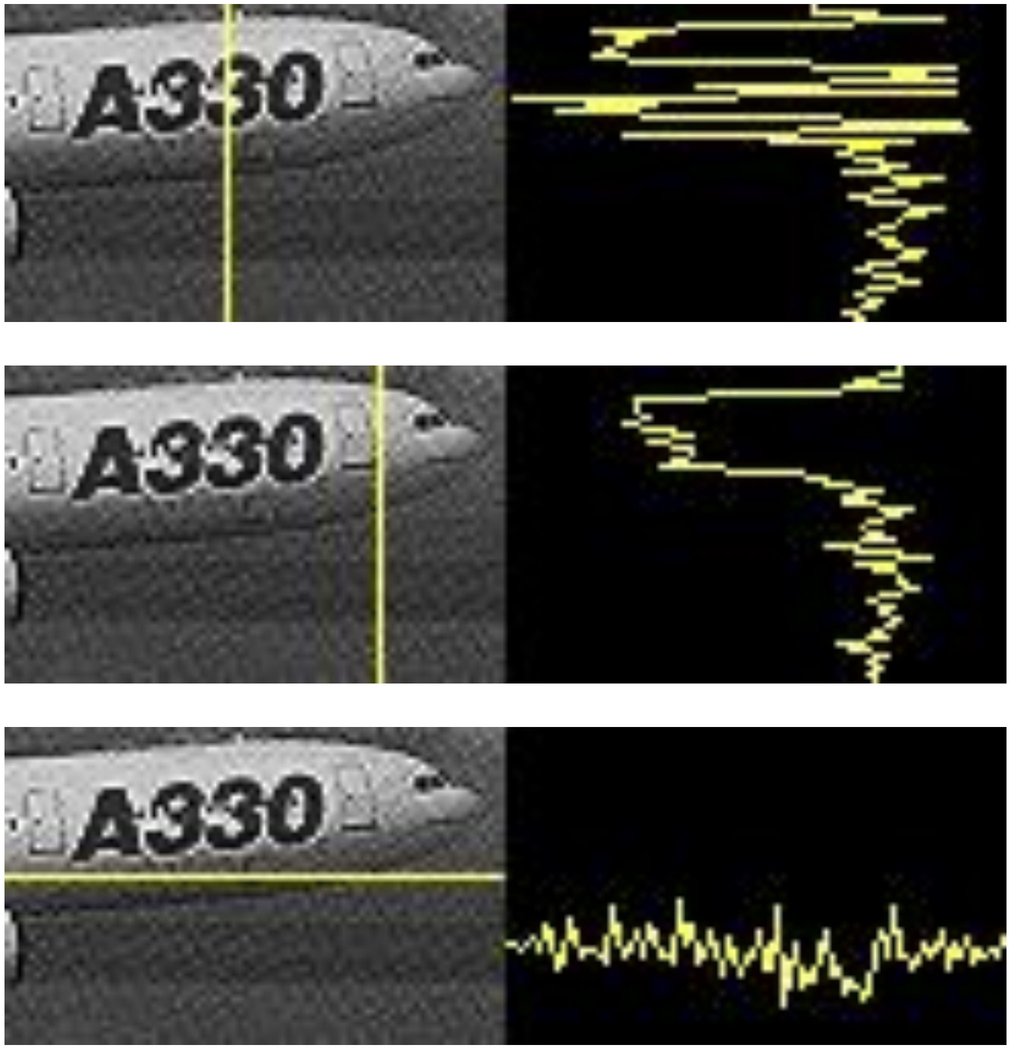
CONCAVE vs CONVEX



A POWERFUL MECHANISM

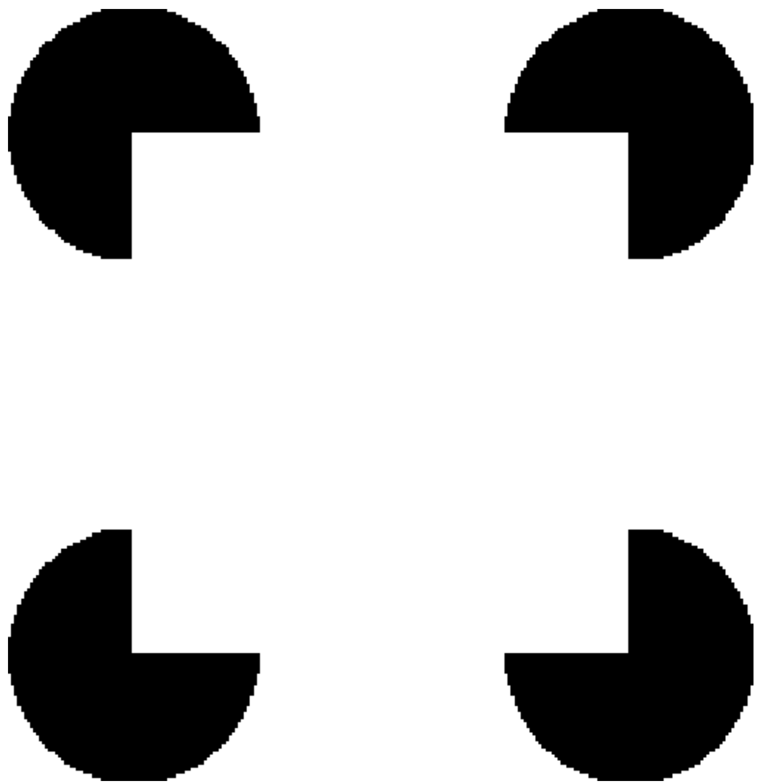
136	134	161	159	163	168	171	173	173	171	166	159	157	155
152	145	136	130	151	149	151	154	158	161	163	163	159	151
145	149	149	145	140	133	145	143	145	145	145	146	148	148
148	143	141	145	145	145	141	136	136	135	135	136	135	133
131	131	129	129	133	136	140	142	142	138	130	128	126	120
115	111	108	106	106	110	120	130	137	142	144	141	129	123
117	109	098	094	094	094	100	110	125	136	141	147	147	145
136	124	116	105	096	096	100	107	116	131	141	147	150	152
152	152	137	124	113	108	105	108	117	129	139	150	157	159
159	157	157	159	135	121	120	120	121	127	136	147	158	163
165	165	163	163	163	166	136	131	135	138	140	145	154	163
166	168	170	168	166	168	170	173	145	143	147	148	152	159
168	173	173	175	173	171	170	173	177	178	151	151	153	156
161	170	176	177	177	179	176	174	174	176	177	179	155	157
161	162	168	176	180	180	180	182	180	175	175	178	180	180

A POWERFUL MECHANISM



CONTEXT & MODELS

ILLUSORY CONTOURS



KANIZSA'S TRIANGLE

Closure of Good Form Hypothesis: Illusory contours represent an example of the closure of good form (Osgood 1953), (Pastore 1971), (Kanizsa 1976, 1979).

Figural-Cue Hypothesis: Illusory contours are responses to partial figural cues in the same way that meaning is abstracted from simple outline drawings or cartoons (Gregory 1972), (Piggins 1975), (Rock and Anson 1979).

Cues-to-Depth Hypothesis: Illusory contours are produced by the monocular depth cue of interposition to perceive a plane in depth (Coren 1972).

Organizational-Attentional Effects Hypothesis: Emphasizes that illusory contours are not totally stimulus-bound (Bradley and Dumais 1975), (Bradley and Petry 1977), (Kennedy 1976).

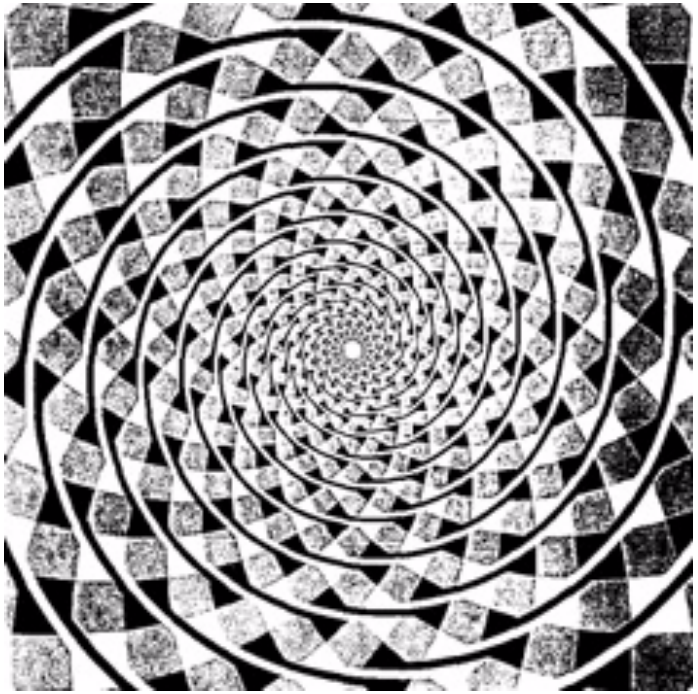
Retinal-Smearing Hypothesis: The edge effects created by the inducing areas are smeared over the retina during the course of normal eye movements to produce illusory contours (Kennedy and Chattaway 1975). This theory has been found to be untenable.

Brightness-Contrast Hypothesis: Illusory contour formation is secondary to the perception of brightness differences between the illusory figure and its background (Brigner and Gallagher 1974), (Frisby and Clatworthy 1975), (Day and Jory 1978,1979,1980).

Feature Analyzers Hypothesis: Illusory contours result from the partial triggering of contour-specific neural units by the physically present edge along the inducing areas (Stadler and Dieker 1972), (Smith and Over 1975,1979). Others have suggested neural networks that generate continuous contours (filling-in contours) from discontinuous stimulus (Ullman 1976), (Grossberg and Mingolla 1985).

Spatial-Frequency-Analysis Hypothesis: Existence of a stimulus correlate for illusory contours based on a Fourier analysis of the stimulus display (Ginsburg 1975) .

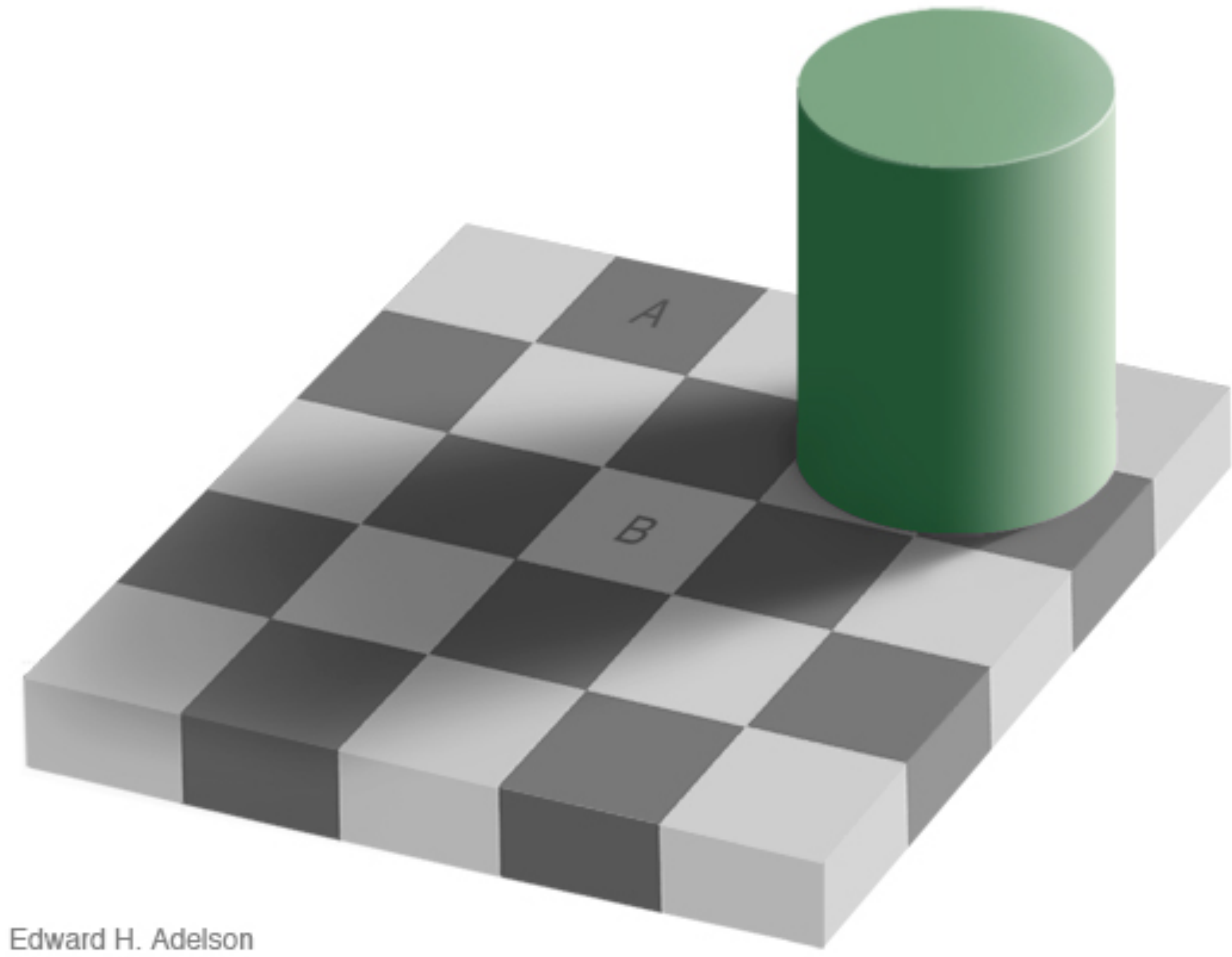
OPTICAL ILLUSIONS



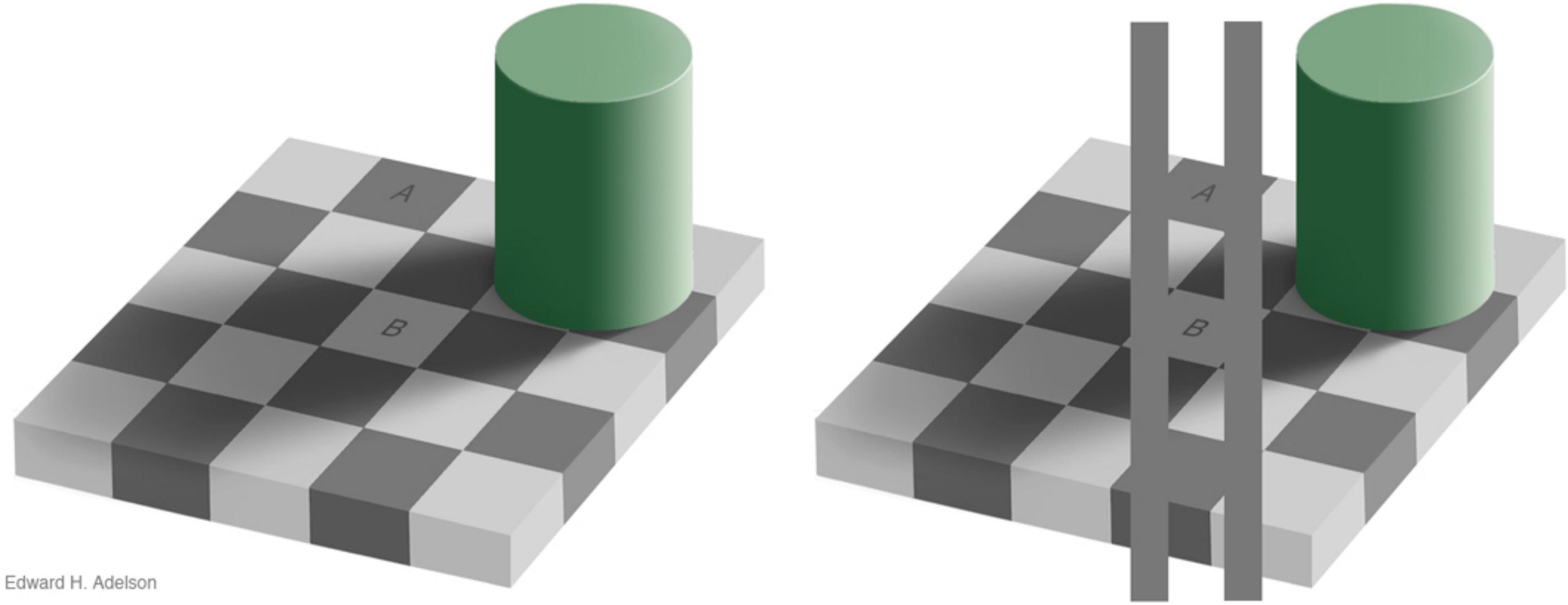
Every image is the image of thing merely to him who knows how to read it, and who is enabled by the aid of the image to form an idea of the thing.

Handbook of Physiological Optics
H. von Helmholtz

PHOTOMETRIC ILLUSION



NO MORE ILLUSION



The human eye measures relative rather than absolute intensity values.

CHALLENGES

Vision involves dealing with:

- Noisy images
- Many-to-one mapping
- Aperture problem

→ Requires:

- Assumptions about the world
- Object models
- Training data

COMPUTERS vs HUMANS

True image understanding seems to involve a great deal of human intelligence:

- Automated systems are still very far from achieving human performances;
- But can be very effective in a sufficiently constrained context.

→ **Good interfaces key to effective systems**

APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

Medical Imagery
- Microscopy

CARTOGRAPHY ON MARS

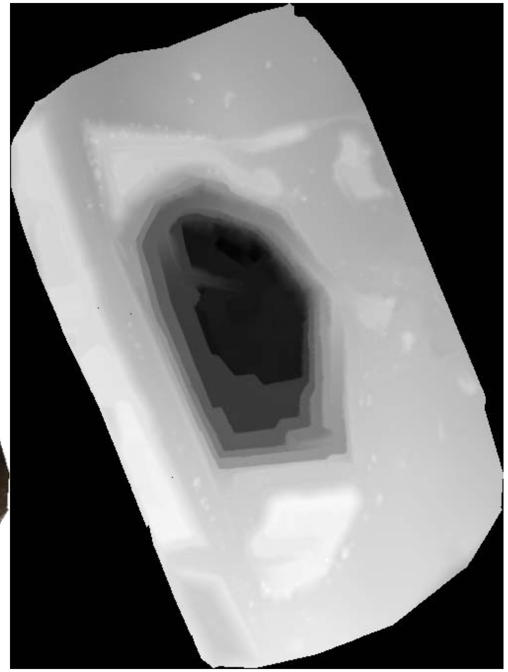
CARTOGRAPHY ON EARTH

ACQUISITION PLATFORMS

VIRTUAL MATTERHORN

Drone: https://www.sensefly.com **Mapping:** http://pix4d.com/

MINING SITE



- Fully automated.
- Accurate.
- Inexpensive.

GCP statistics

	X[m]	Y[m]	Z[m]
RMS	0.086	0.074	0.053
σ	0.040	0.061	0.053

http://www.pix4d.com/

SAILS AND WINGS

Useful for

- analysis of structural behavior under realistic strains,
- guiding design choices.

Salzmann et al. , CVPR'09

APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

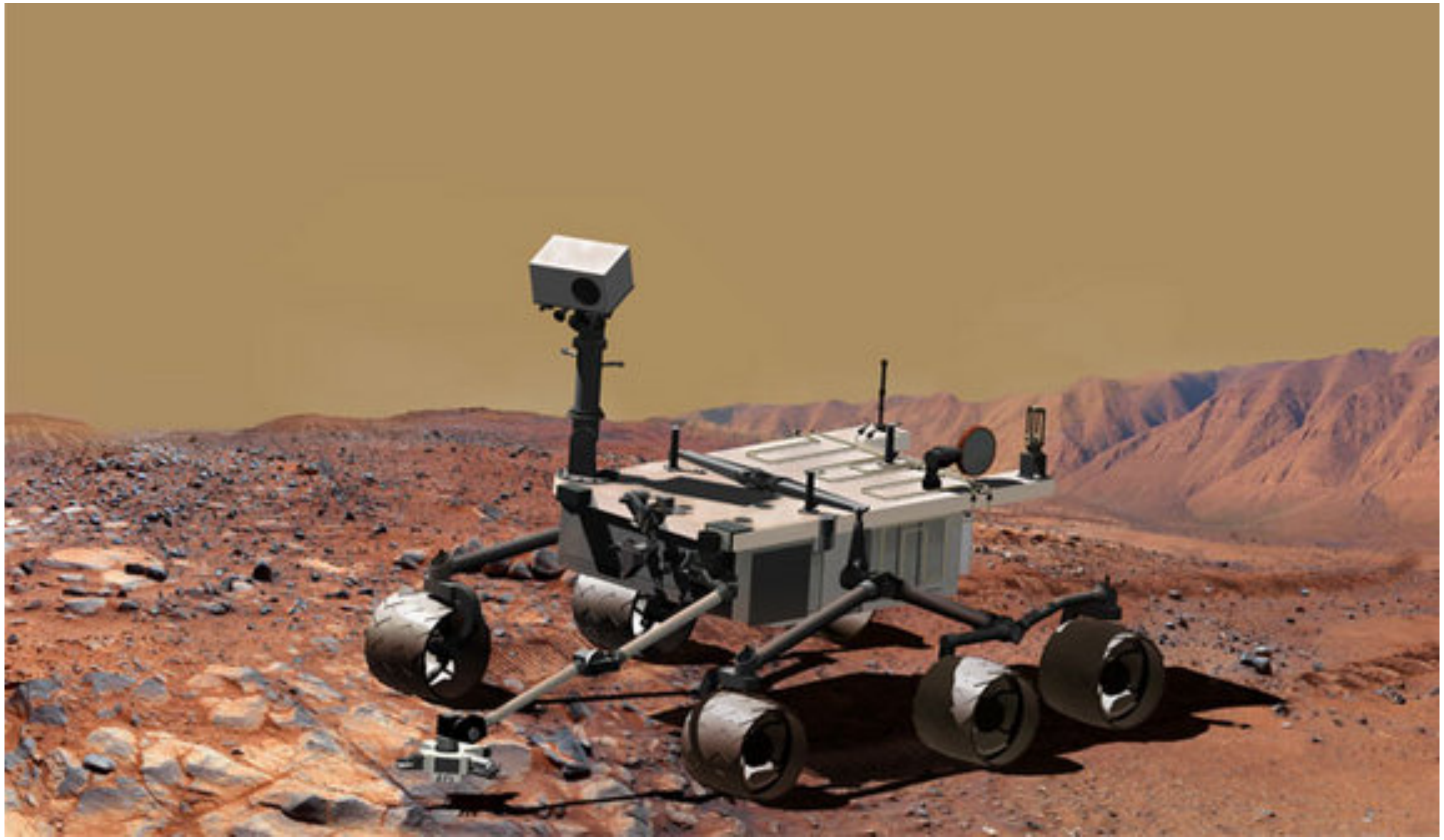
Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

Medical Imagery
- Microscopy

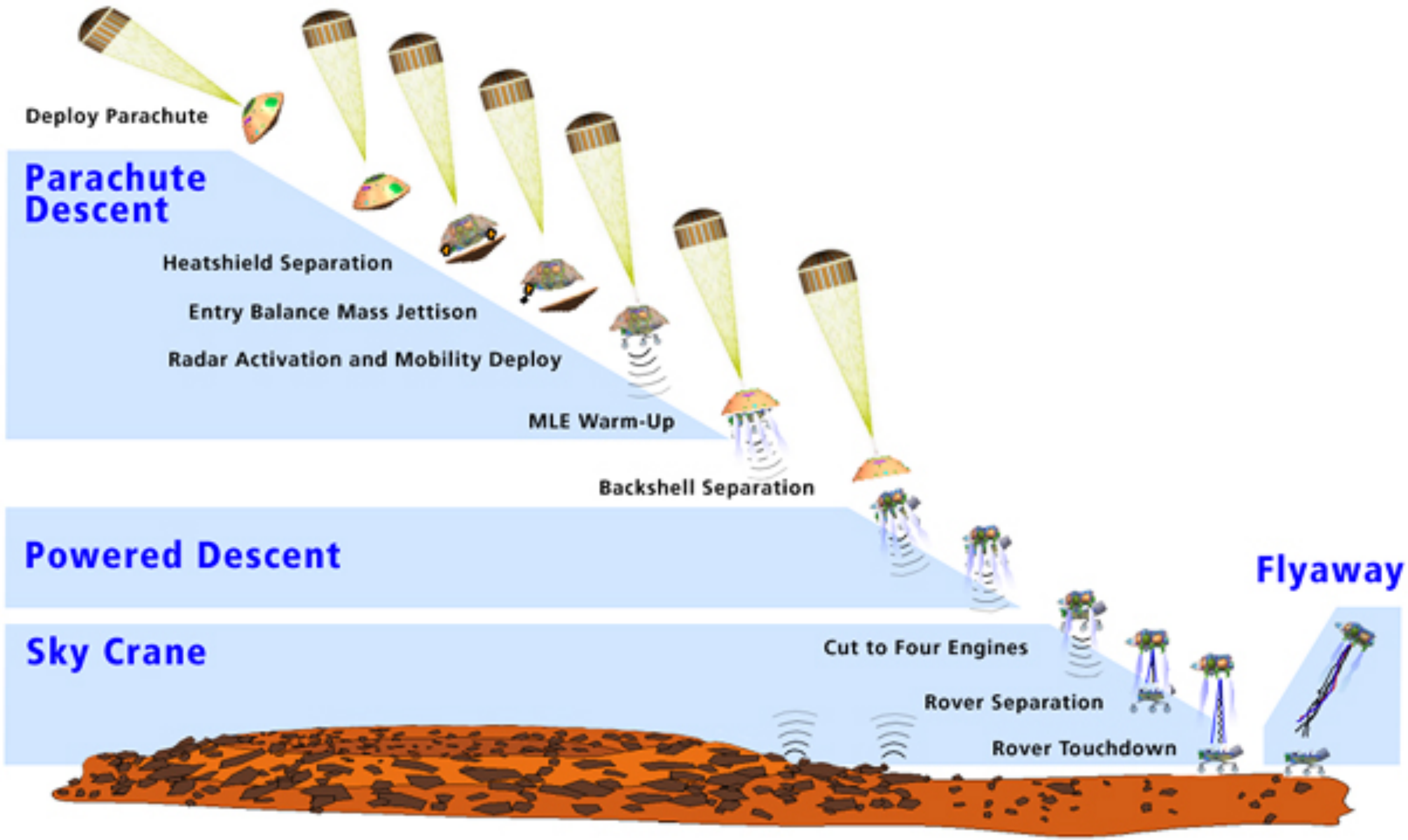
MARS ROVERS



Opportunity
Launched 2003
Landed 2004

Curiosity
Launched 2011
Landed 2012

DESCENT PROFILE



POWERED DESCENT

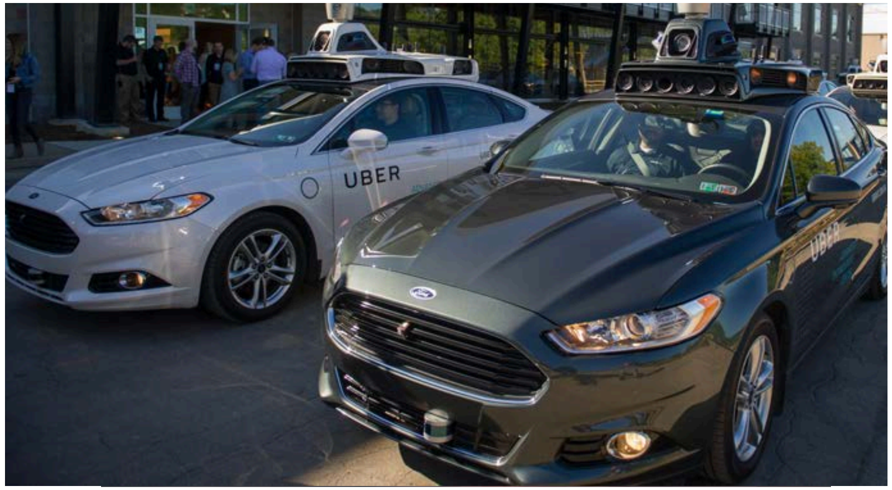
- Track feature points.
- Estimate drift.
- Combine with IMU output.
- Fire retro rockets as needed.

- Relatively simple computations.
- Space hardened hardware now fast enough.

EARTH ROVERS

1985
DARPA ALV

2007
DARPA Urban Challenge

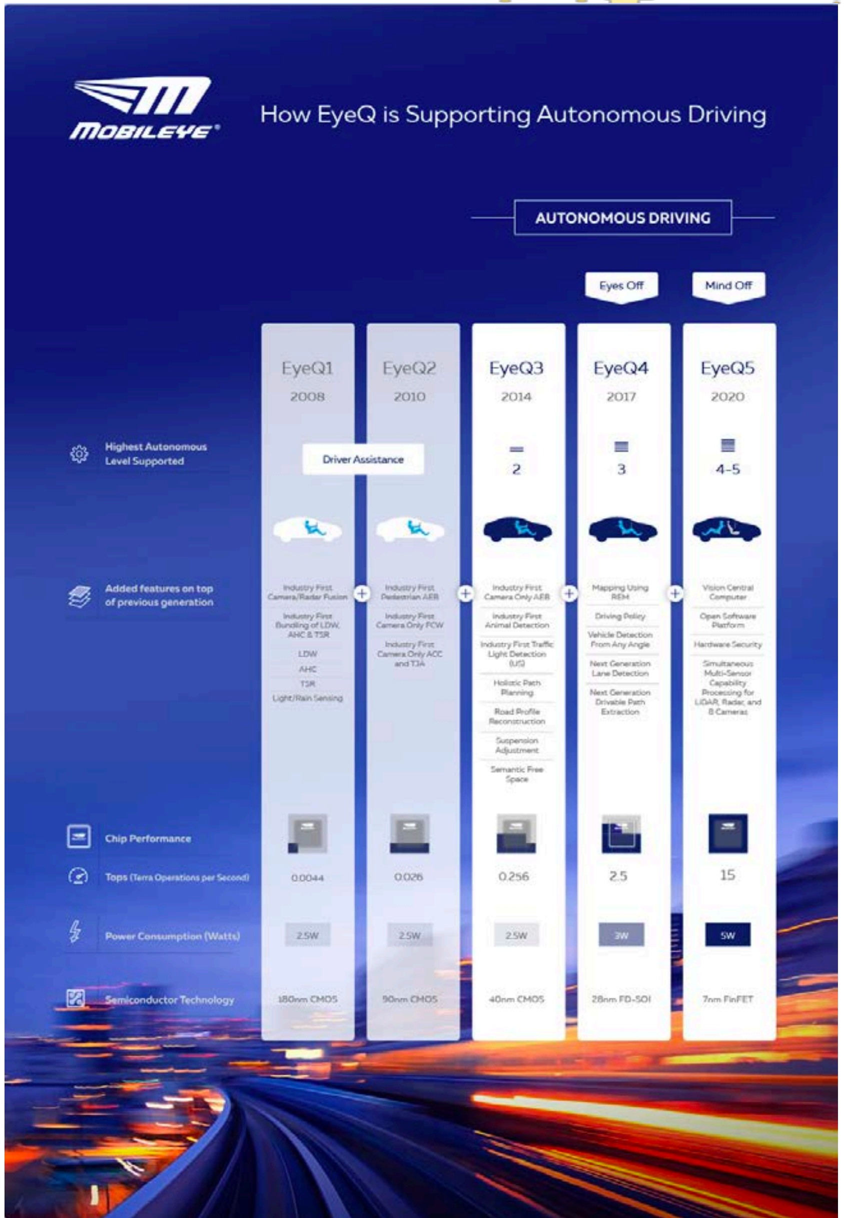


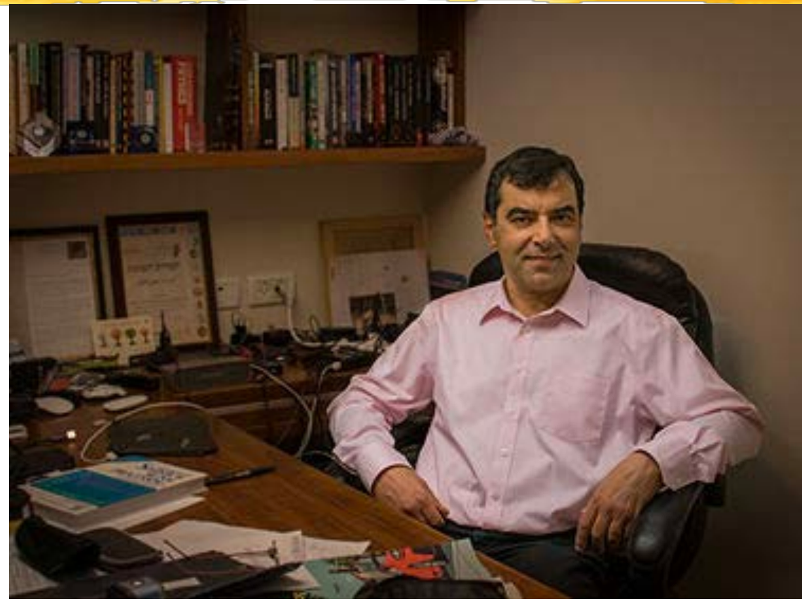
2014
Google Cars

2016
Uber

- Much more computing power.
- More reliable sensors.
- Detailed maps and models of the environment.

MOBILEYE





A. Shashua, co-founder 1999.

Intel Buys Mobileye in $15.3 Billion Bid to Lead Self-Driving Car Market.

NEW YORK TIMES, 13/03/17

APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

Medical Imagery
- Microscopy

VISUAL INSPECTION OF ASSEMBLED DEVICES

AKAmedix. www.taeym.be

Software embedded in the camera to find and read serial numbers

- Localization
- Illumination changes
- Generality

APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

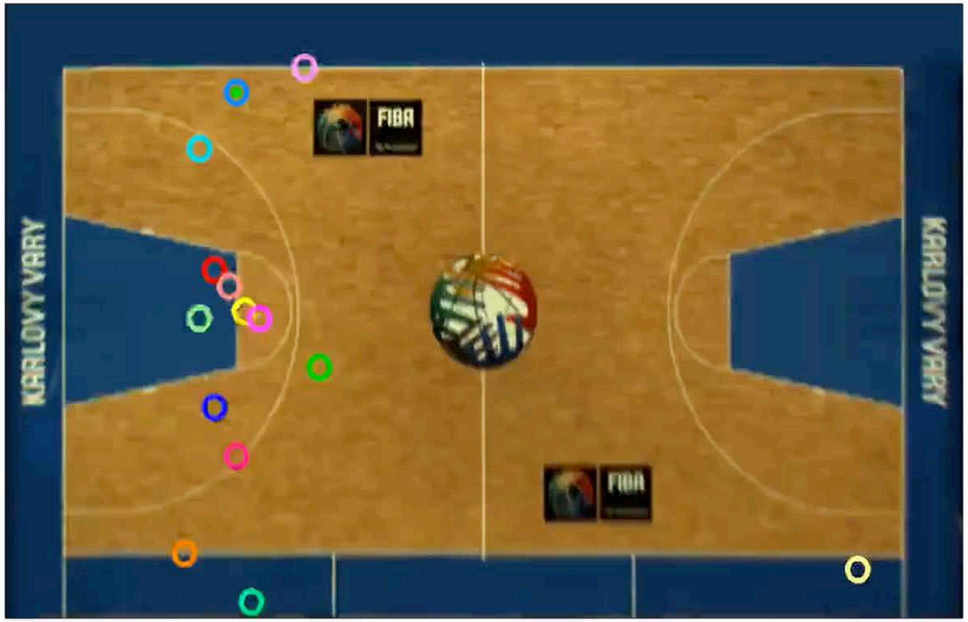
Medical Imagery
- Microscopy

LICENCE PLATES

Appian Technologies PLC

http://www.appian-tech.com/products/anpr_overview.html

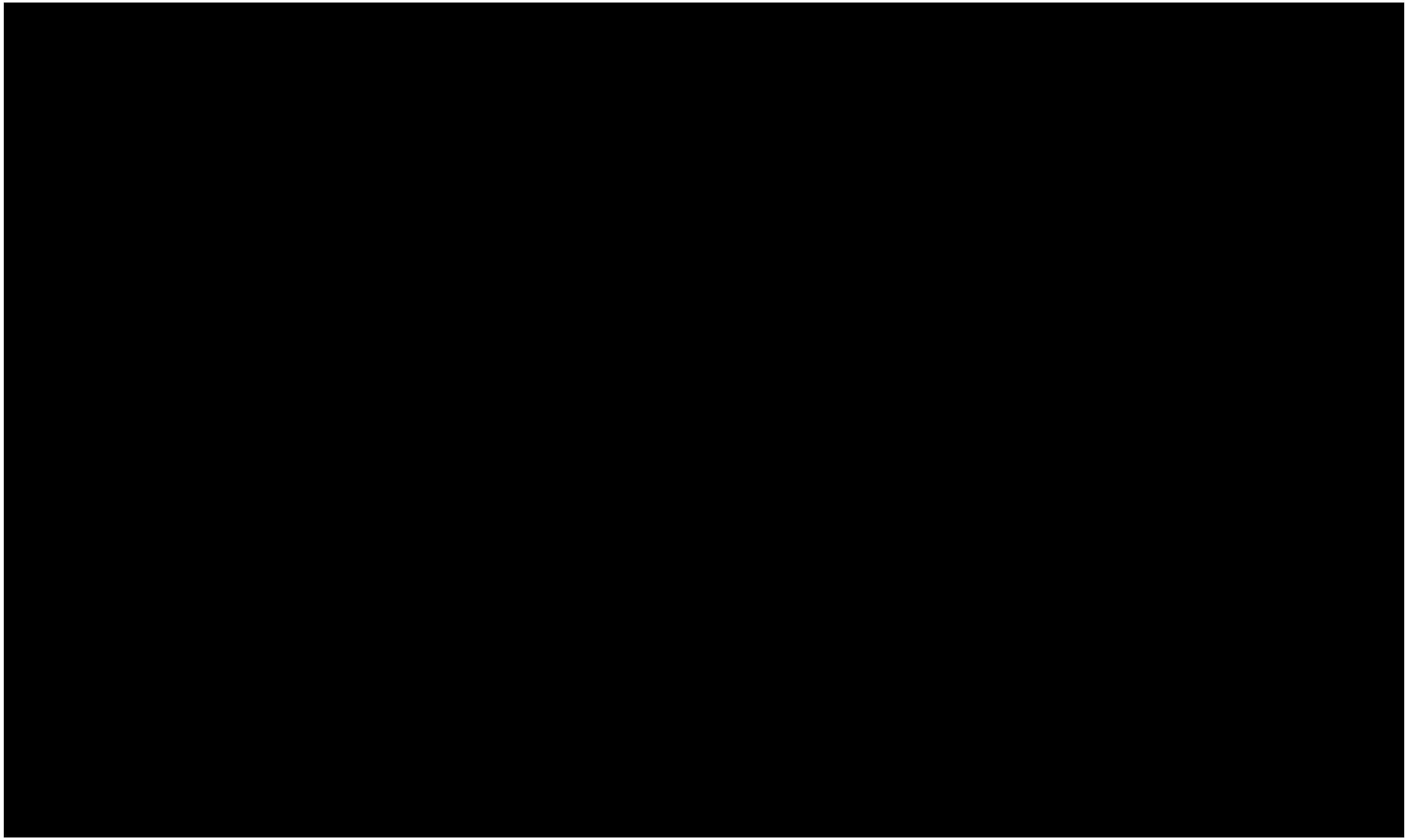
TRACKING PEOPLE



... and the ball → Behavioral analysis.

BenShitrit et al. PAMI’13

TECH TRANSFER



https://www.secondspectrum.com/

APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

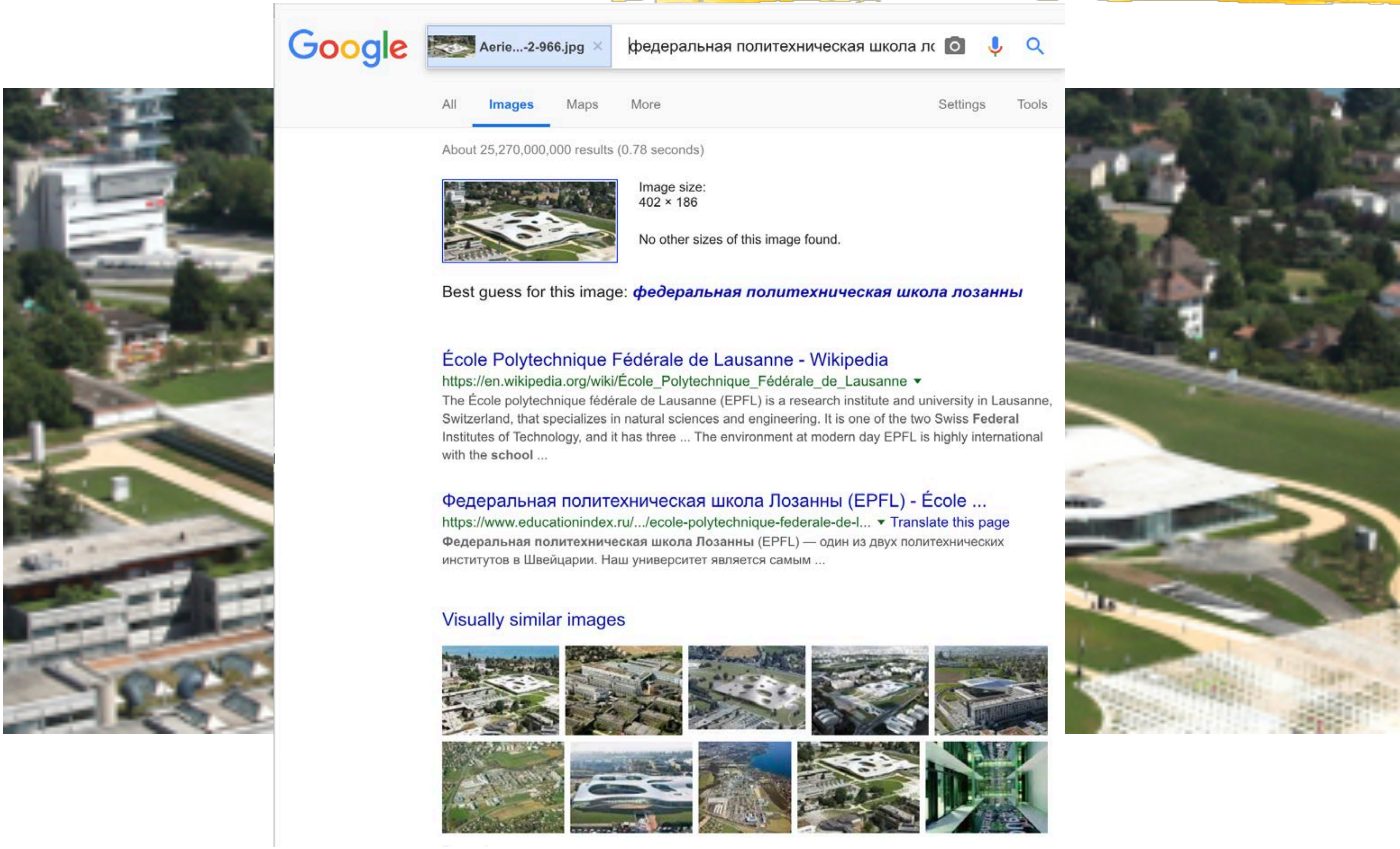
Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

Medical Imagery
- Microscopy

IMAGE RETRIEVAL



APPLICATIONS

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

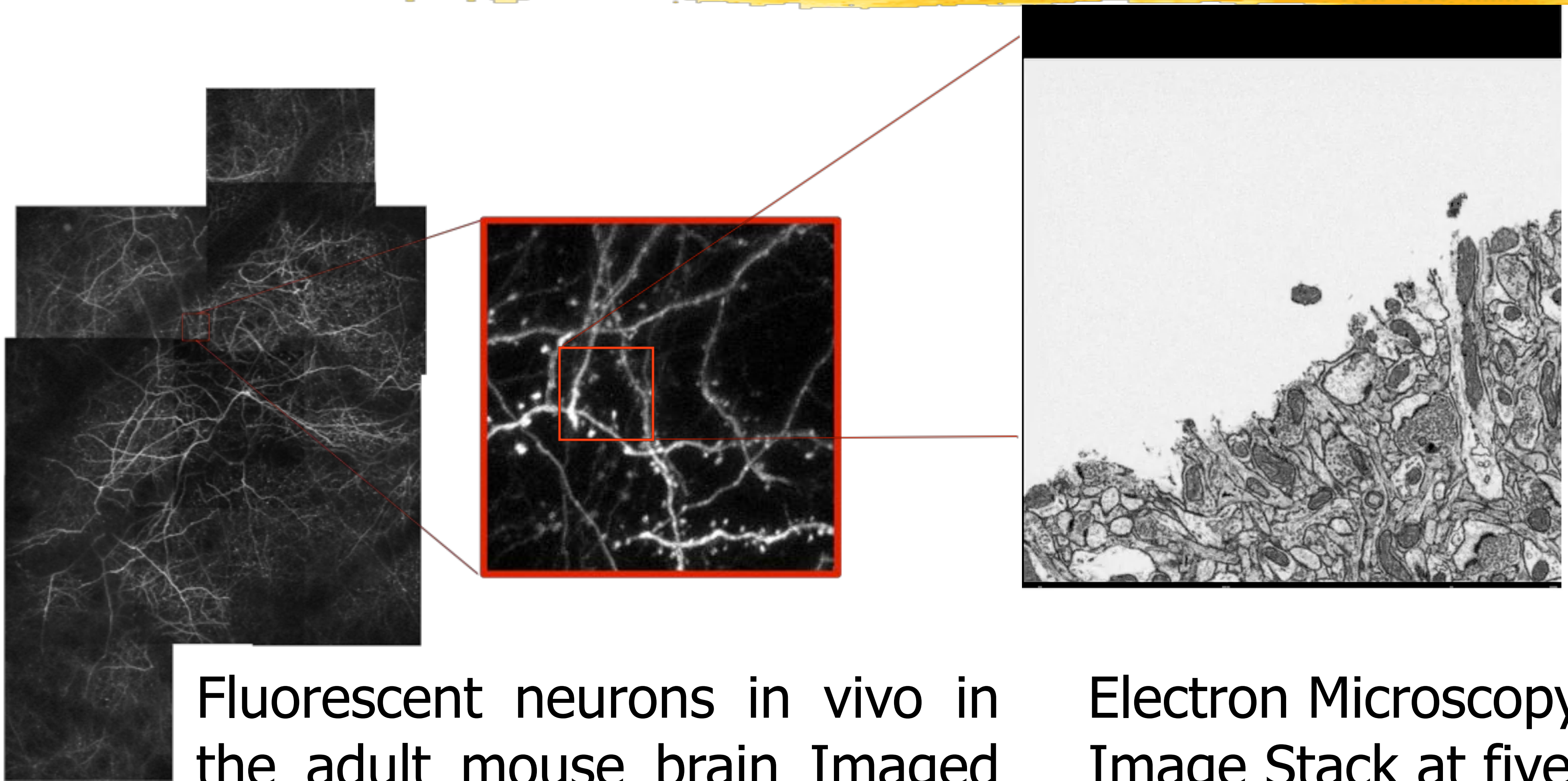
Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
- Retrieval and Annotation

Medical Imagery
- Microscopy

MICROSCOPY

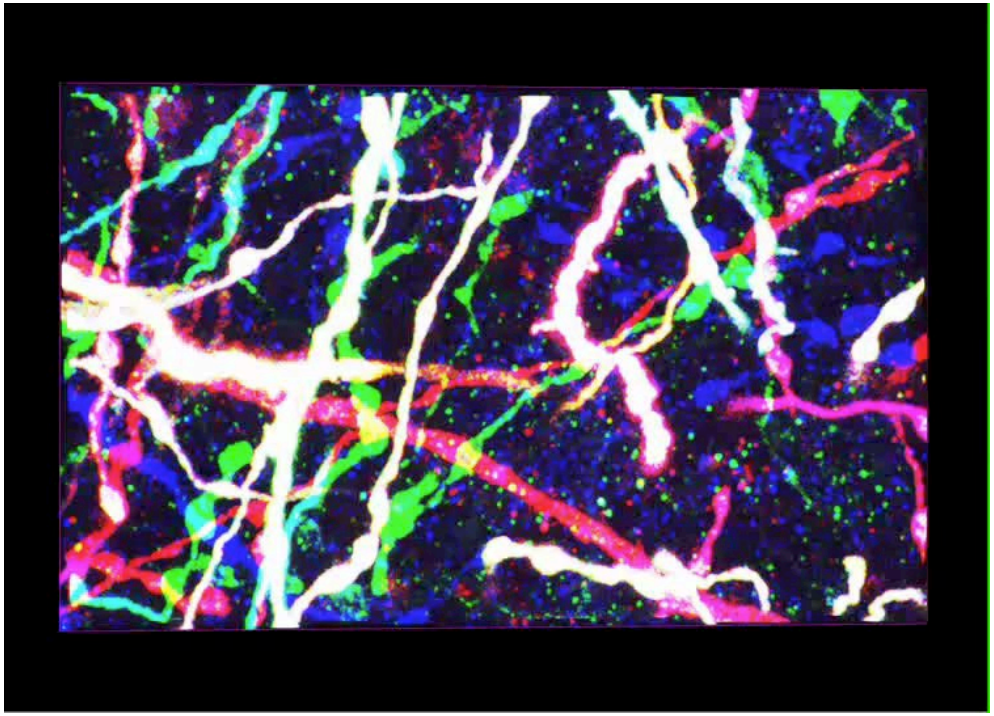


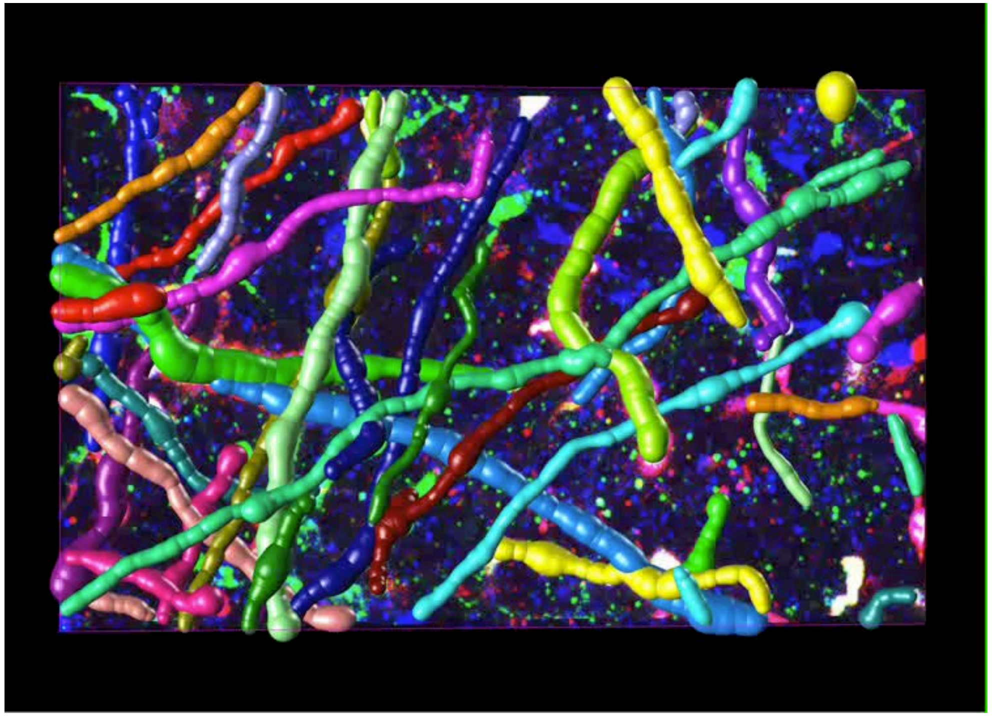
Fluorescent neurons in vivo in the adult mouse brain Imaged through a cranial window using a 2-photon microscope.

Electron Microscopy Image Stack at five nanometer resolution.

Courtesy of G. Knott

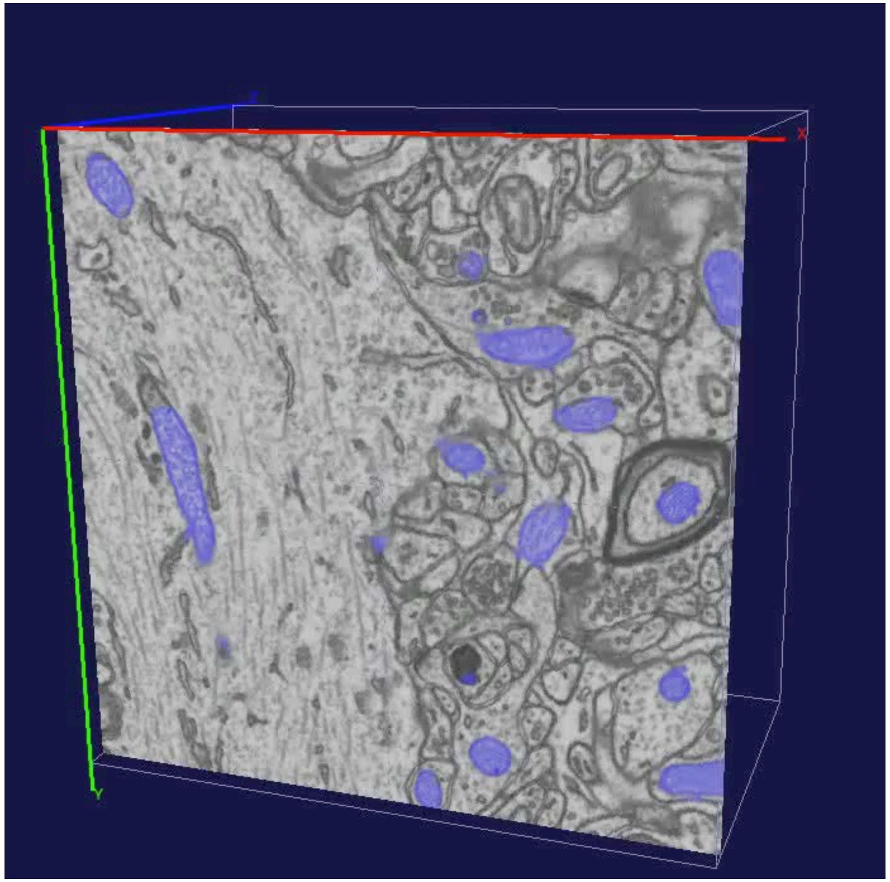
DELINEATING DENDRITIC TREES

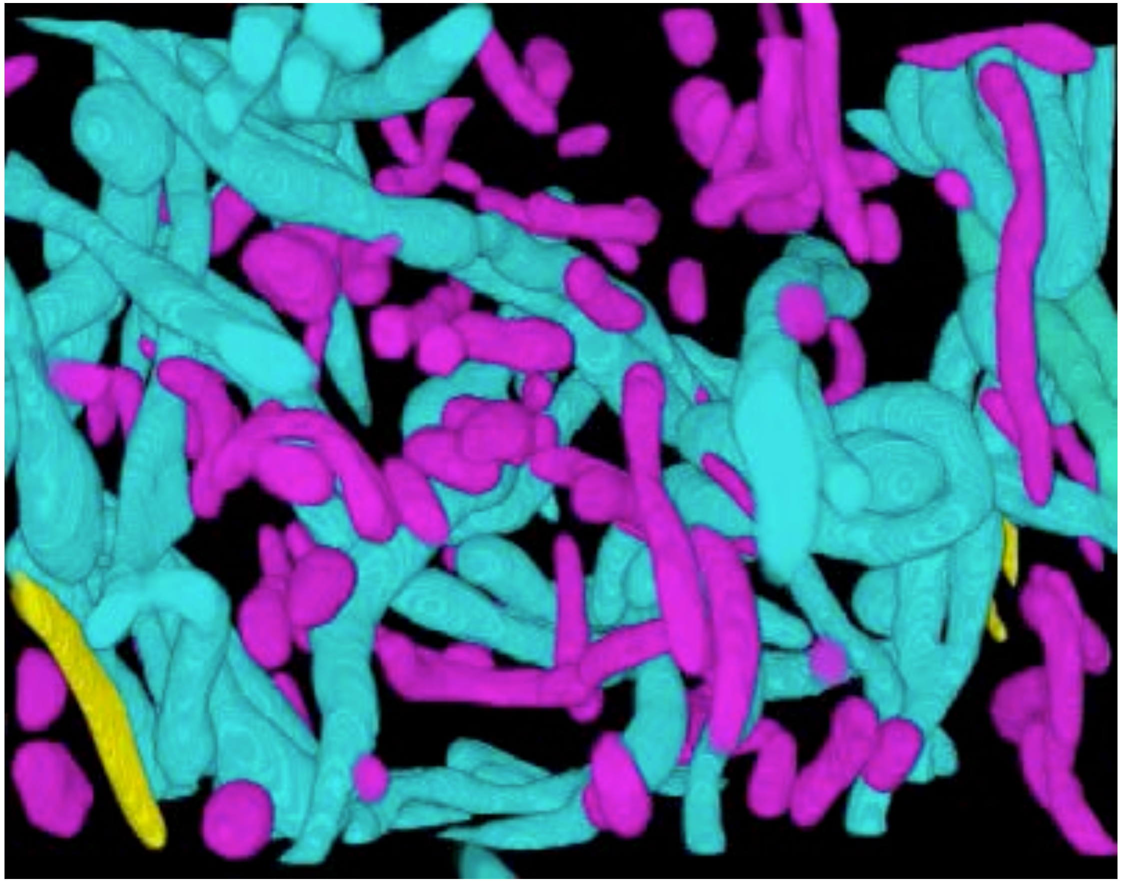




Turetken et al., PAMI'16

FINDING MITOCHONDRIA





Lucchi et al. TMI'11

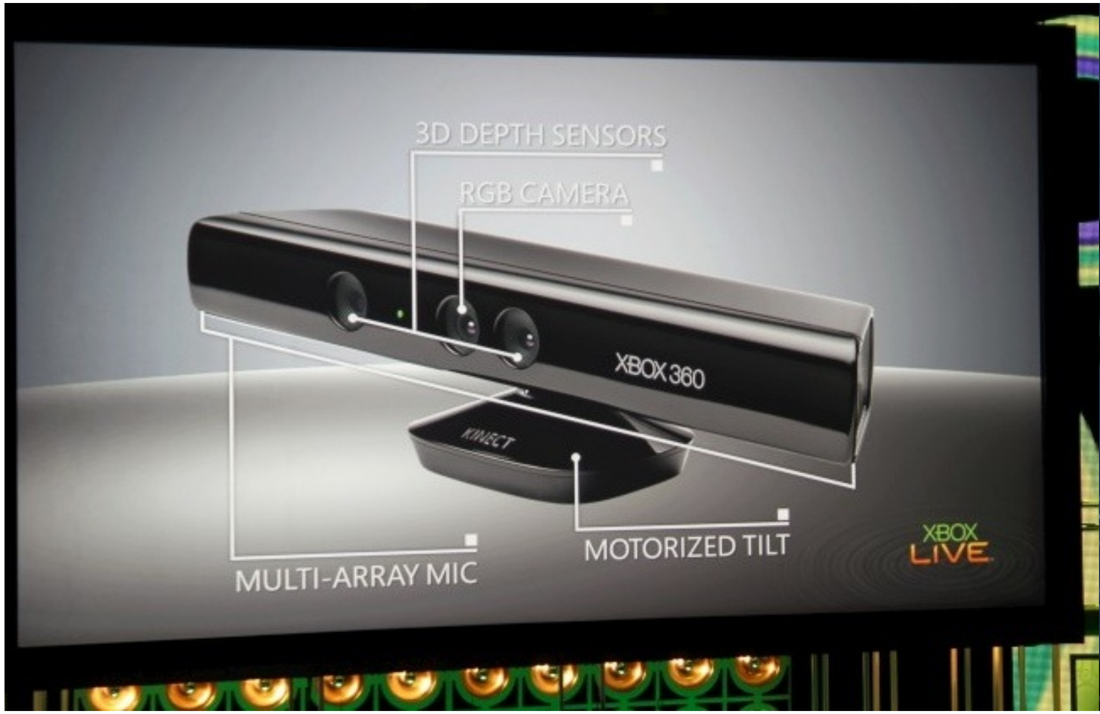
GOOGLE EARTH FOR THE BRAIN

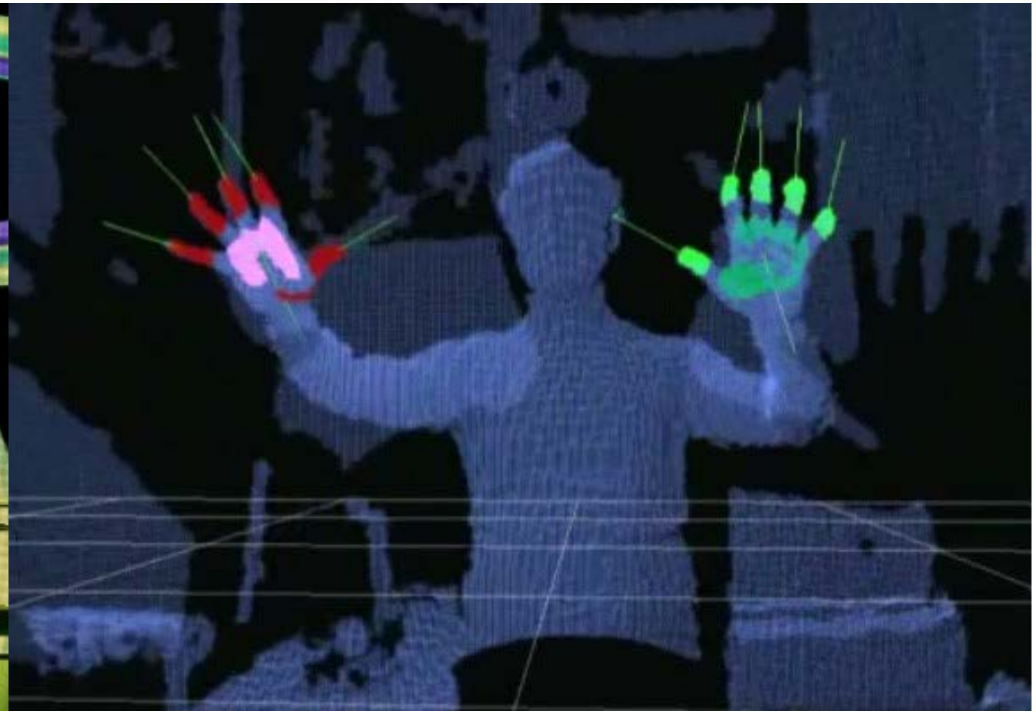
• A human brain contains approximately 100 billion neurons and 100 trillion synapses.
• It would take 1000 Exabytes to store an uncompressed digitization at 5nm resolution.

amazon x 500!

—> Seriously big data!

KINECT



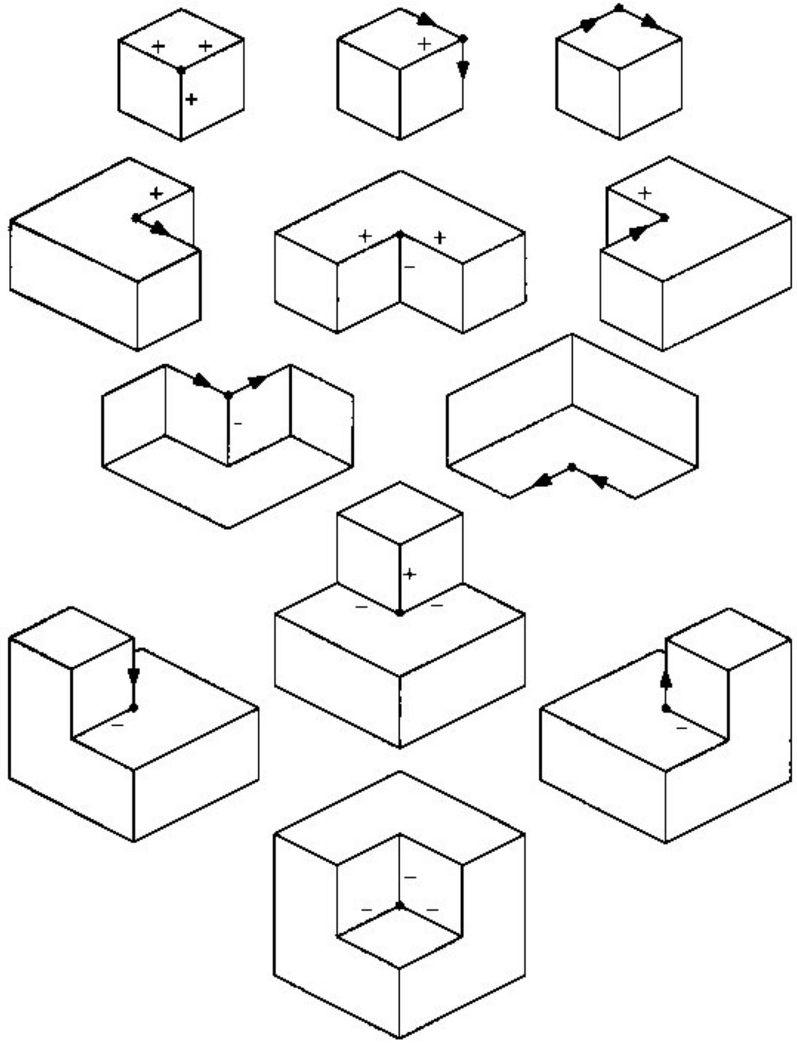


→ Image whose pixel values are distances

HOW IT BEGAN

- Computer Vision started in 1965 at MIT as a short term project.
- A world of perfect blocks and strong assumptions.

→ The real world is not like that!



Roberts, PhD 65

HISTORICAL PERSPECTIVE

1960s: Beginnings in artificial intelligence, image processing and pattern recognition.

1970s: Foundational work on image formation.

1980s: Vision as applied mathematics, geometry, multi-scale analysis, control theory, optimization.

1990s: Physics-based models, Probabilistic reasoning.

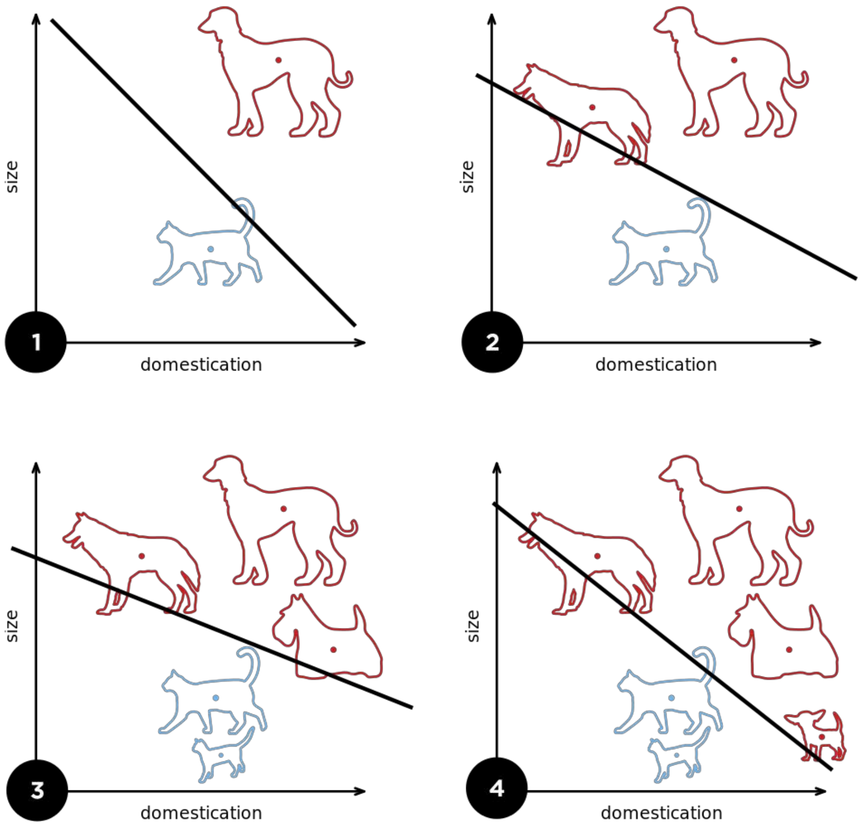
2000s: Machine learning.

2010s: Deep Learning.

2020s: ?????

--> Improved understanding and successful applications in graphics, mapping, biometrics, and others but still far from human performance.

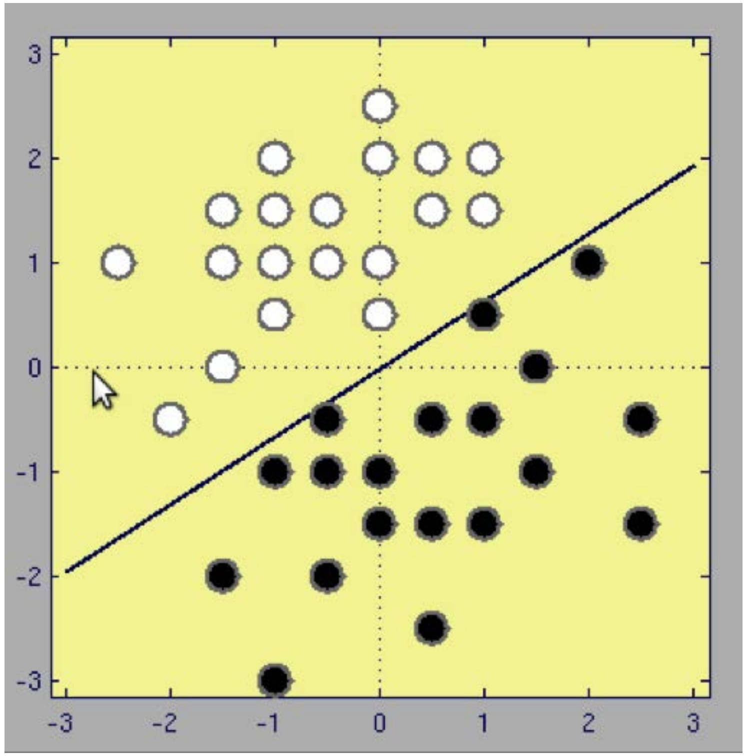
LINEAR CLASSIFICATION



$$f(\mathbf{x};\mathbf{w},b) = \begin{cases} 1 \text{ if } \mathbf{w}\cdot\mathbf{x}+b \geq 0, \\ 0 \text{ otherwise.} \end{cases}$$

Learning: Finding w and b so that the number of misclassified instances in the training set is minimized.

PERCEPTRON



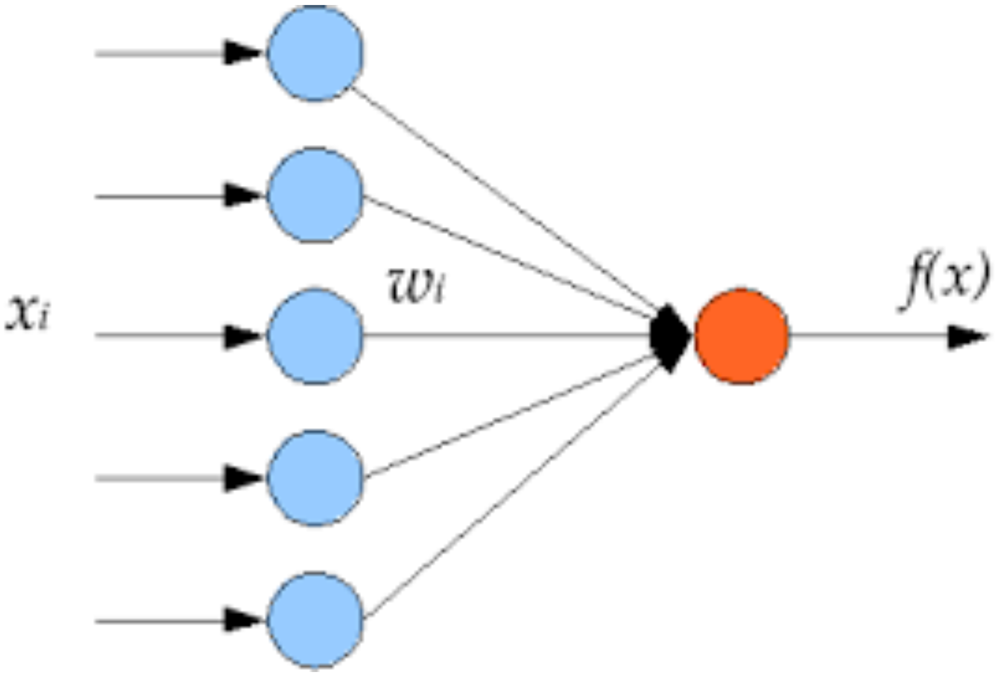
- Center the $\mathbf{x}_i$s so that $b = 0$.
- Pick a random $\mathbf{w}_0$.
- Iteratively, pick a random index i.
  - If $\mathbf{x}_i$ is correctly classified, do nothing.
  - Otherwise, $\mathbf{w}_{t+1} = \mathbf{w}_t + y_i \mathbf{x}_i$.

MORE ANCIENT HISTORY

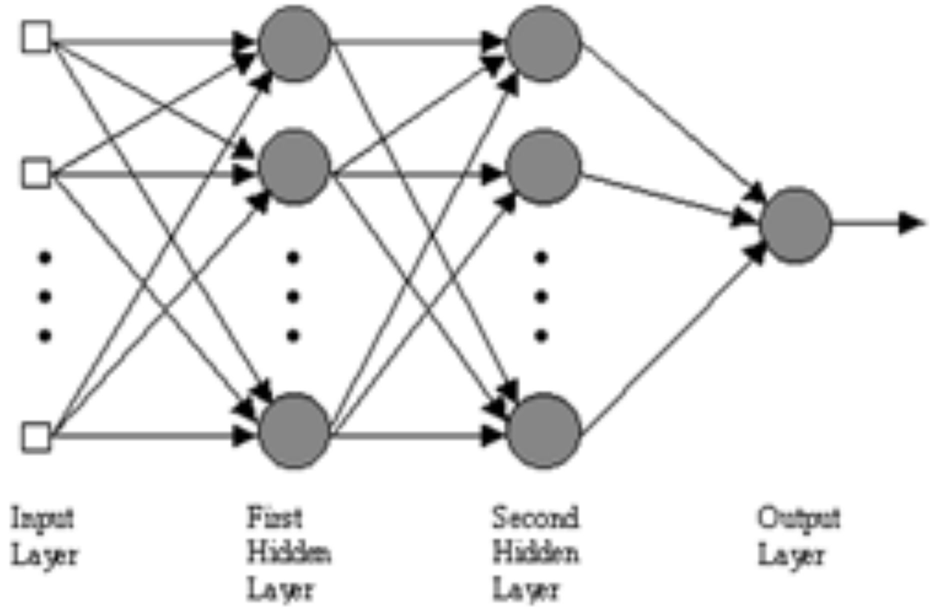
- The perceptron is a simple algorithm, but imagine coding it on this IBM 704, which Frank Rosenblatt used to to implement it in 1957.

- There was much initial enthusiasm. But, it was later realized there were serious limitations, such as the linear separability requirement.
- The perceptron eventually evolved to have multiple layers and smooth activation functions.

LAYERED PERCEPTRONS



Single Layer

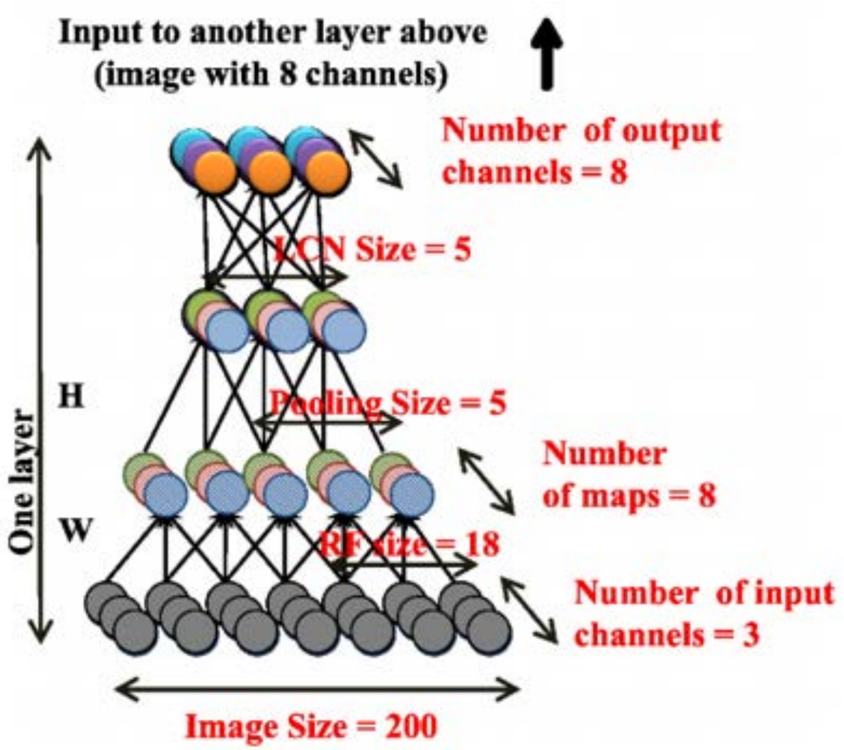


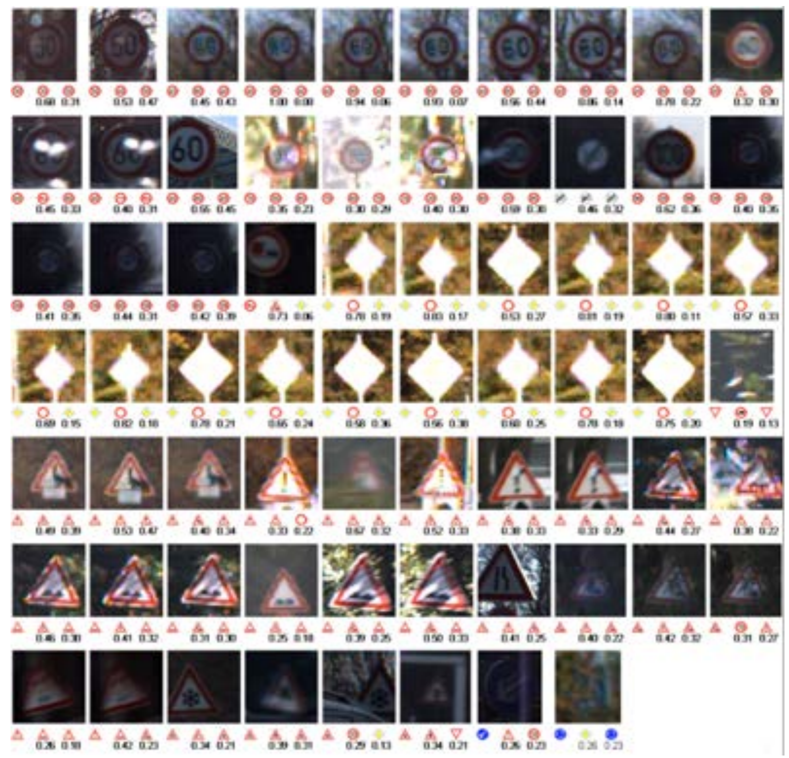
Multiple Layers

Requires many wide layers to be effective:

- Mostly impractical in the 1980s due to computational limitations.
- Shown to be highly effective in 2012 by using GPUs.

CONVOLUTIONAL NETWORKS

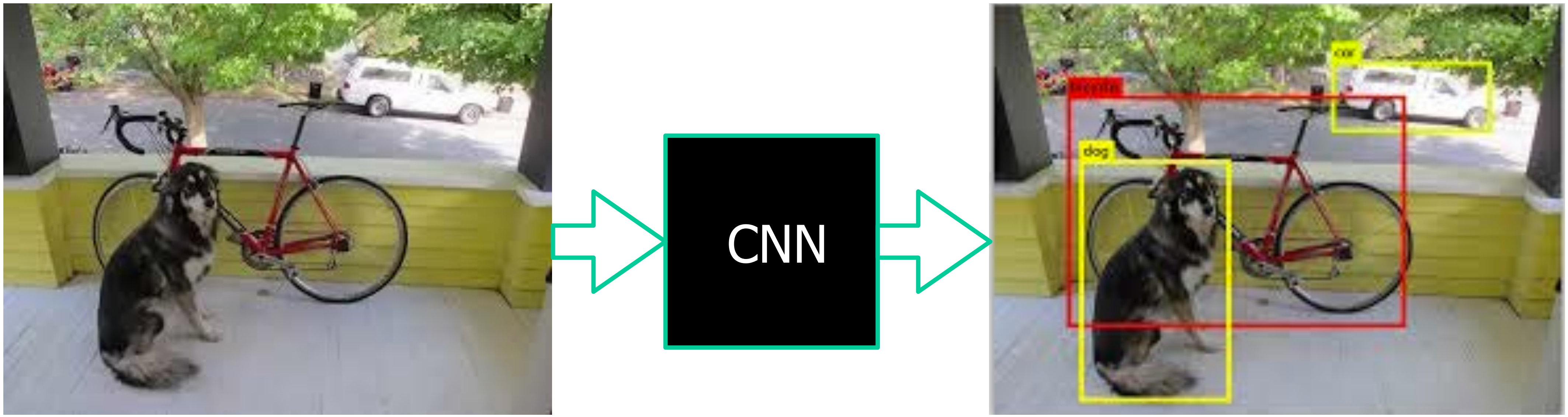




Multi-Layer Perceptron --> Deep Convolutional Network:
Effective use not only of GPUs but also of the geometry of images.

Le et al. ICML'12
Ciresan et al. , NIPS'12

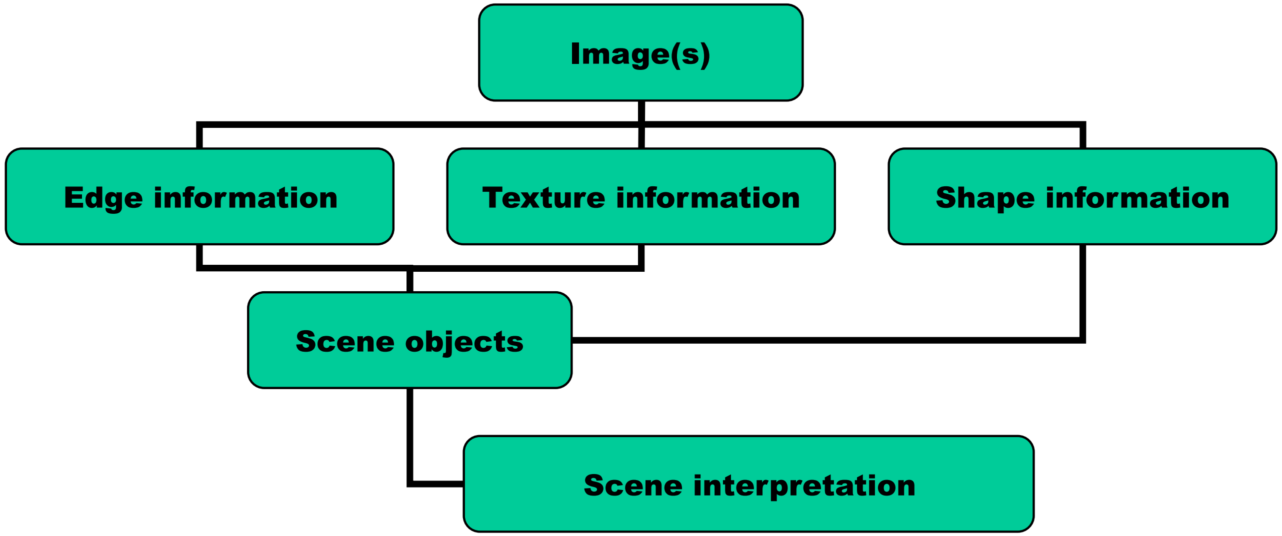
COMPUTER VISION IN THE ERA OF DEEP NETS



- Extremely effective in many cases but does not shed much light on the vision process.
- The best algorithms combine Deep Nets with more traditional techniques.

Redmon et al. , CVPR'16

A TEACHABLE SCHEME



Decomposition of the vision process into smaller manageable and implementable steps.

--> Paradigm followed in this course

--> May not be the one humans use

COURSE OUTLINE

Introduction:
- Definition
- Image formation

Extracting features:
- Contours
- Texture
- Regions

Shape recovery:
- From one image
- Using additional images

COURSE ORGANIZATION

- Formal lectures every week (Friday)
- Exercises every other week (Tuesday)
- Written examination

WEBSITES

Projects:

- http://cvlab.epfl.ch/projects

Research activities:

- http://cvlab.epfl.ch/research

COURSE MATERIAL

Textbooks:

- R. Szeliski, Computer Vision: Computer Vision: Algorithms and Applications, 2010.
- A. Zisserman and R. Hartley, Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.

Web pages:

- moodle.epfl.ch (Computer Vision, introcv)
- cvlab.epfl.ch (Teaching & Projects)