Computer Vision

P. Fua
IC-CVLab
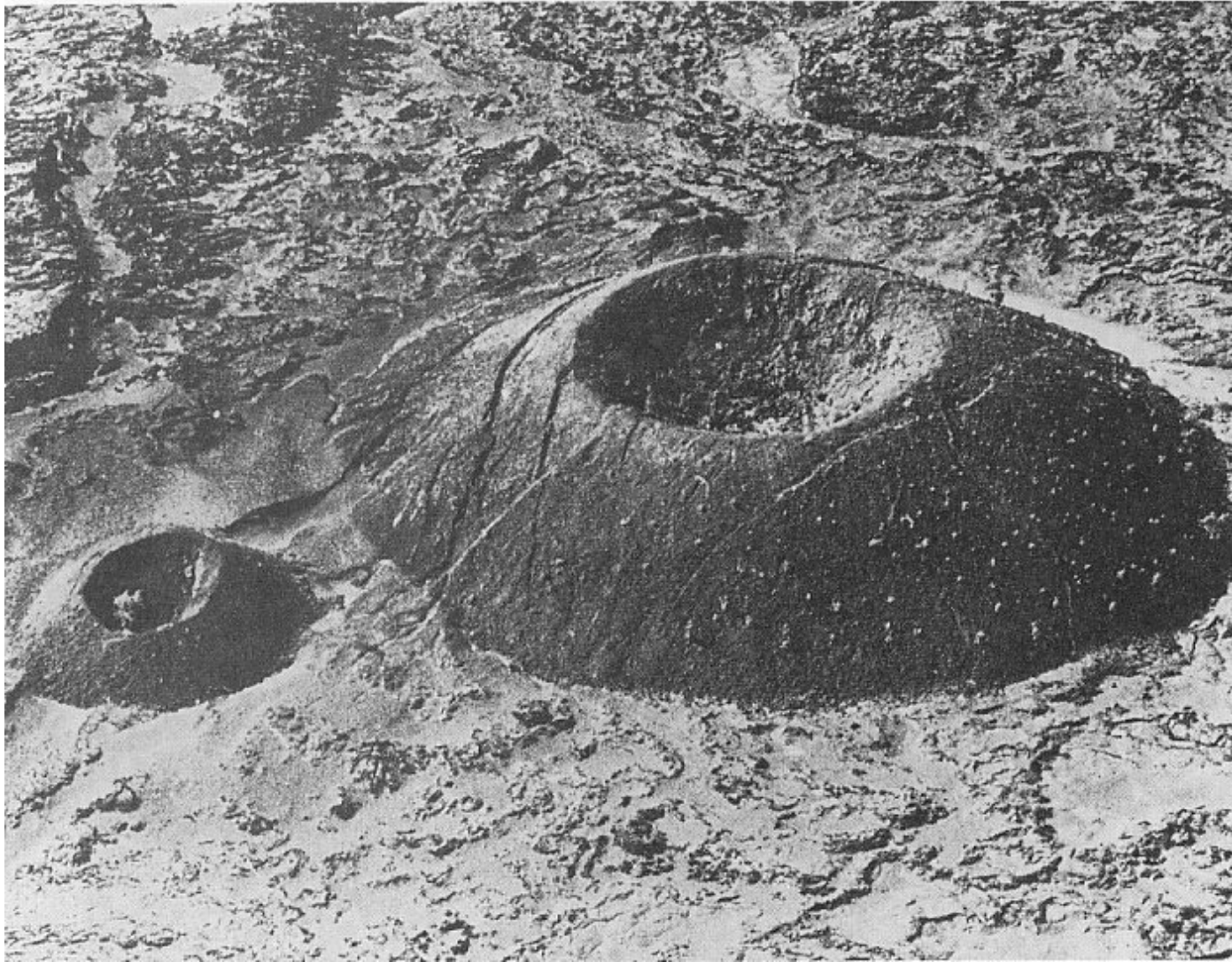EPFL

Computer Vision

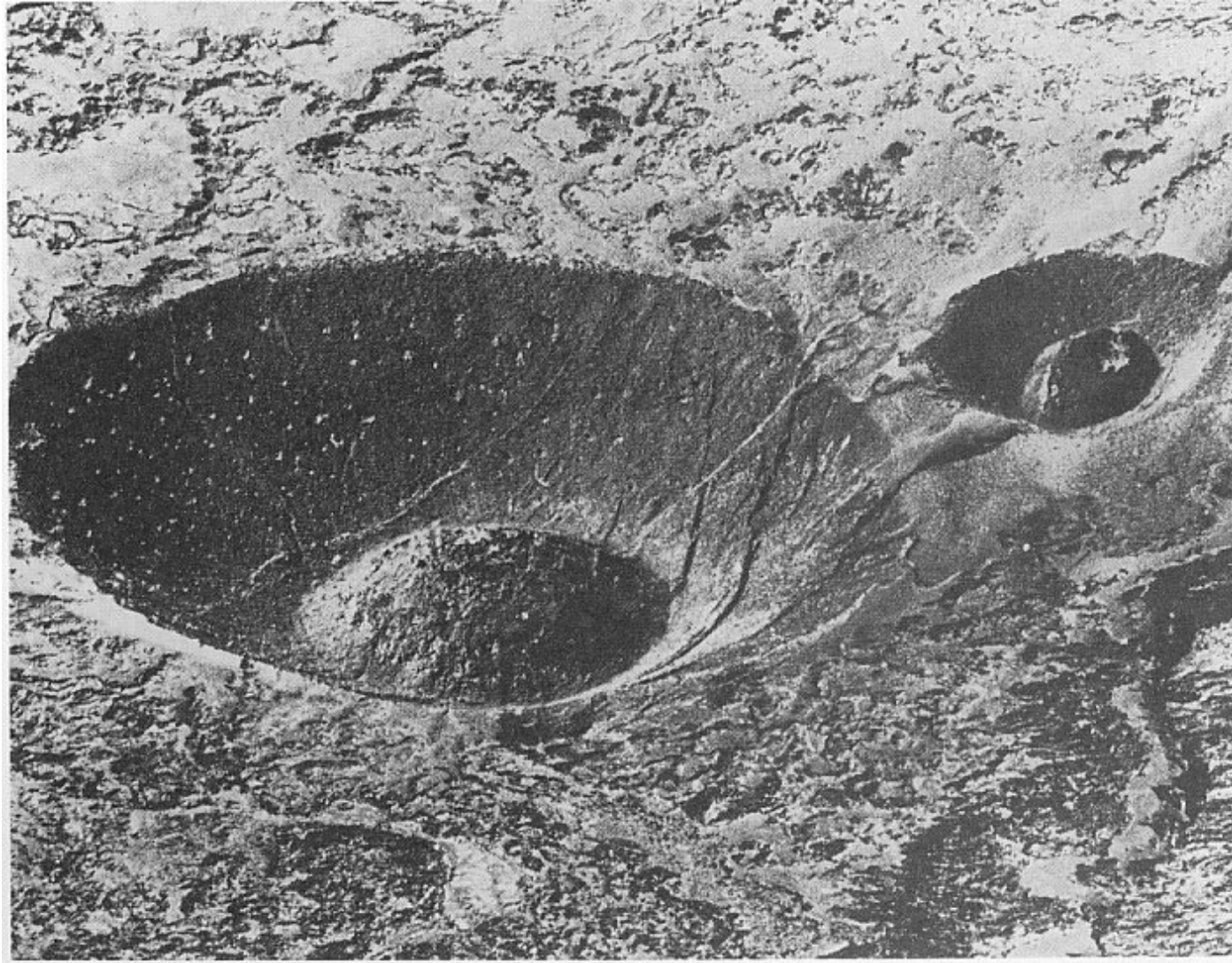**Goal:** Inferring the properties of the world from one or more images

- Photographs
- Video Sequences
- Medical images
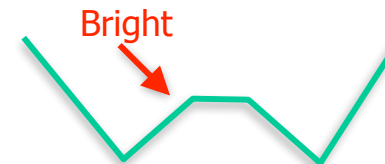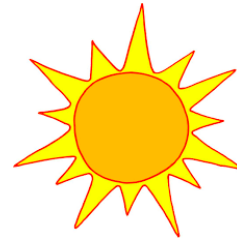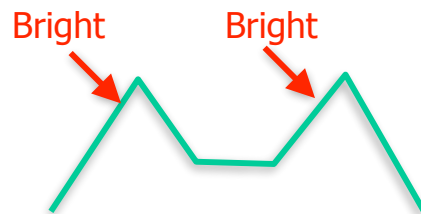- Microscopy data

→ **Image Understanding**

What do You See?

And Now?

Potential Interpretation
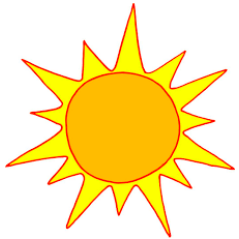
Shape from Contours

- Later in the class, we will formalize this in terms of a set of differential equations.
- This is probably not what we do in our heads when we look at images.

A Powerful Mechanism

A Powerful Mechanism

Illusory Contours

Kanizsa's Triangle

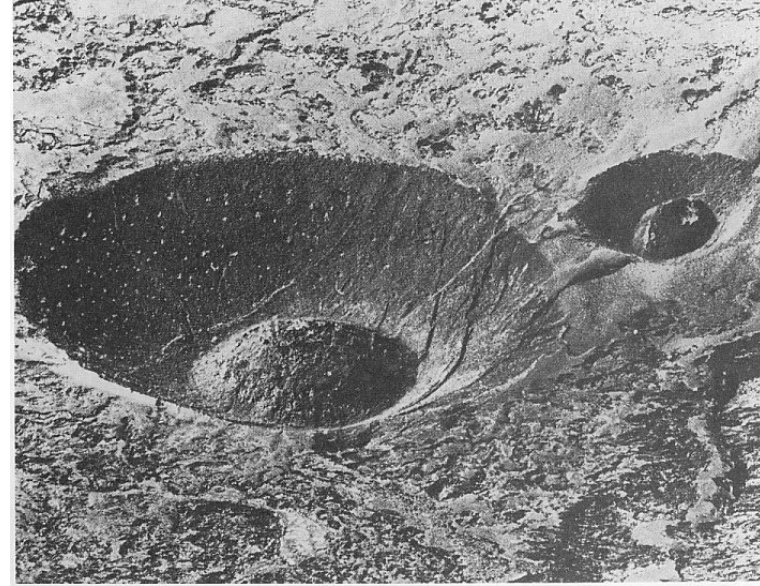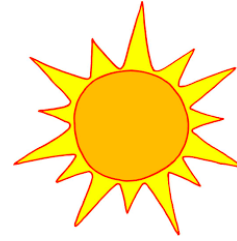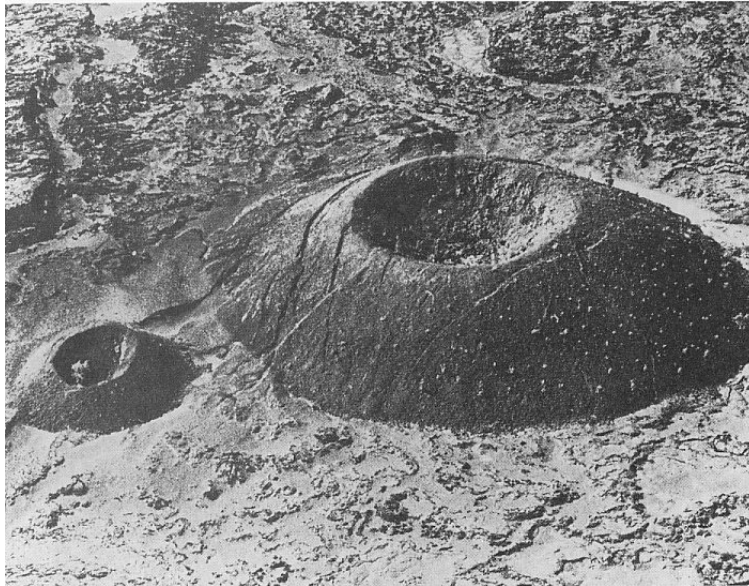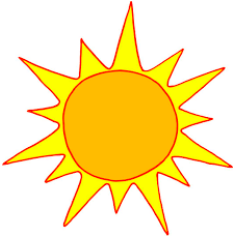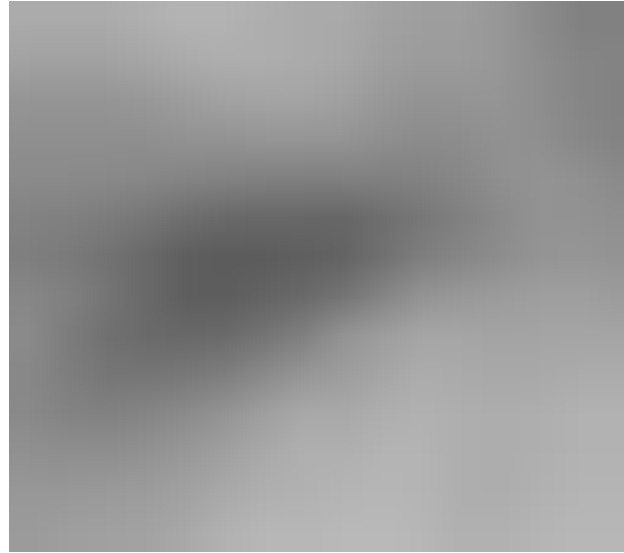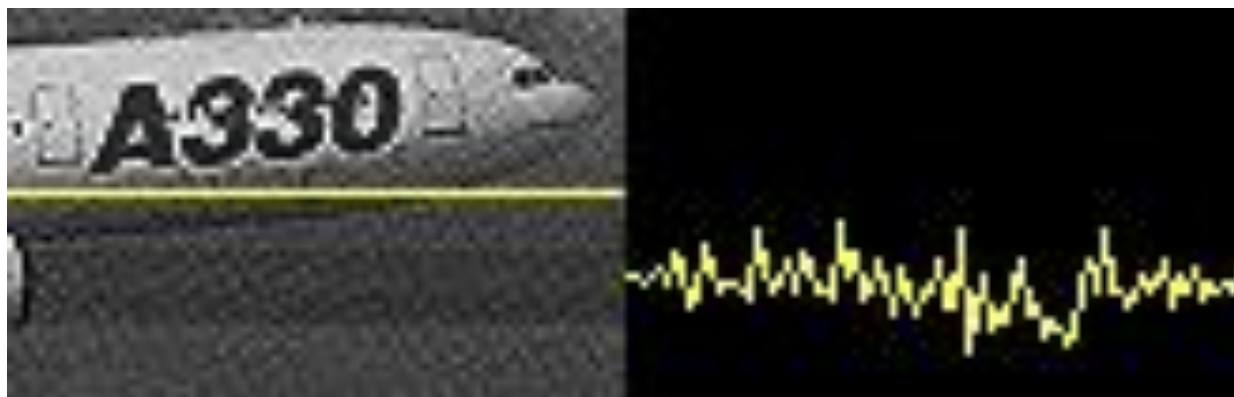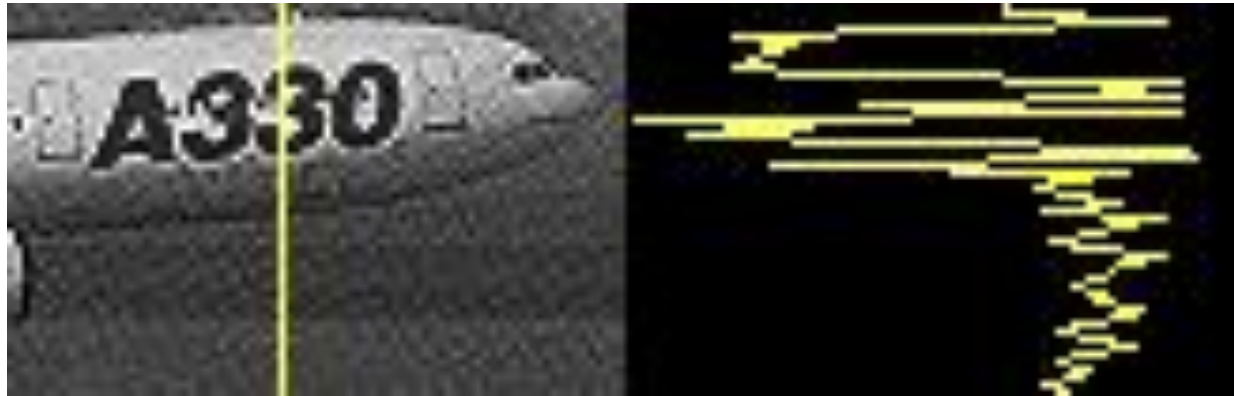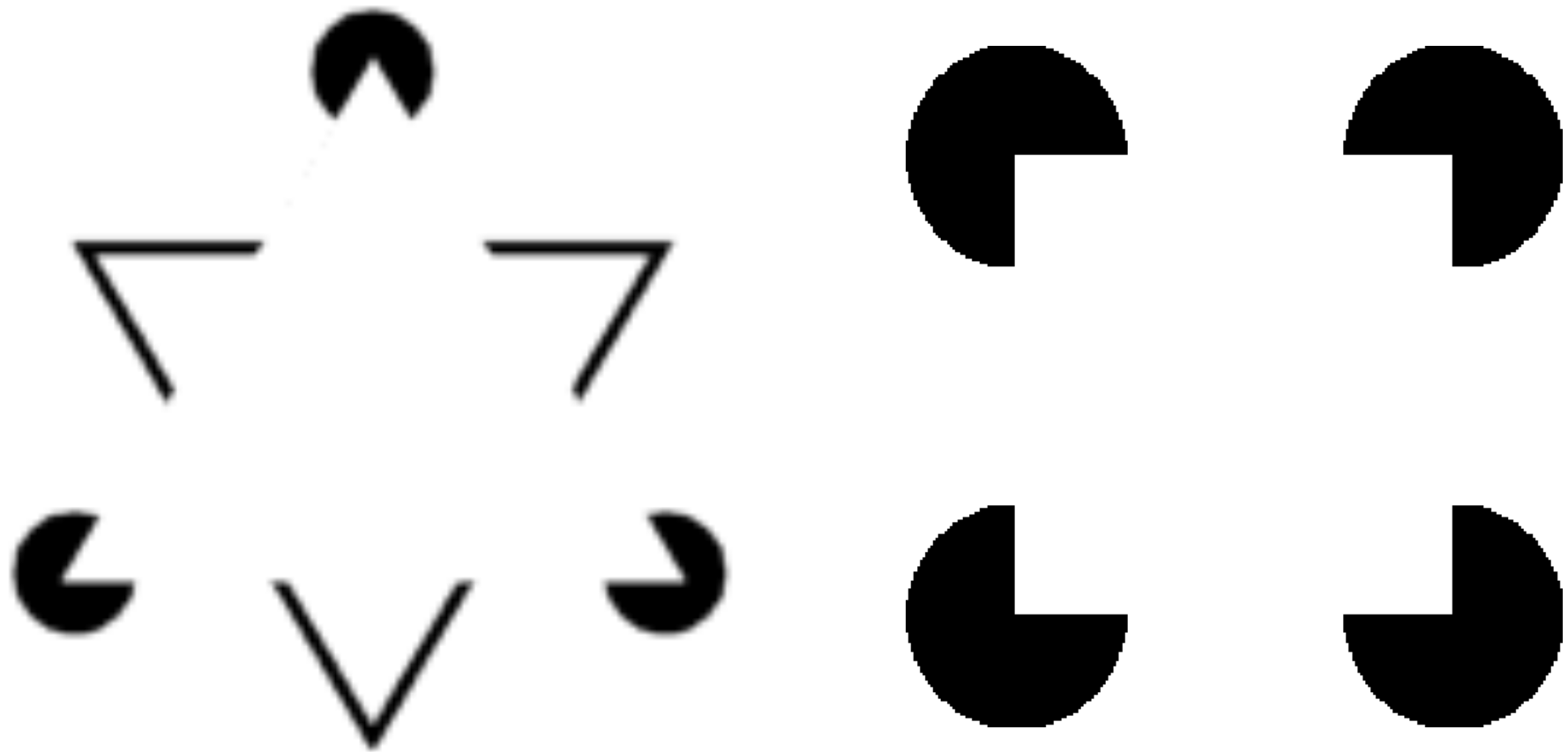**Closure of Good Form Hypothesis:** Illusory contours represent an example of the closure of good form (Osgood 1953), (Pastore 1971), (Kanizsa 1976, 1979).

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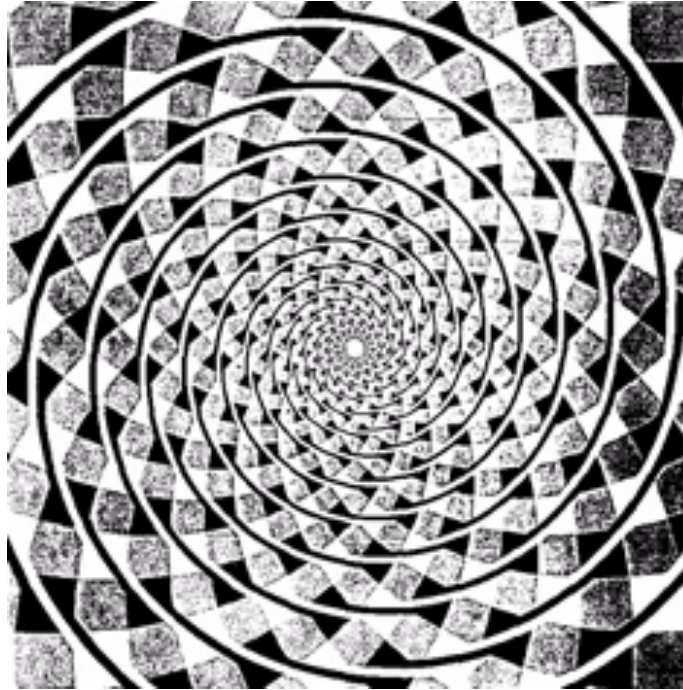
Every image is the image of thing merely to him who knows how to read it, and who is enabled by the aid of the image to form an idea of the thing.

Handbook of Physiological Optics
H. von Helmholtz

Photometric Illusion
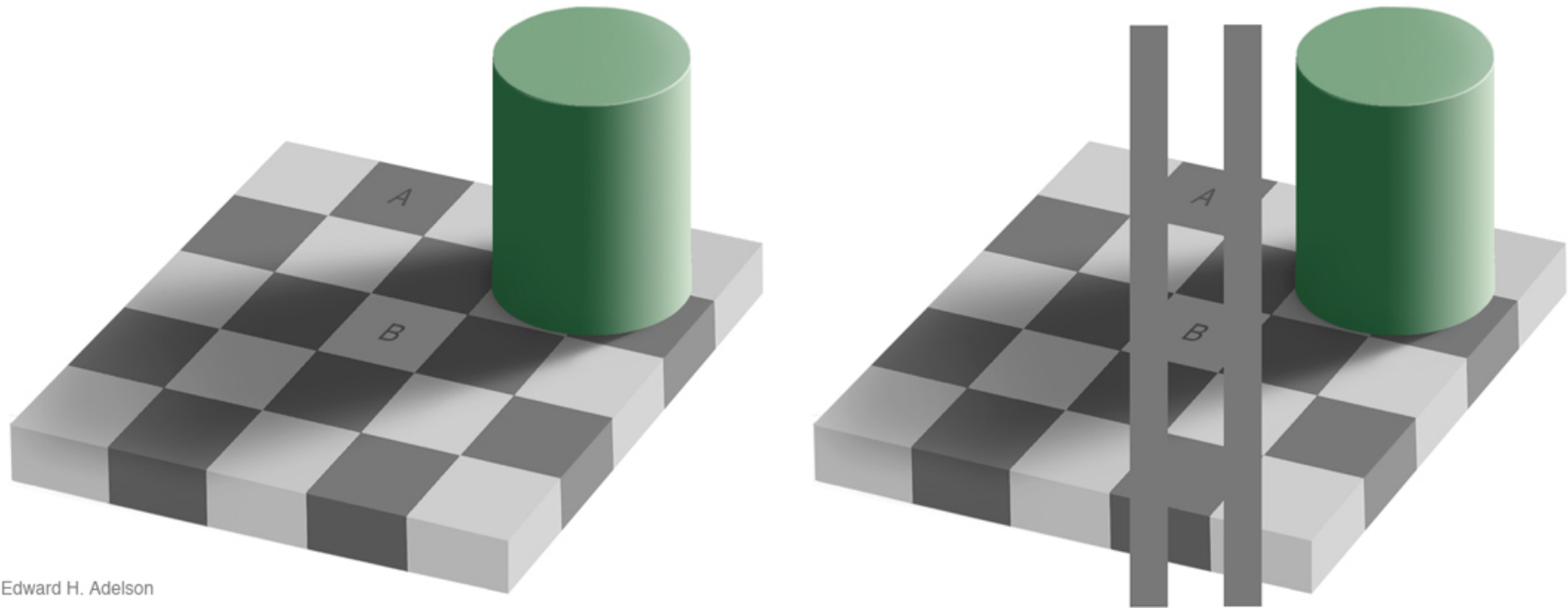
Edward H. Adelson

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
- Statistical and physics-based models
- Training data

True image understanding seems to require a great deal of thinking. We are not quite there yet.

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

Satellite view
(real)

Ground view
(synthetic)

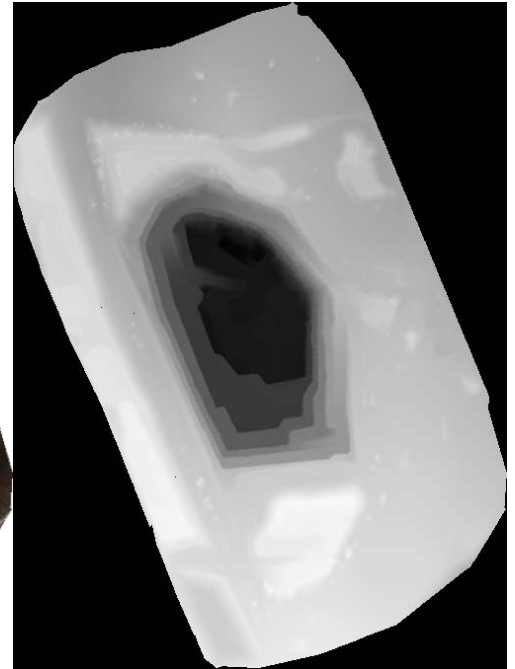# Cartography on Earth

Virtual Matterhorn

Mining Site

- Fully automated.
- Accurate.
- Inexpensive.

GCP statistics

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

EPFL

Applications

Cartography:
- Maps from aerial and satellite images

Robotics:
- Autonomous navigation
- Visual servoing

Industrial inspection
- Quality control

Security applications
- Access control
- Surveillance

Databases
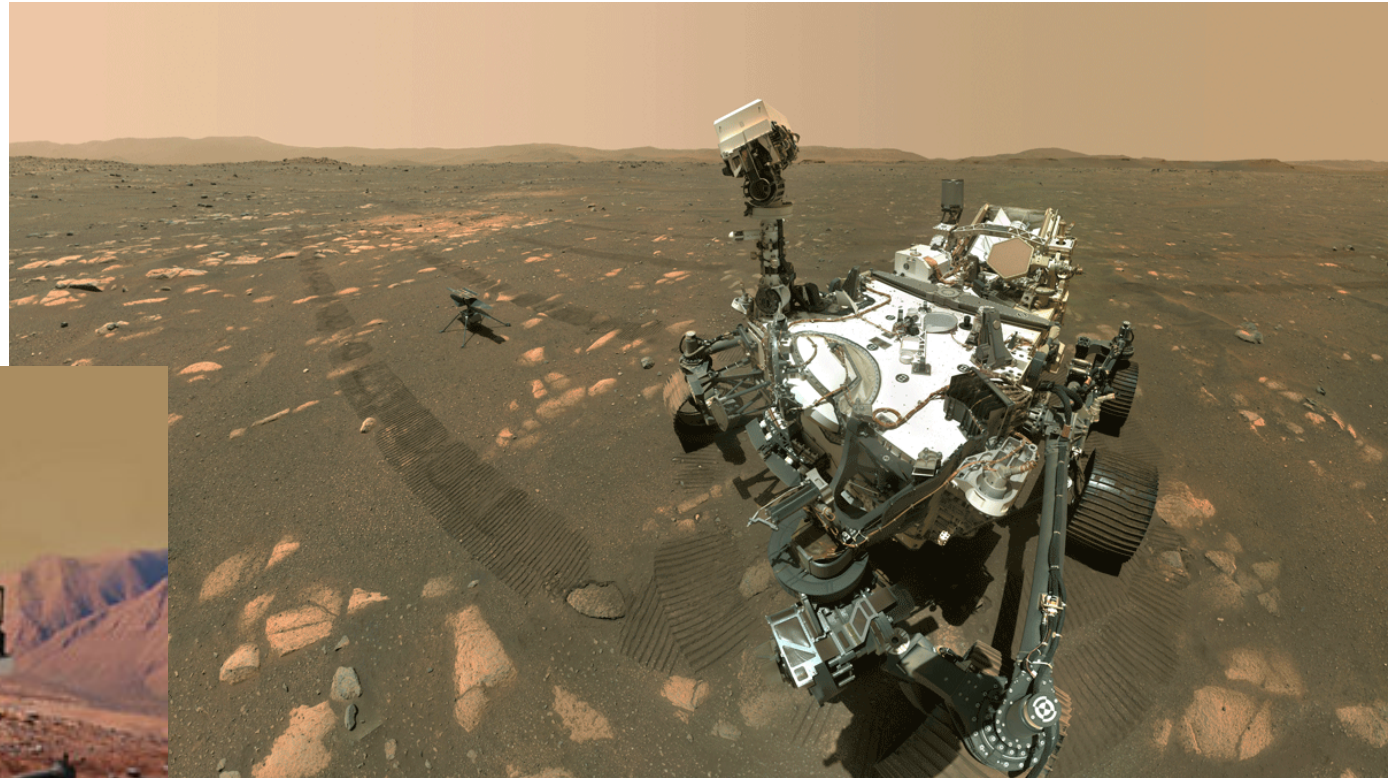- Retrieval and Annotation

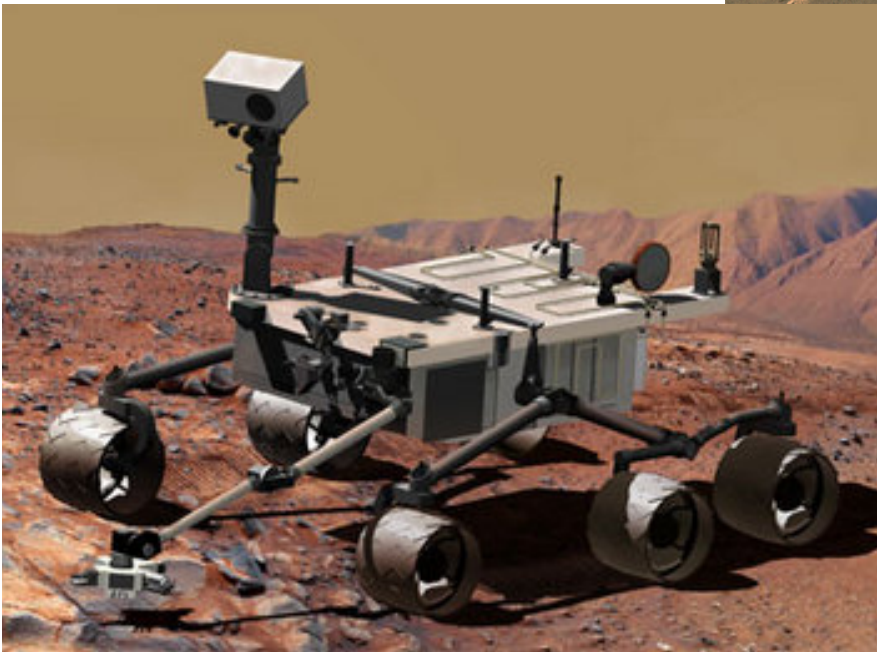**Medical Imagery**
- Microscopy
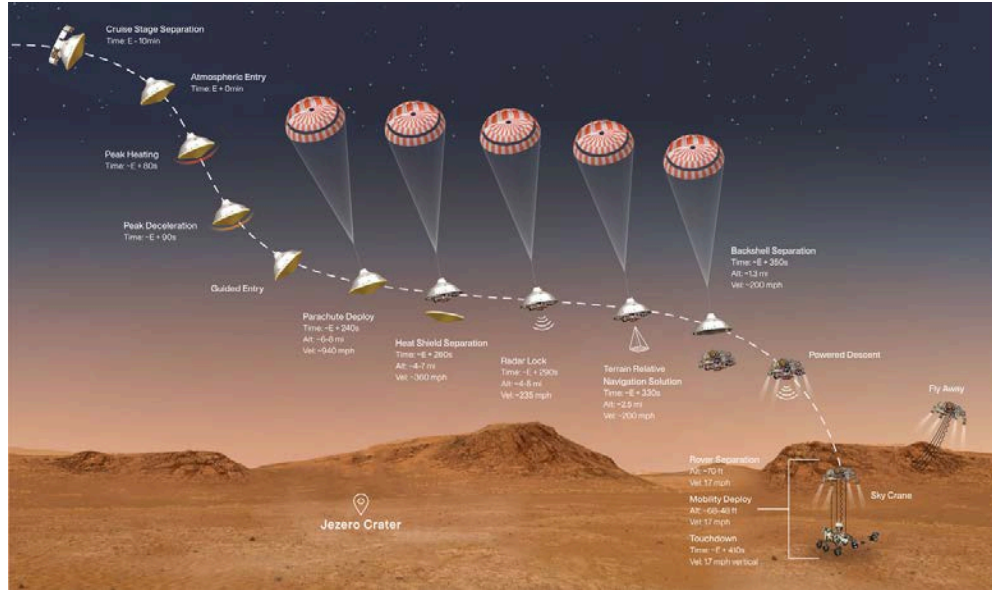
Mars Rovers

Opportunity
Landed 2004

Perseverance
Landed 2021

Curiosity
Landed 2012

Landing System Valid

Andrew Johnson (NASA)

Computer vision will play an unprecedented role in the landing, ensuring that the rover avoids such obstacles as boulder fields, dunes and crater walls in the final seconds of its seven-month journey to Mars.

https://www.ri.cmu.edu/cmu-robotics-alum-leads-development-of-critical-landing-technology-computer-vision-system-will-enable-safe-martian-landing-for-nasas-perseverance-rover/

Earth Rovers

| 1985 | 2007 | 2024 |
| DARPA ALV | DARPA Urban Challenge | Waymo Taxis |

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

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

Software embedded in the camera to find and read serial numbers

- Accurate localization
- Robustness to 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
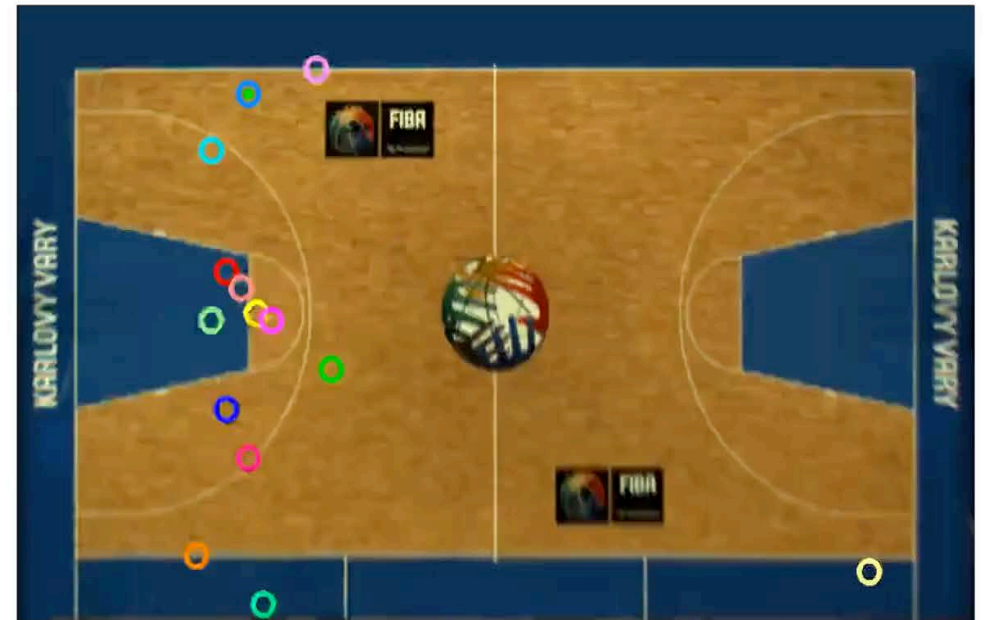
Databases
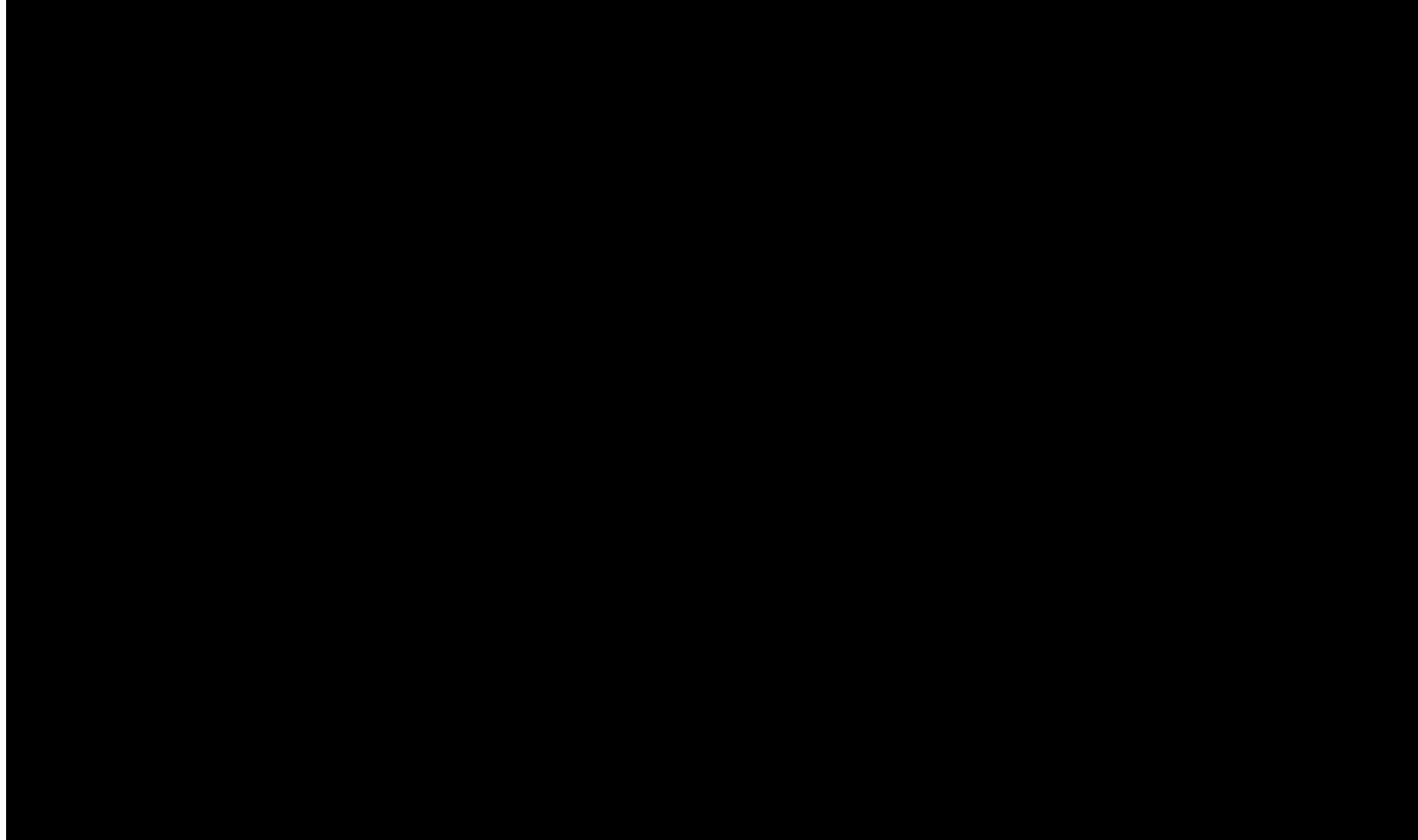- Retrieval and Annotation

Medical Imagery
- Microscopy

Crowd Counting

EPFL at lunchtime: The colors denote crowd density.

Liu et al. , CVPR'22

Tracking People

... and the ball → Behavioral analysis.

BenShitrit et al. PAMI'14

Tech Transfer

- 2005: First ICCV paper published.
- 2017: System deployed in NBA arenas.
- 2019: Premier League Optical Tracking Provider.

—> It takes time!

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

Identify a picture of EPFL in a large database or on the web **without** words.

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
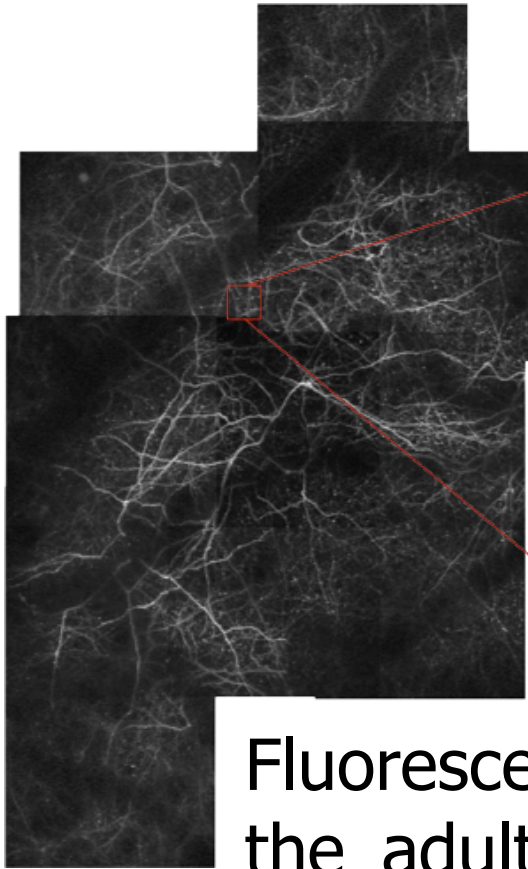
Fluorescent neurons in vivo in the adult mouse brain Imaged through a cranial window using a 2-photon microscope.
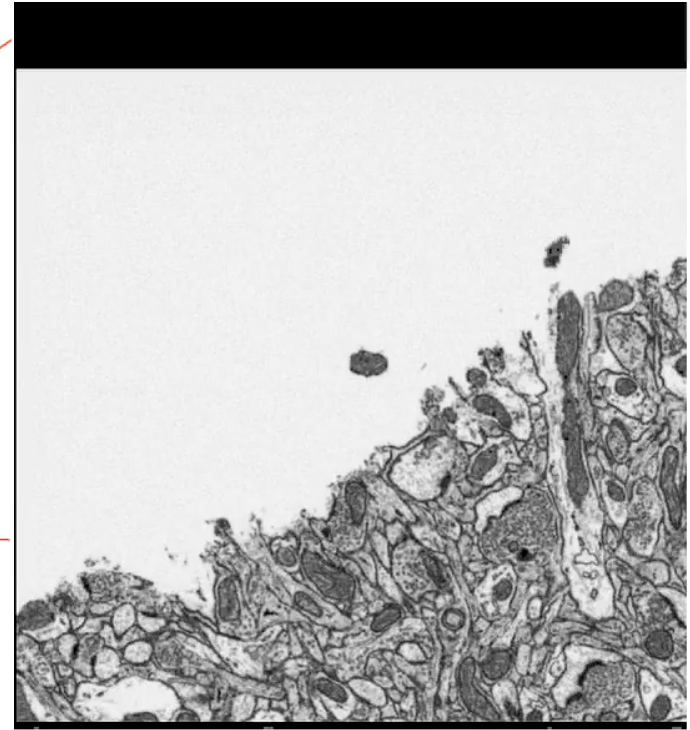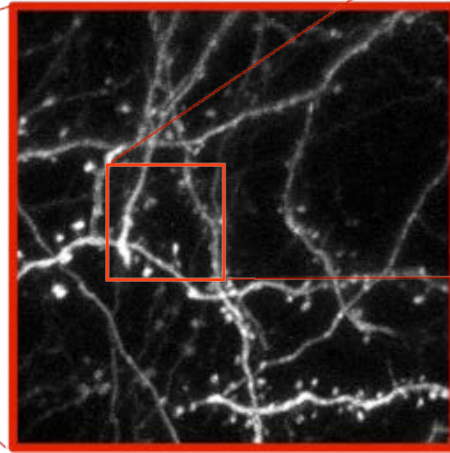
Electron Microscopy Image Stack at five nanometer resolution.

Delineating Dendritic Trees
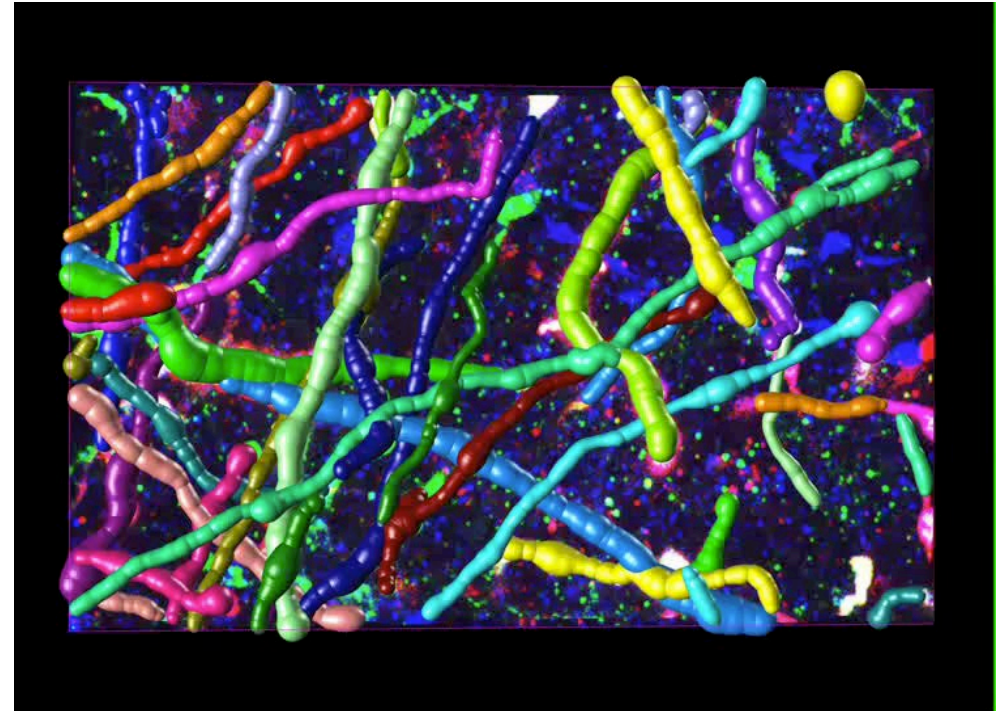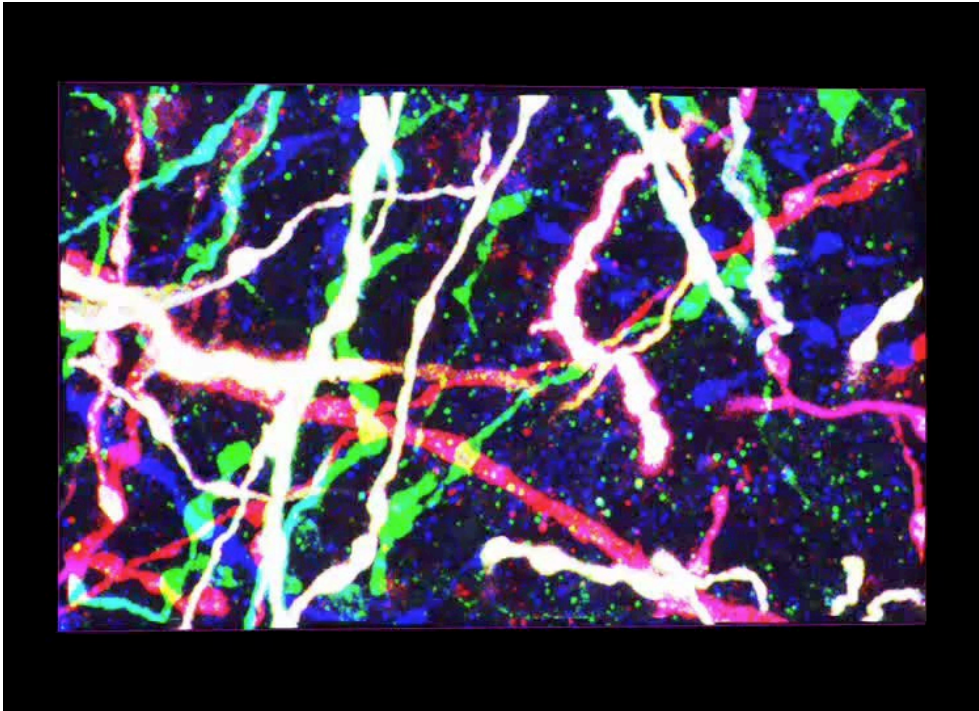
Finding Mitochondria
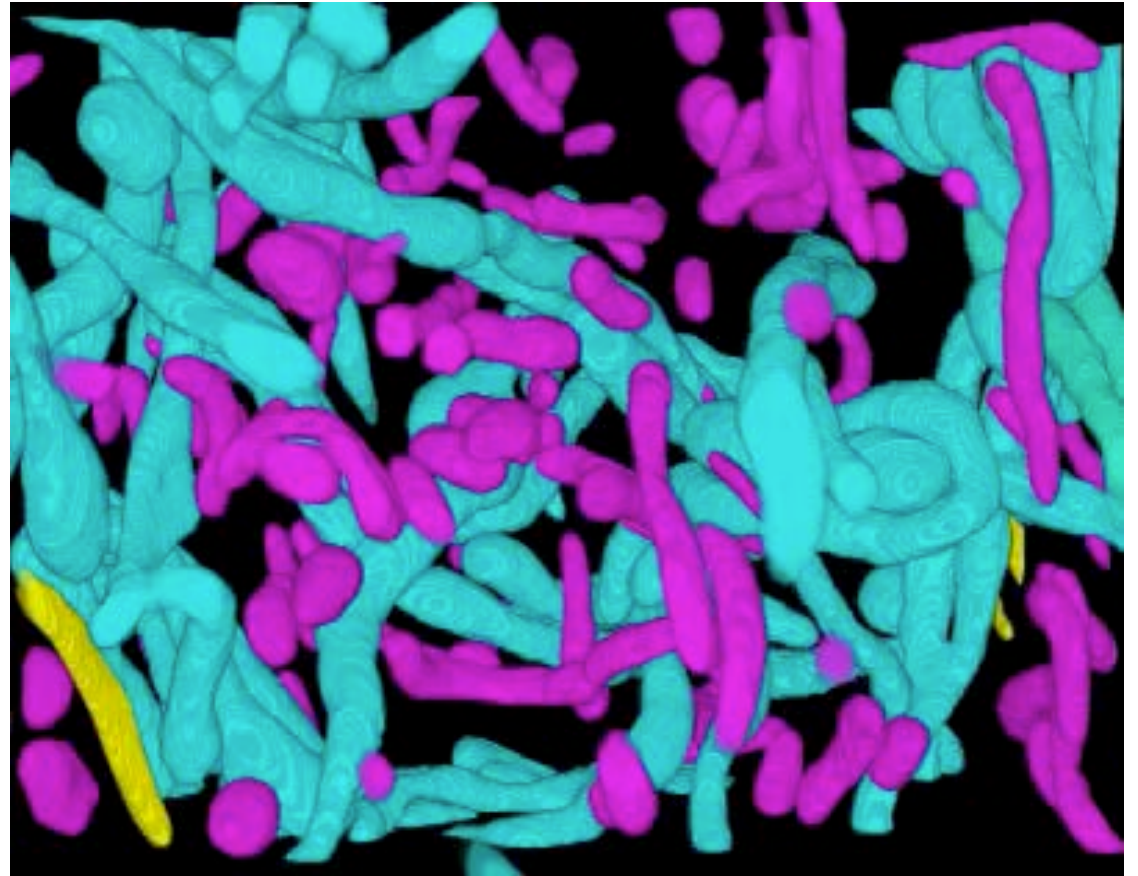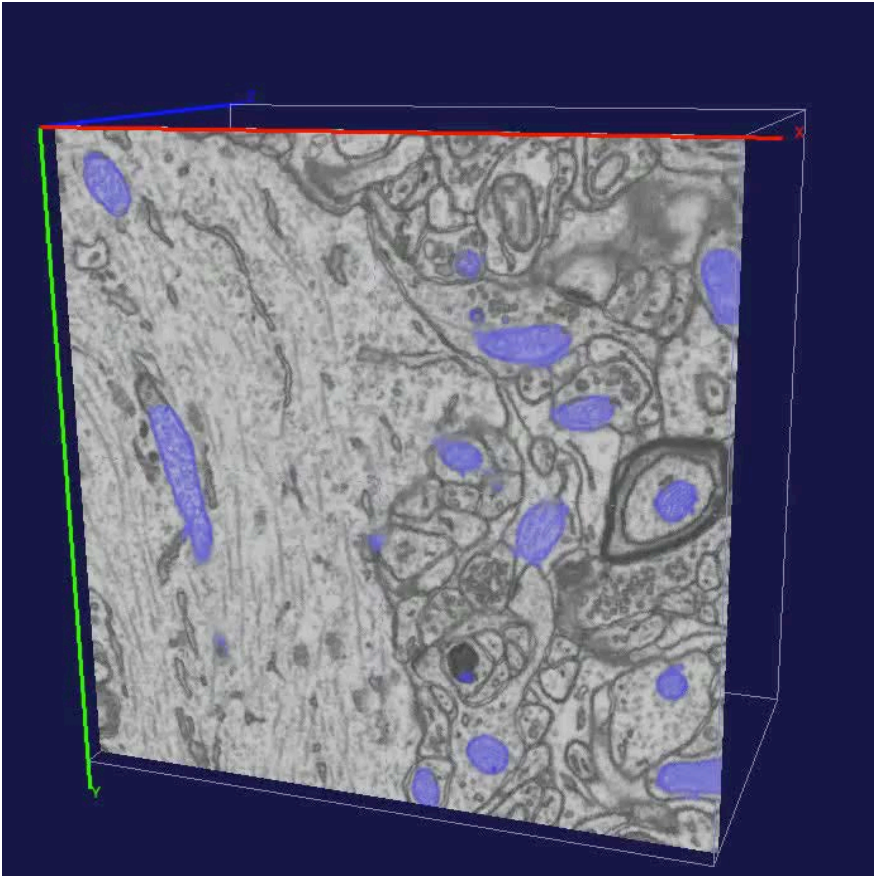
Google Earth For The Brain

- 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 20!

—> Seriously big data!

KINECT
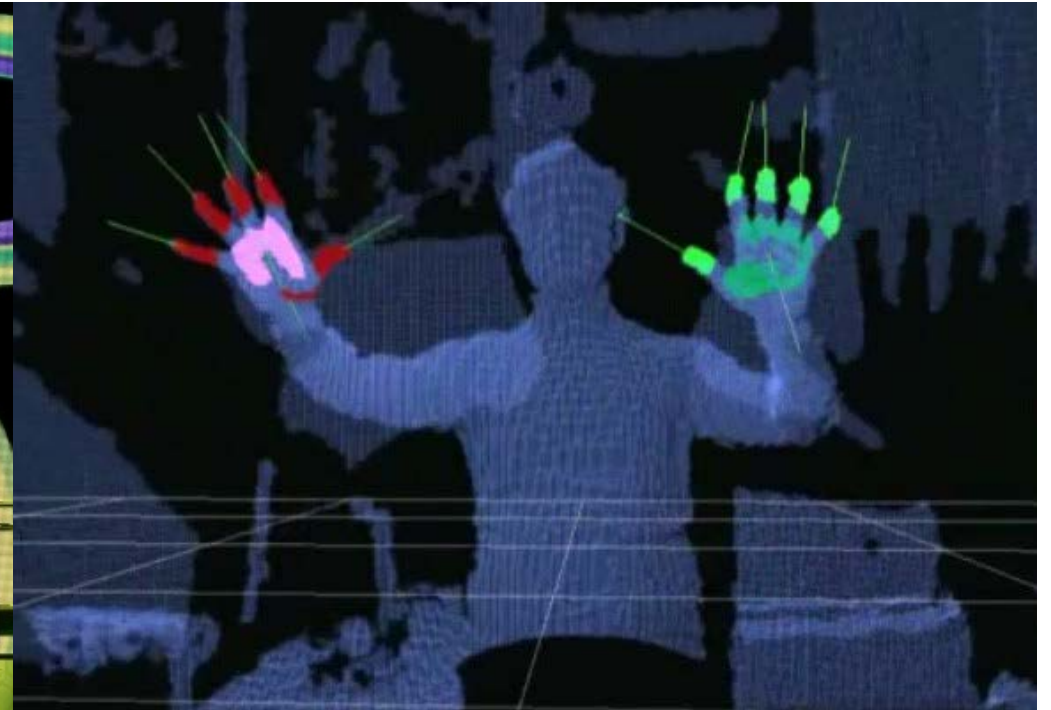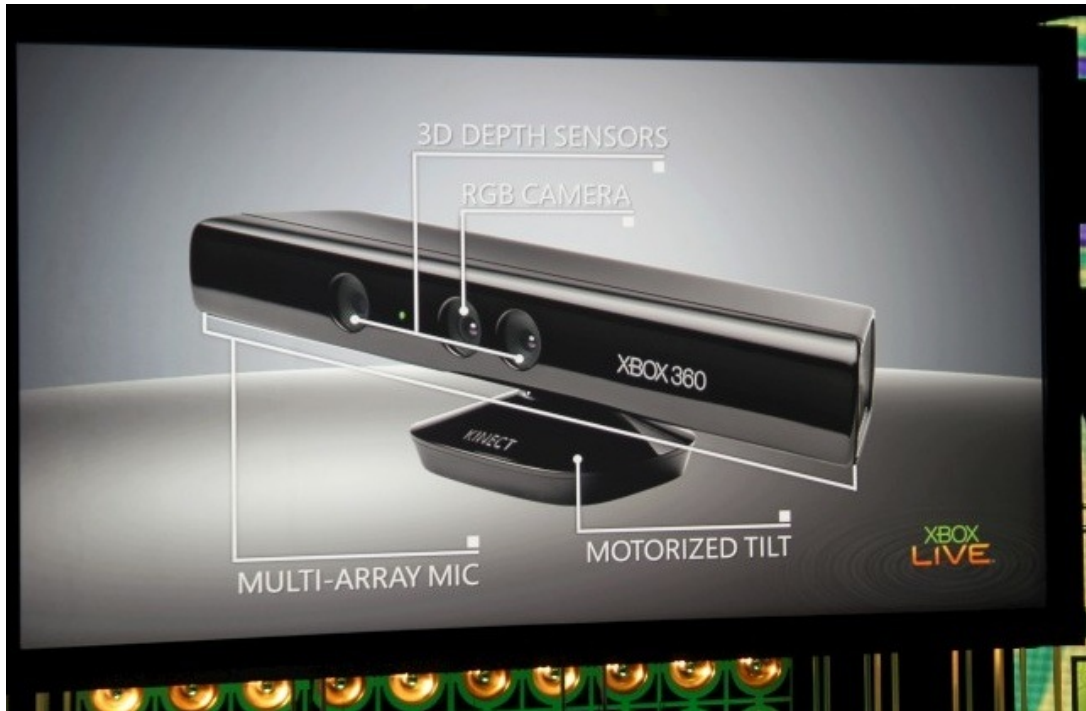
→ Image whose pixel values are distances

How it Began

- Computer Vision started i[n] short term project.

- A world of perfect blocks assumptions.

→ The real world is not like
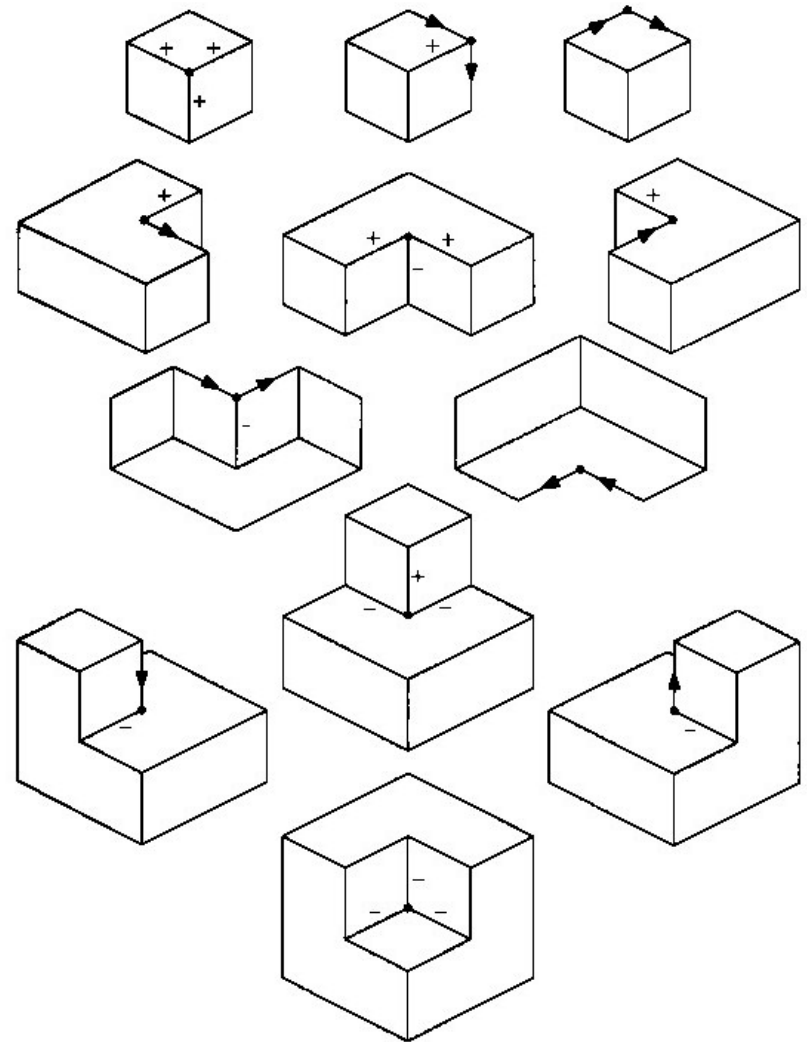
Historical Perspective

- 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: Vision-Language Models ??

--> Improved understanding and successful applications in graphics, mapping, biometrics, and many others.

Human vs Machine Learning

Learn from experience

Learn from experience

Recognizing Hand-Written Digits
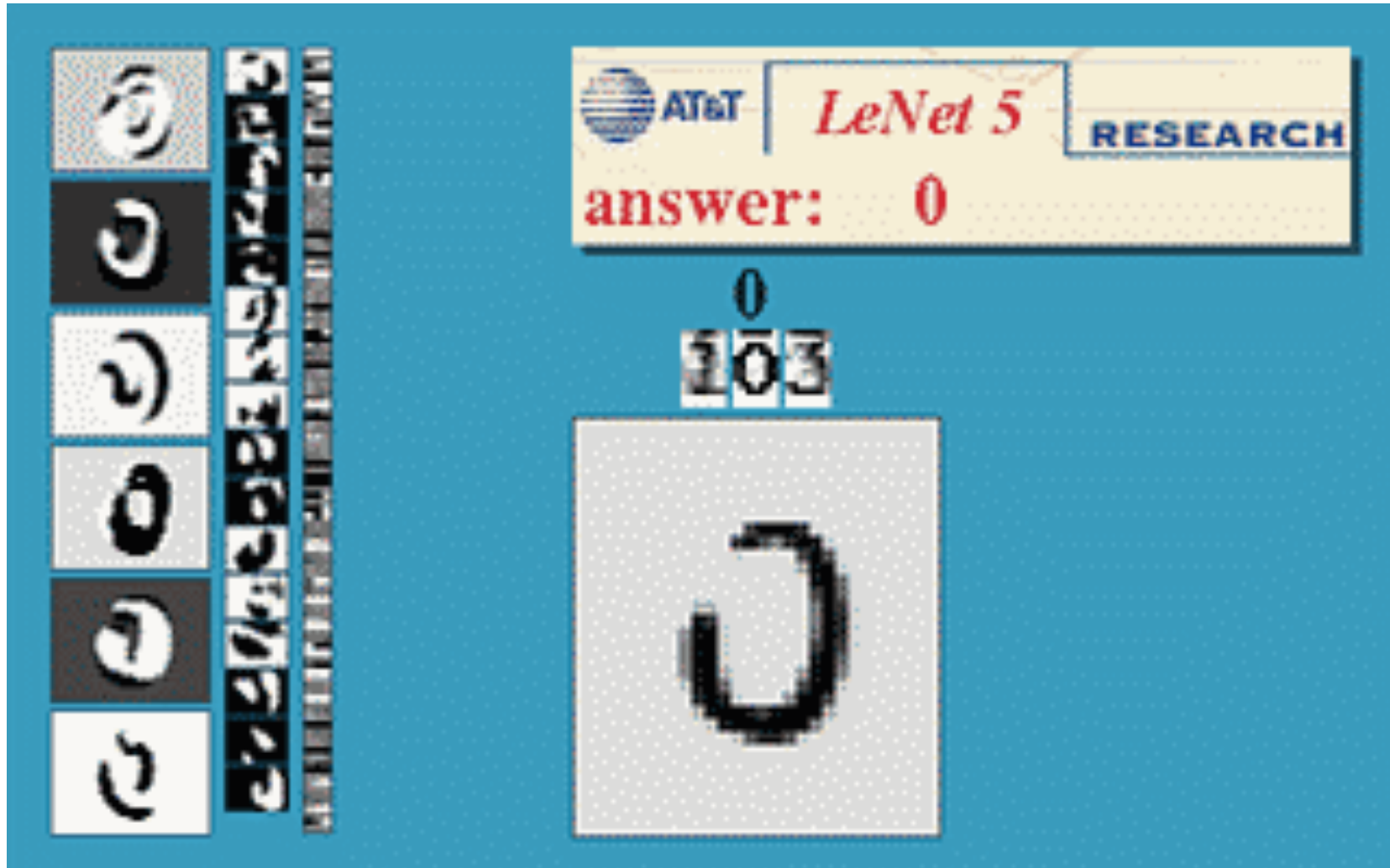
LeNet (1989-1999)

Recognizing Hand-Written Digits

Predictor and Labels

28x28 pixels

X is a 784-D Vector
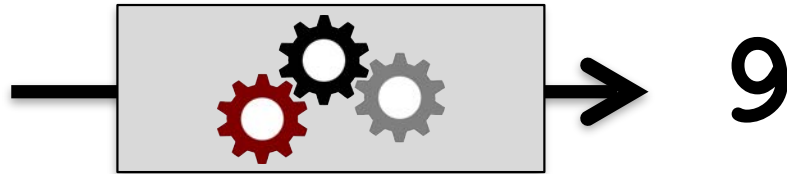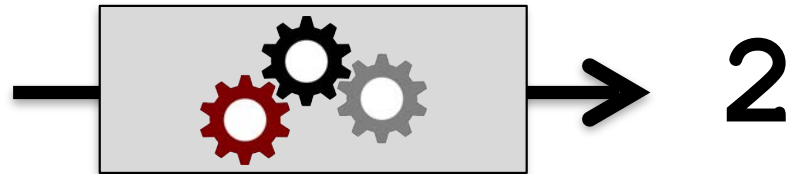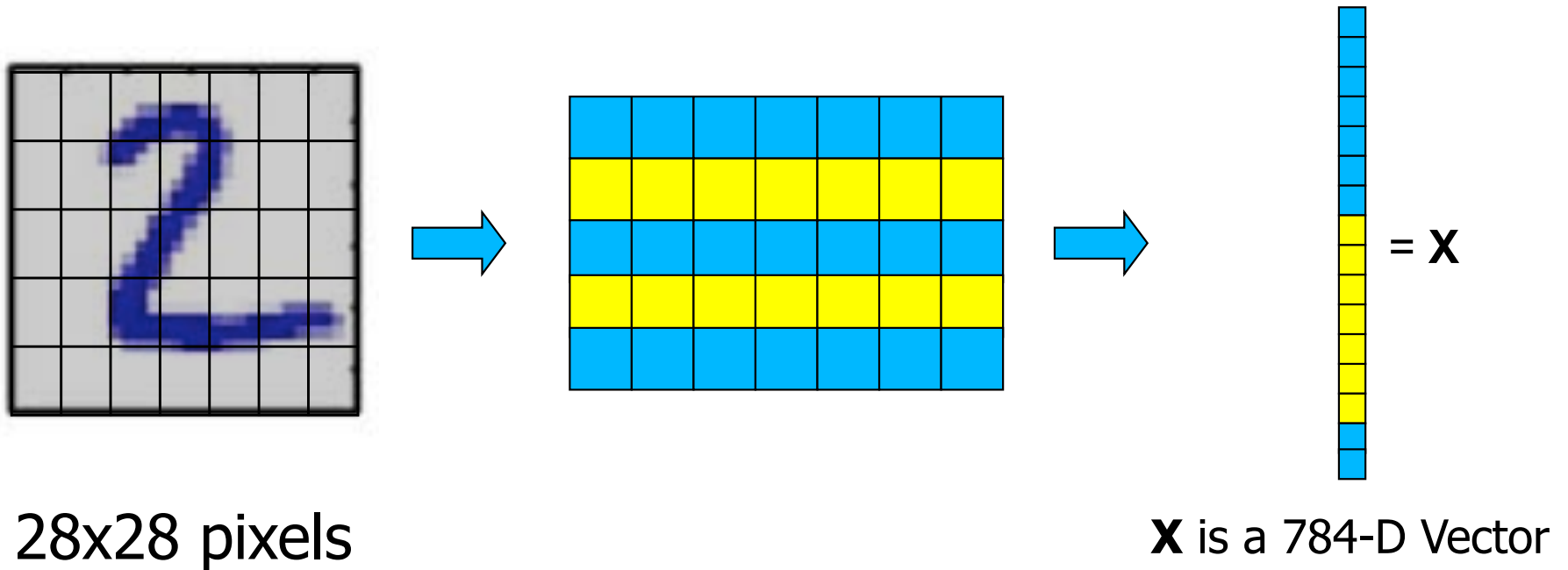
$$y : \mathbf{x} \in \mathbb{R}^{784} \rightarrow \{0, 1, 2, \cdots, 9\} \quad ?$$

Predictor

Labels

Convolutional Neural Networks
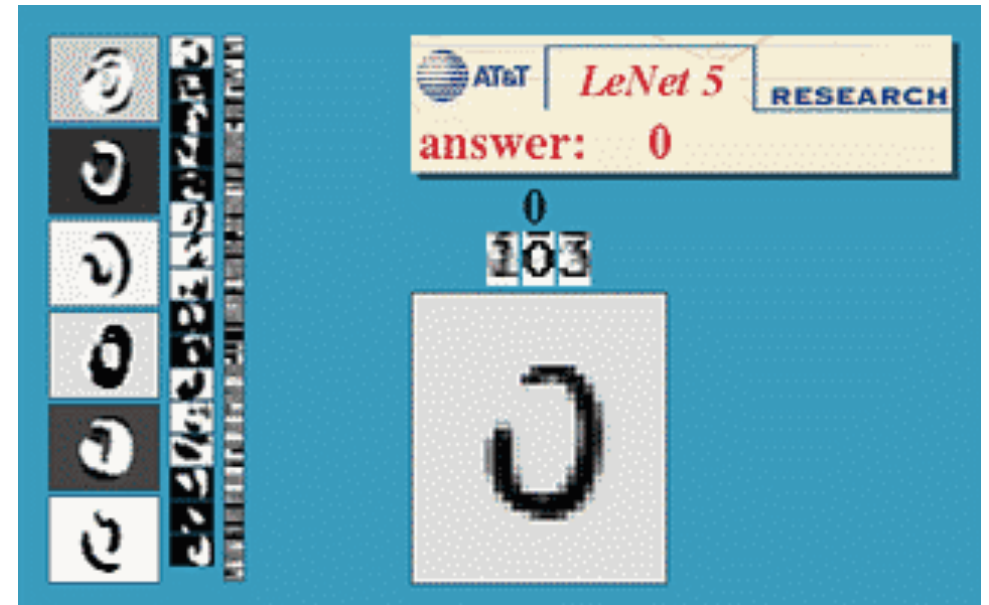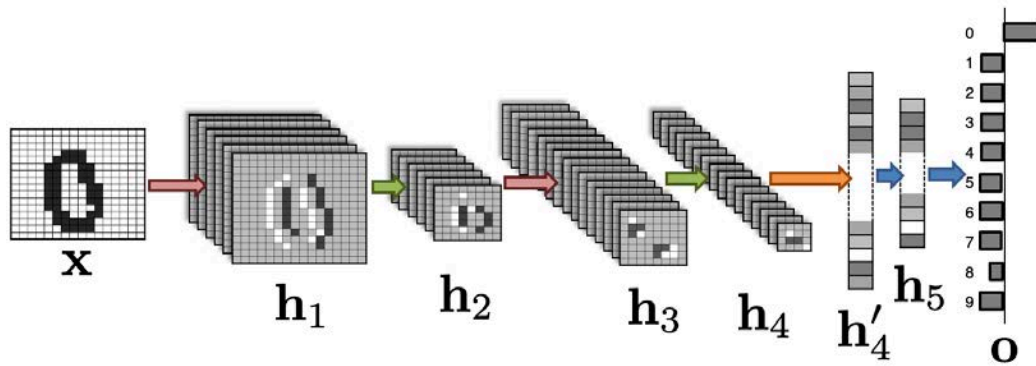
- Powerful way to encode the function y.
- But require much training data.

—> We will discuss them later in the class.

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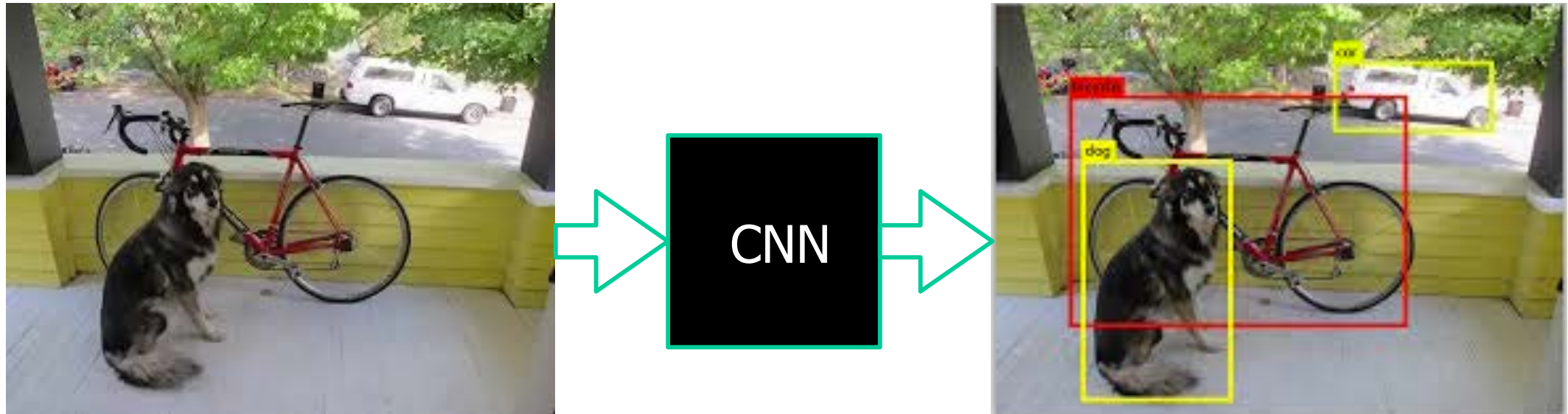
- Extremely effective in many cases but does not shed much light on the vision process.
- The best algorithms combine Deep Nets with more traditional techniques.

Redmon et al. , CVPR'16

A Teachable Scheme

```
                        ┌─────────────┐
                        │   Image(s)  │
                        └──────┬──────┘
          ┌────────────────────┼────────────────────┐
┌──────────────────┐ ┌──────────────────┐ ┌──────────────────┐
│ Edge information │ │Texture information│ │ Shape information │
└─────────┬────────┘ └─────────┬────────┘ └─────────┬────────┘
          │         ┌──────────────────┐            │
          └─────────│  Scene objects   │────────────┘
                    └─────────┬────────┘
                    ┌──────────────────────┐
                    │  Scene interpretation │
                    └──────────────────────┘
```
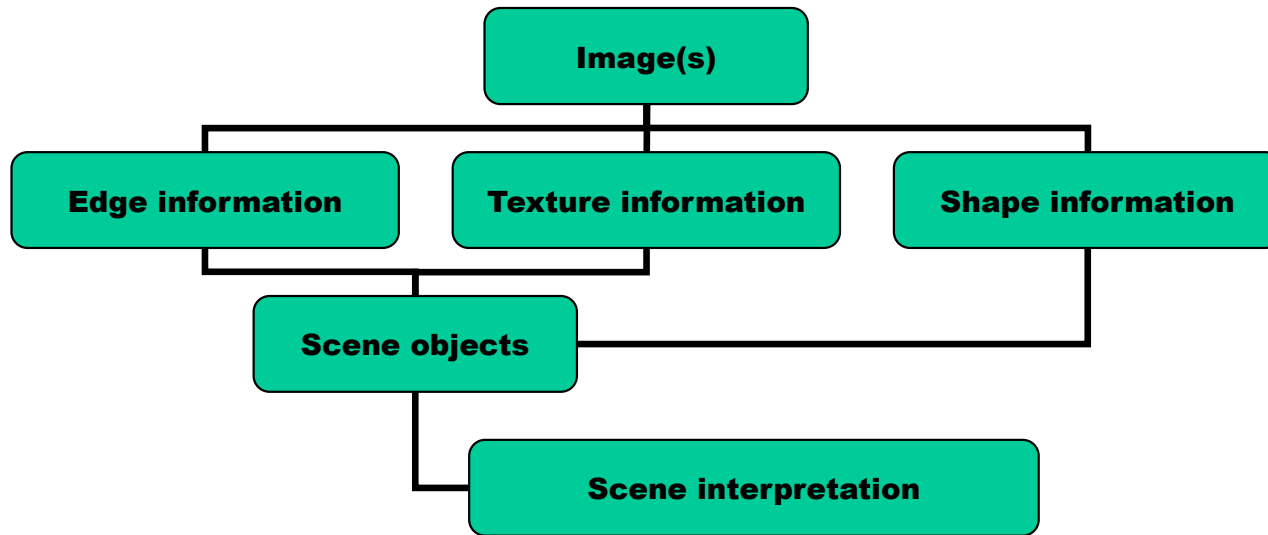
Decomposition of the vision process into smaller manageable and implementable steps.

--> Paradigm followed in this course

--> May not be the one humans use

Course Outline

Introduction:
- Definition
- Human vision
- Image formation

Extracting features:
- Contours
- Texture
- Regions

Shape recovery:
- From one image
- Using additional images

Course Organization

- Formal lectures every week (Monday)
- Exercises every other week (Tuesday)
  - Two of them will be graded (10% of the grade, each).
  - Dates are on the moodle page.
  - Bring your laptops for the exercises.
- Written exam (80% of the grade).

Course Material

Textbooks:

- R. Szeliski, Computer Vision: Computer Vision: Algorithms and Applications, 2021.
- A. Zisserman and R. Hartley, Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.

Web pages:

- moodle.epfl.ch   (Computer Vision, CS-442)
- cvlab.epfl.ch/projects   (Projects)
- cvlab.epfl.ch/research (Research)

Slide Codes

Training vs Testing

Normal slide: It is part of the course and I may ask exam questions about it.

Training vs Testing

Reminder

Reminder slide: We have already covered this earlier in the class. Go back to the appropriate lecture if you do not remember.

Training vs Testing

Optional   Bishop, xxx

Optional slide: This is additional material for people interested in more details. I will not ask direct exam questions on this.

Reference to book or paper for even more details.