universität innsbruck

Lecture 1. Introduction

703142. Computer Vision

Assoz.Prof. Antonio Rodríguez-Sánchez, PhD.

Who, when and where

- Antonio Rodriguez-Sanchez, PhD
  - Antonio.rodriguez-sanchez@uibk.ac.at
  - Room 3N05
- Mondays, 11:15-14:00
  - HS 11

Prerequisites

- Familiarity with
  - Linear algebra
  - Probability
  - Coding: Python or C/C++, OpenCV
  - Mathematical language
  - 703089 Visual Computing

Course requirements

- Four problem sets with lab exercises in C++ or Python with OpenCV.
- Midterm
- Final exam
  - Final exam will include everything covered in the course, specially subjects after midterm

Grading

- Grades are 1-5
  - 1: Very good
  - 5: Fail
- Grade Exam (45%)
  - Final exam
- Grade Assignments (55%)
  - Four sets of problems which include exercises and programming

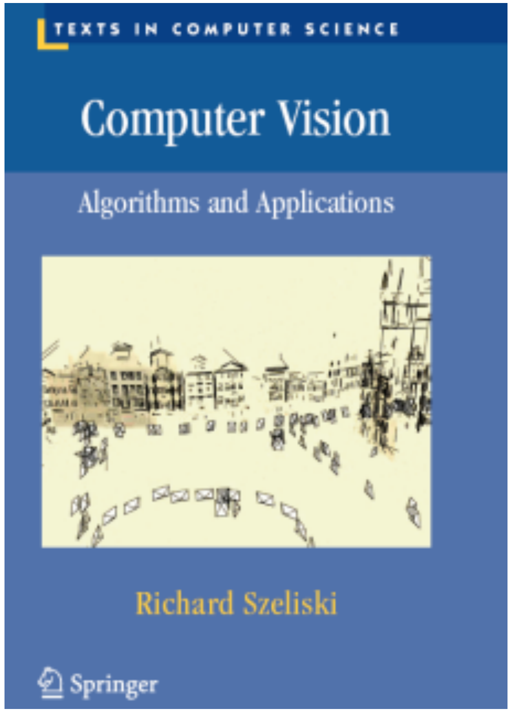
Grading

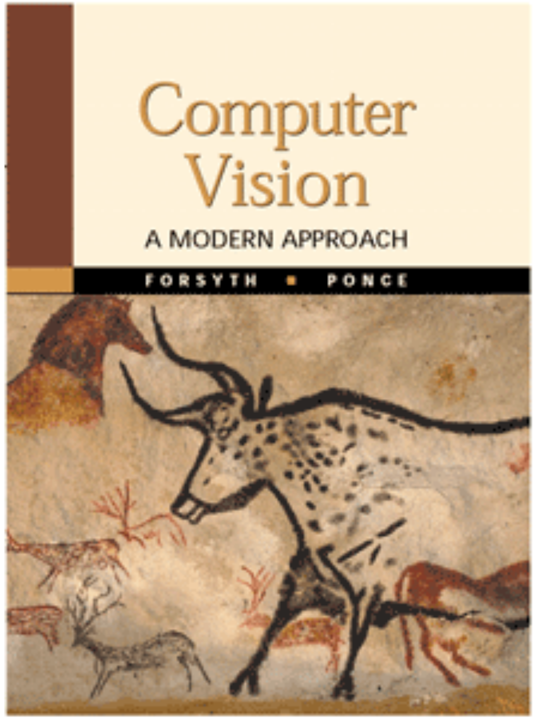
- Attendance is mandatory
  - 2 Jokers available
- Sickness absences
  - Send medical note
  - PCR test

Collaboration policy

- Problems may be discussed, but all written work and coding must be done individually.
- Please note on your problems who you discussed them with.
- Individuals found submitting duplicate or substantially similar materials may get a 5 in this course and other sanctions.

Recommended readings





http://luthuli.cs.uiuc.edu/~daf/book/book.html

http://szeliski.org/Book/

http://www.computervisionmodels.com/

Tentative schedule

Date	Topic
06-03-2021	Introduction.
13-03-2021	Review: Image Formation. Introduction to OpenCV.
20-03-2021	Review: Image processing.
27-03-2021	Review: Feature extraction.
17-04-2021	Motion I: Optical Flow. OpenCV.
24-04-2021	Motion II: Spatiotemporal filters. Spatiotemporal analysis.
08-05-2021	Stereopsis I: Correspondence. RANSAC.
15-05-2021	Stereopsis II: Reconstruction. Rendering.
22-05-2021	Object Recognition I: Categories. Photo editing.
05-06-2021	Object Recognition II: Objects, faces, instances. Viola and Jones
12-06-2021	Object Recognition III: Deep Learning. CNNs in Pytorch.Exam questions
19-06-2021	Invited talk: Neuroscience in stereo and motion.
26-06-2021	FINAL EXAM.

Other courses …

- … you may be interested in the future
  - Master course in Computer Vision (703328). SS. assoz.Prof. Antonio Rodriguez-Sanchez
  - Deep Learning (703134). WS. assoz.Prof. Antonio Rodriguez-Sanchez
  - Bachelor thesis: Talk to me
  - Check the IIS webpage: https://iis.uibk.ac.at/

Why Vision?

- What does it mean to see?
  - “To know what is where by looking” (Aristotle)
- How to discover from images what is present in the world, where things are, what actions are taking place

Vision. Marr, 1982. Freeman Publishers

Why Vision?

- Vision is an important source of information
  - About the surrounding world
    - Object identity, location and dynamics
  - Supports intelligent interaction with the environment
    - Navigation, manipulation, decision making
  - Information derived without physical contact
    - Optical data is acquired at a distance
    - Enables unobtrusive sensing

Source: Richard Wildes slides

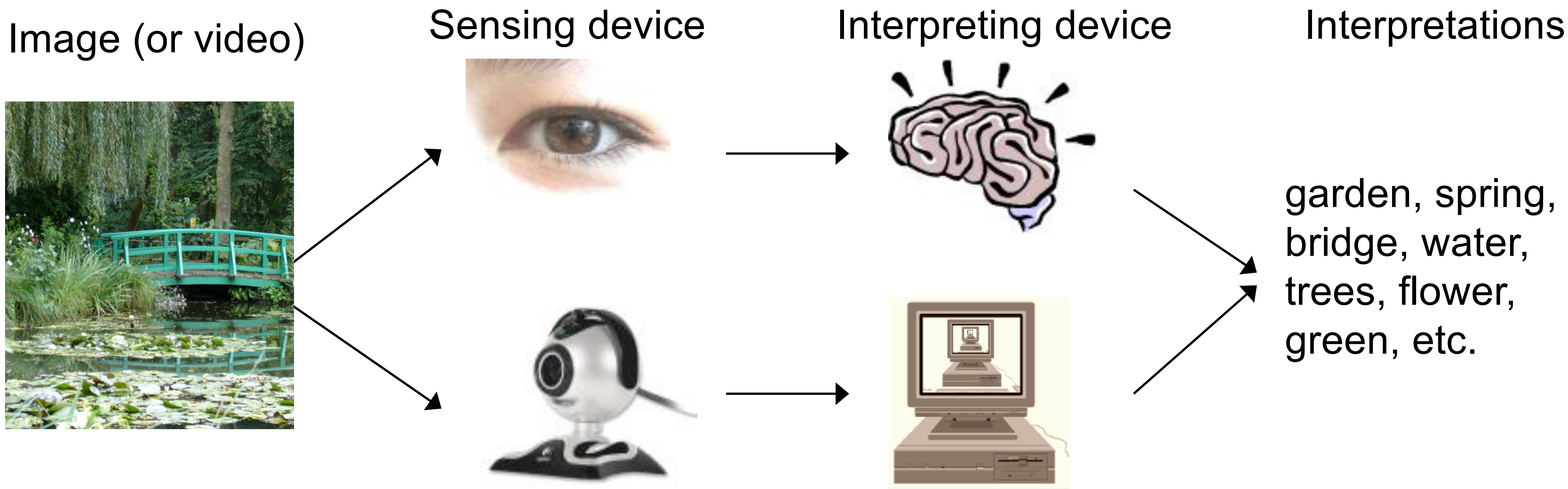
Properties of Vision

- 3D representations are easily constructed
  - There are many different cues.
  - Useful
    - to humans (avoid bumping into things; planning a grasp; etc.)
    - in computer vision (build models for movies).
  - Cues include
    - multiple views (motion, stereopsis)
    - texture
    - shading

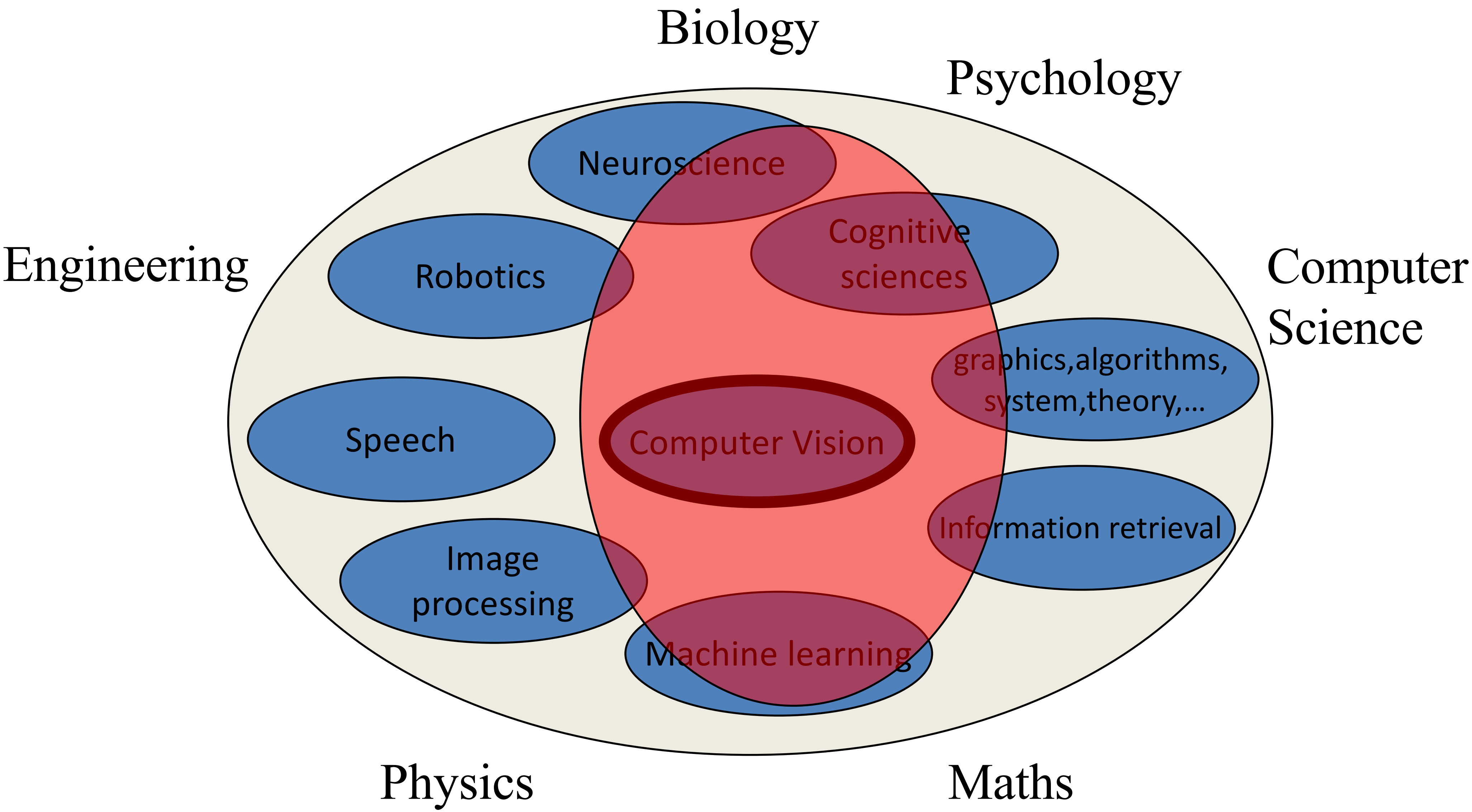
Properties of Vision

- People draw distinctions between what is seen
  - “Object recognition”
  - This could mean “is this a fish or a bicycle?”
  - It could mean “is this emperor Charles V?”
  - It could mean “is this poisonous or not?”
  - It could mean “is this slippery or not?”
  - It could mean “will this support my weight?”
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

What is (computer) vision?



What is computer vision related to?



The goal of computer vision

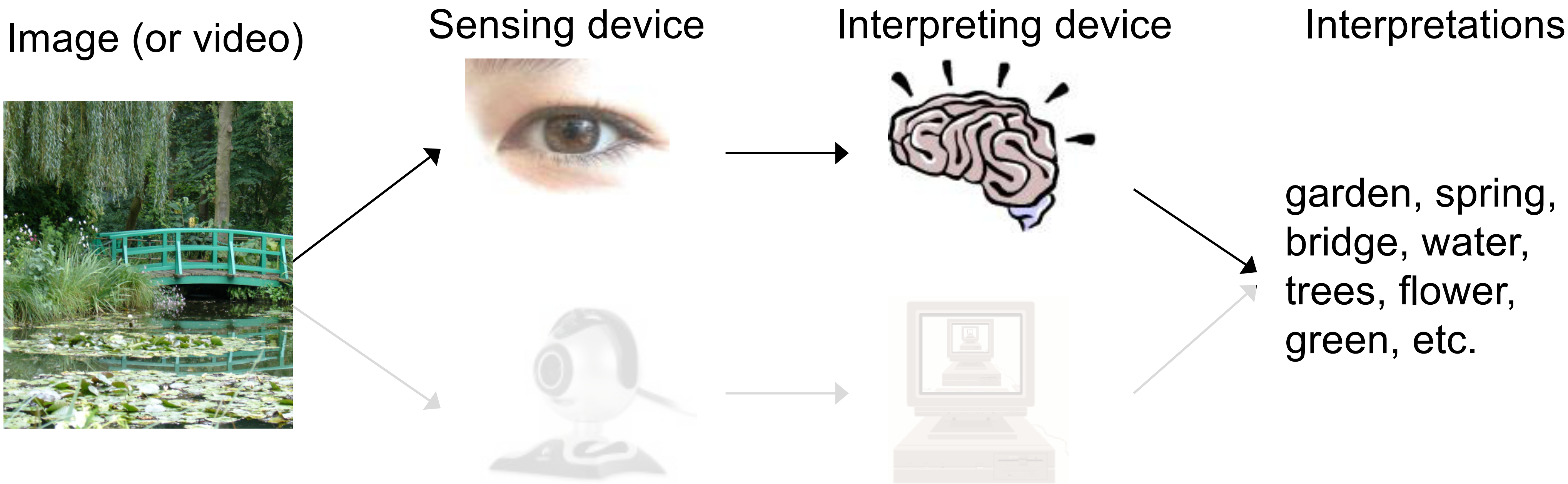
- To bridge the gap between pixels and “meaning”

What we see

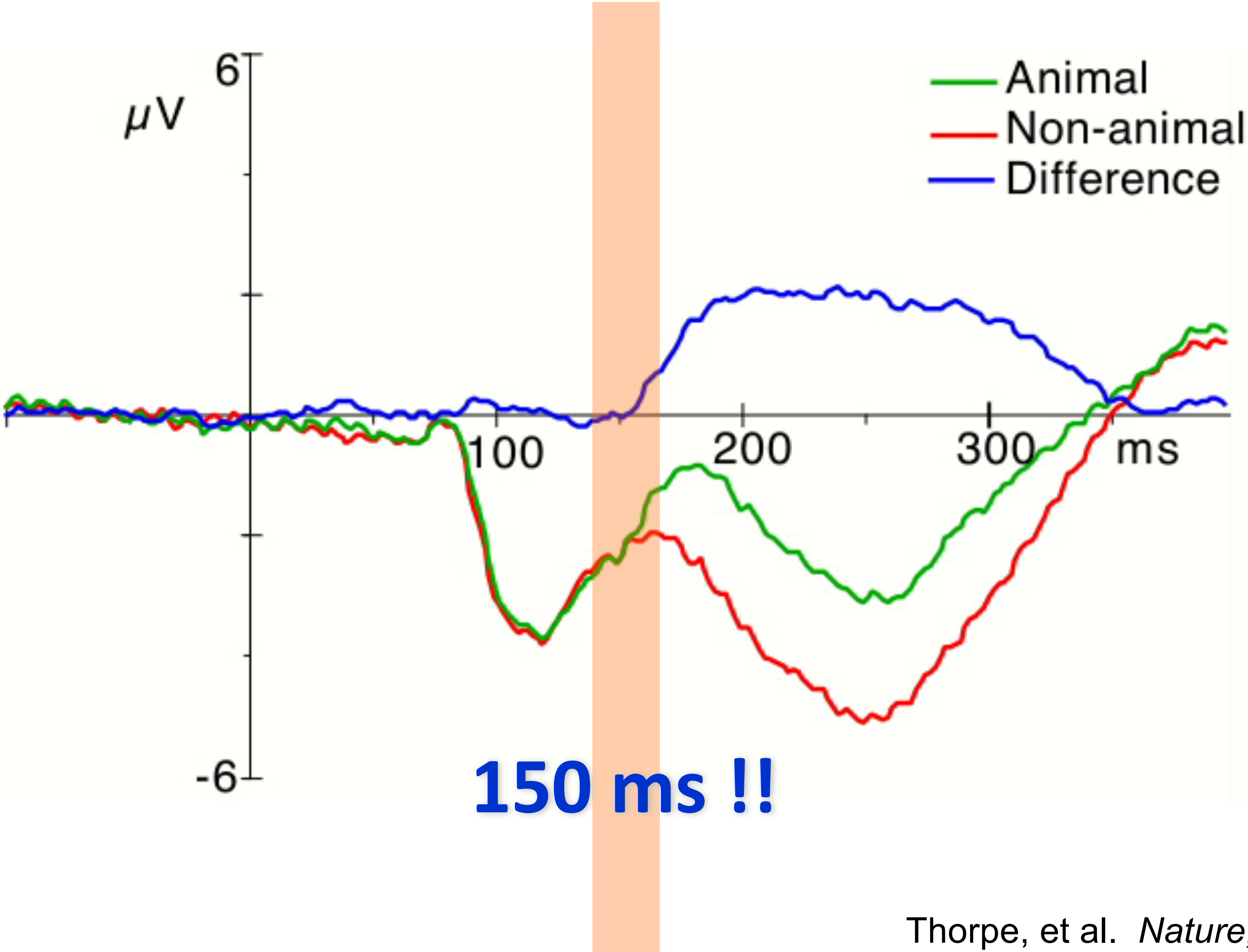
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

What is vision?



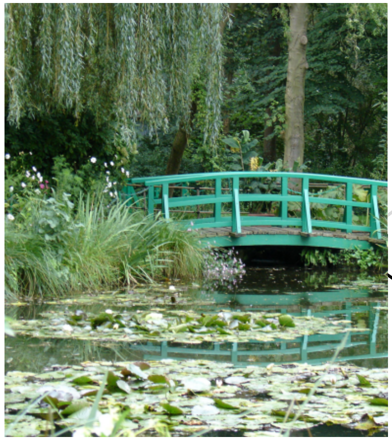
What is vision?



Thorpe, et al. *Nature,* 1996

What is computer vision?

Image (or video)



Sensing device

Interpreting device

Interpretations

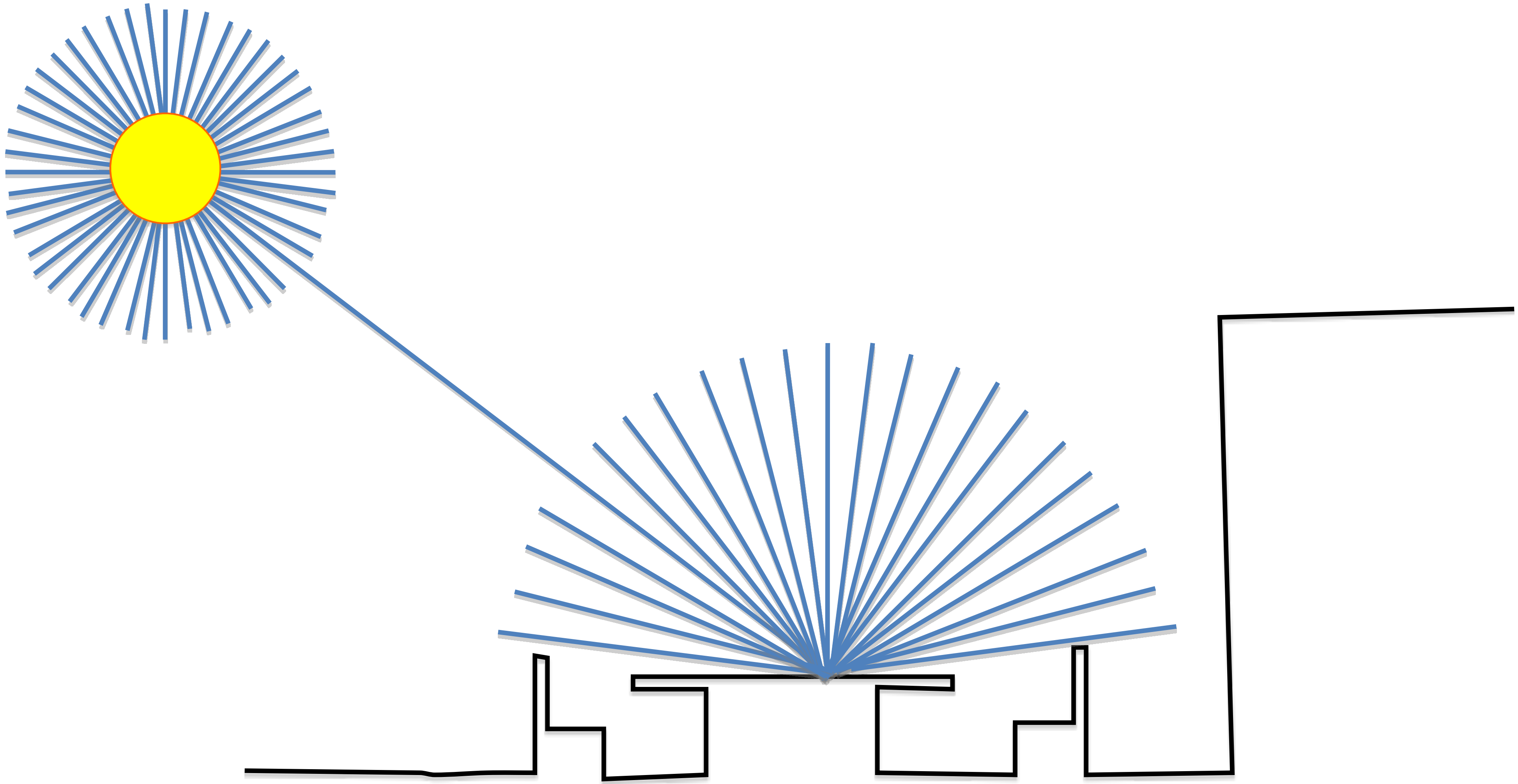
garden, spring, bridge, water, trees, flower, green, etc.

Outline

- Why is vision hard?
- Applications
- Allied areas
- History

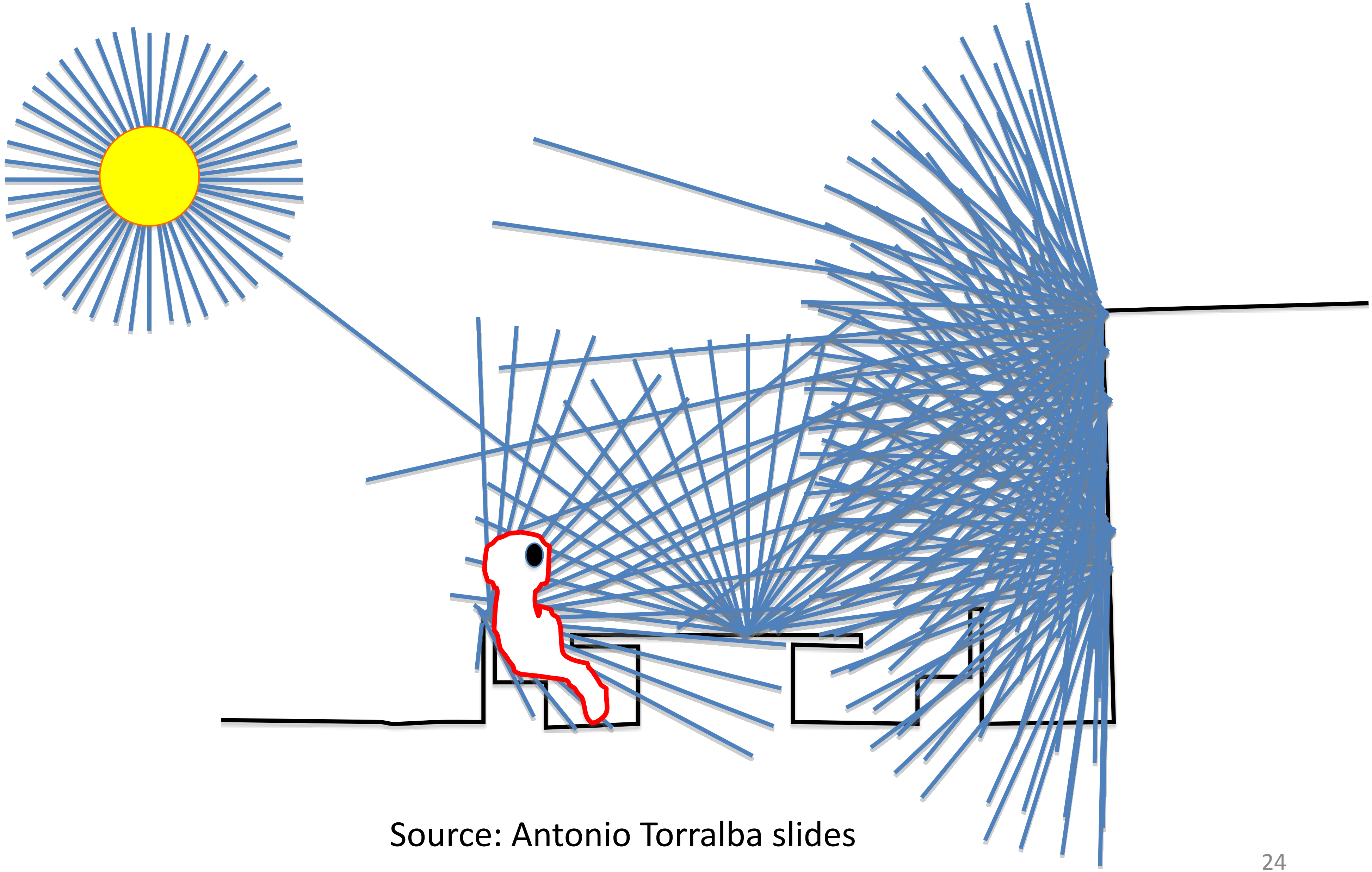
Why is vision hard?

The structure of ambient light



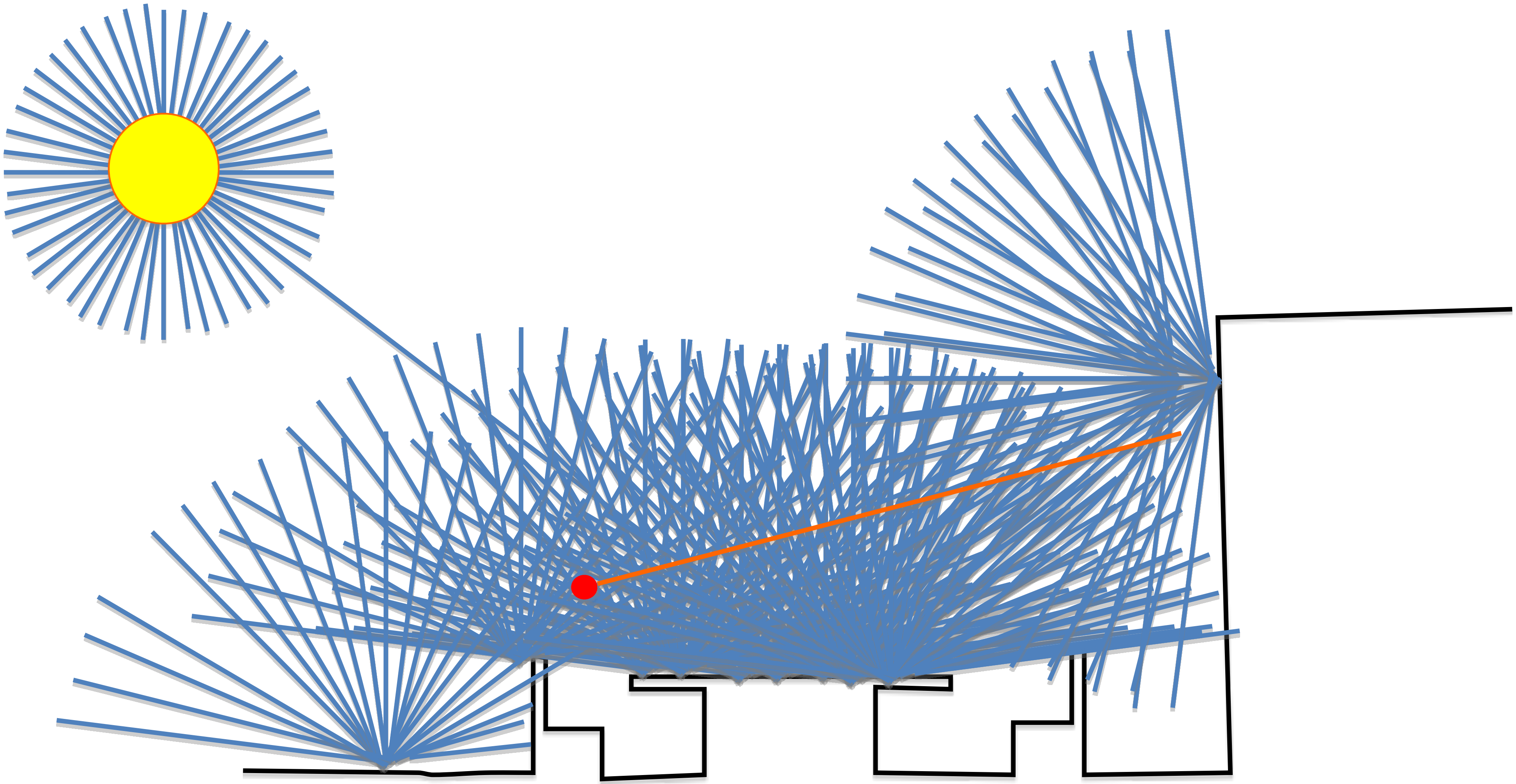
Source: Antonio Torralba slides

The structure of ambient light



Source: Antonio Torralba slides

The plenoptic function



Source: Antonio Torralba slides

The plenoptic function

- What information about the world is contained in the light filling a region of space?
  - Space is filled with a dense array of light rays of various intensities.
  - The set of rays passing through any point in space is mathematically termed a *pencil.*

Adelson and Bergen, 1991

The plenoptic function

- Leonardo da Vinci refers to this set of rays as a "radiant pyramid":

  The body of the air is full of an infinite number of radiant pyramids caused by the objects located in it. These pyramids intersect and interweave without interfering with each other during the independent passage throughout the air in which they are infused. (Kemp, 1969)

Adelson and Bergen, 1991

The plenoptic function

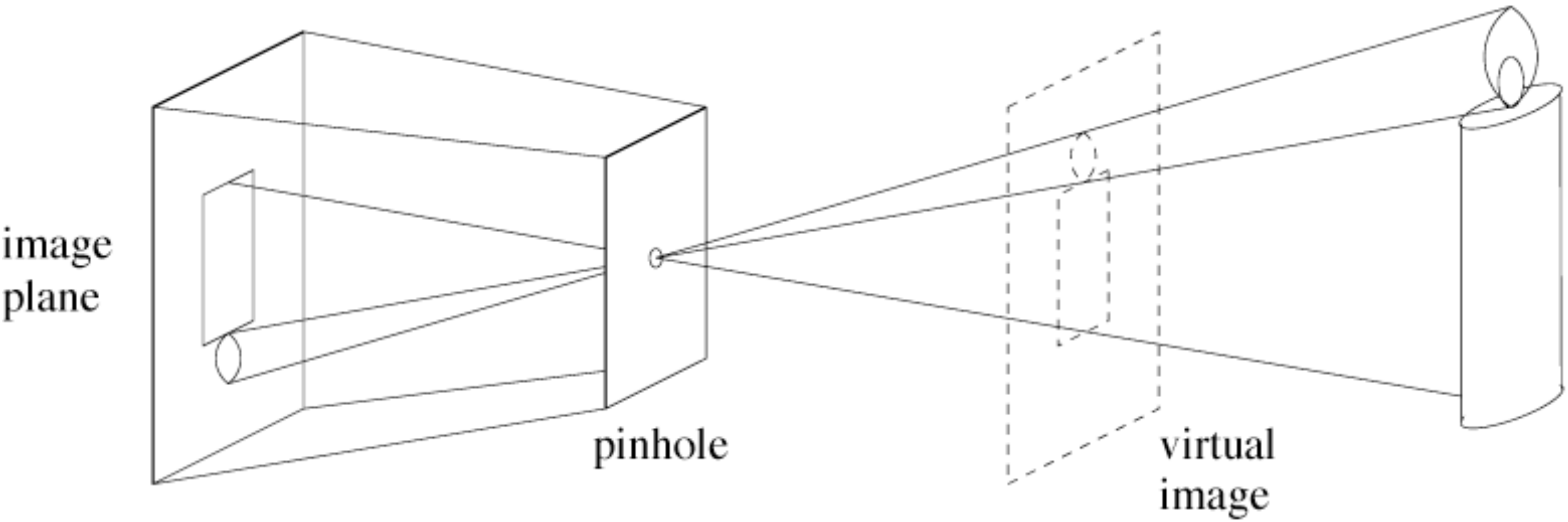
- "The significance of the plenoptic function is this: The world is made of 3D objects, but these objects do not communicate their properties directly to an observer. Rather, the objects fill the space around them with the pattern of light rays that constitutes the plenoptic function, and the observer takes samples from this function."

$$P = P(\theta, \phi, \lambda, t, V_x, V_y, V_z).$$

Adelson and Bergen, 1991

Pinhole cameras

- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice



Why is vision hard?

Some things have strong variations in appearance

Source: Antonio Torralba slides

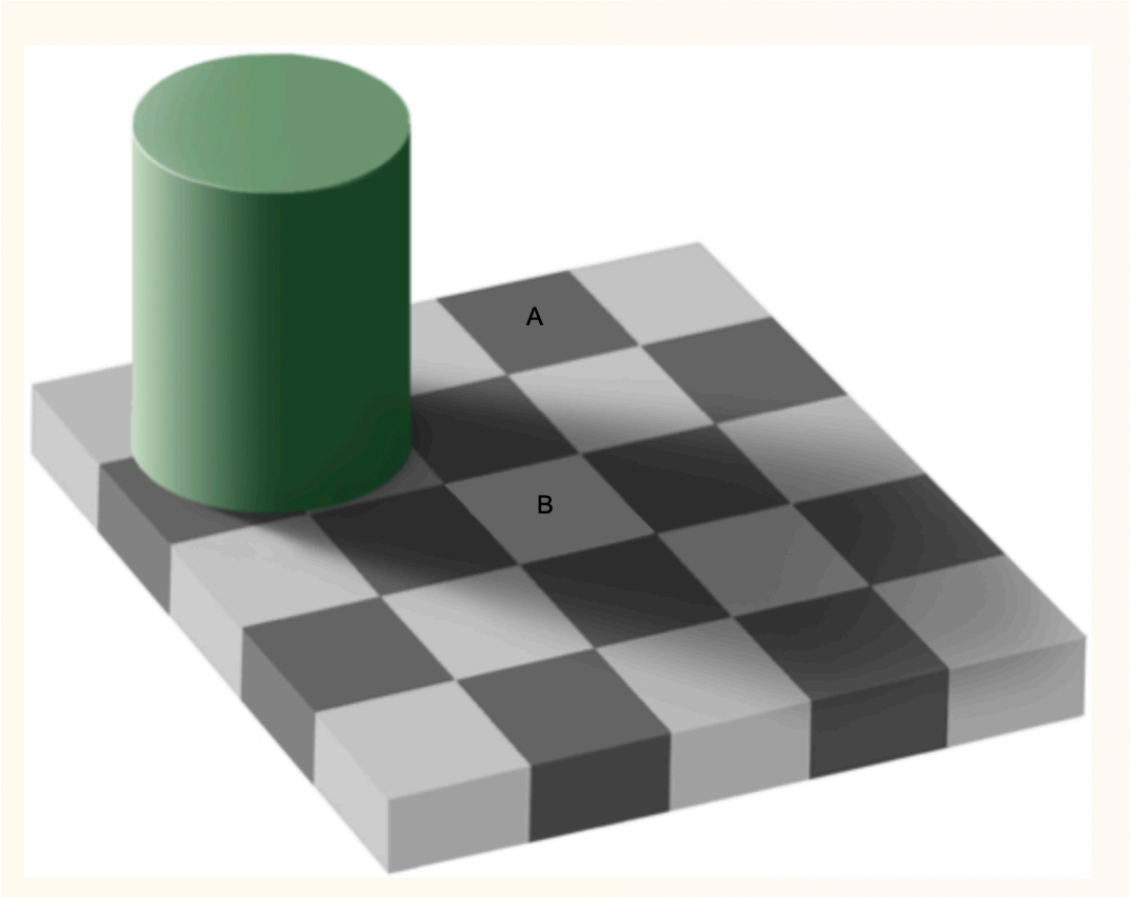
Some things know that you have eyes

Source: Antonio Torralba slides

Some things know that you have eyes

Source: Antonio Torralba slides

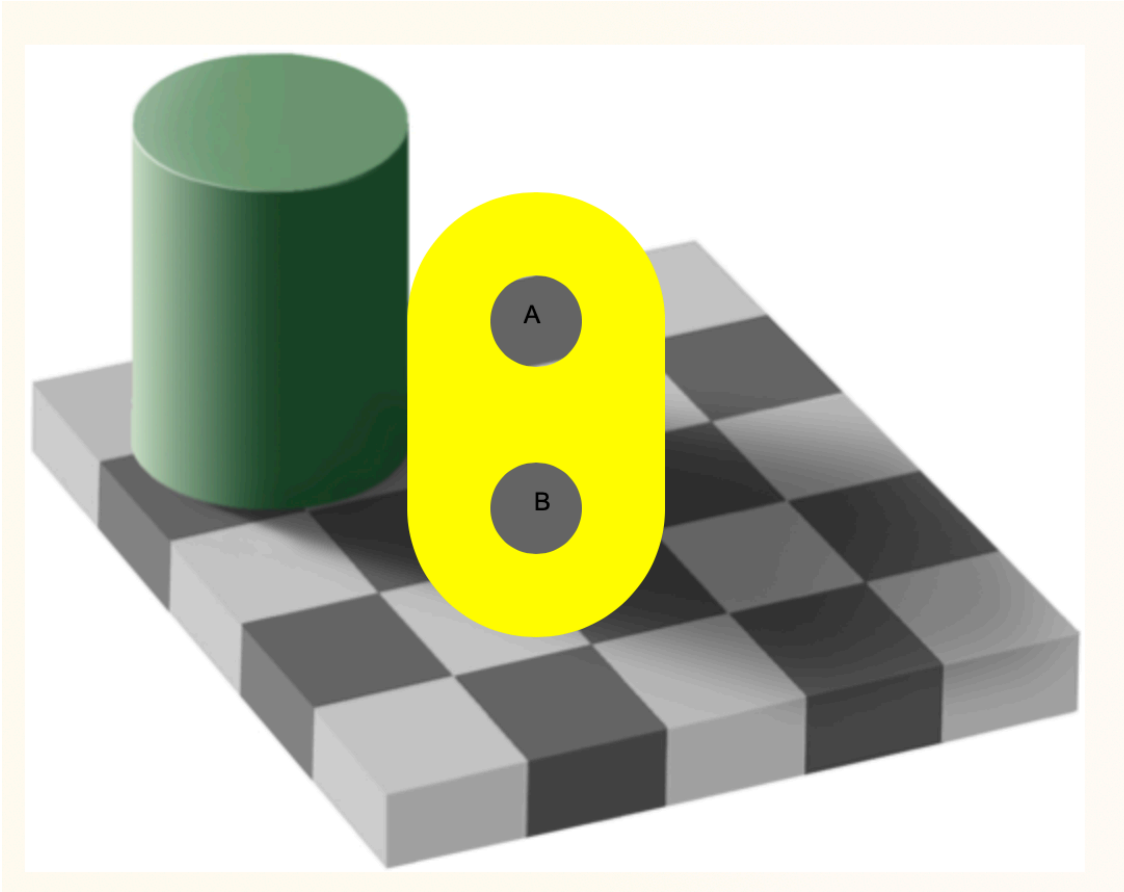
Measure light or perceive nature



http://www.michaelbach.de/ot/lum_adelsonCheckShadow/index.html

Adelson checker illusion: http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

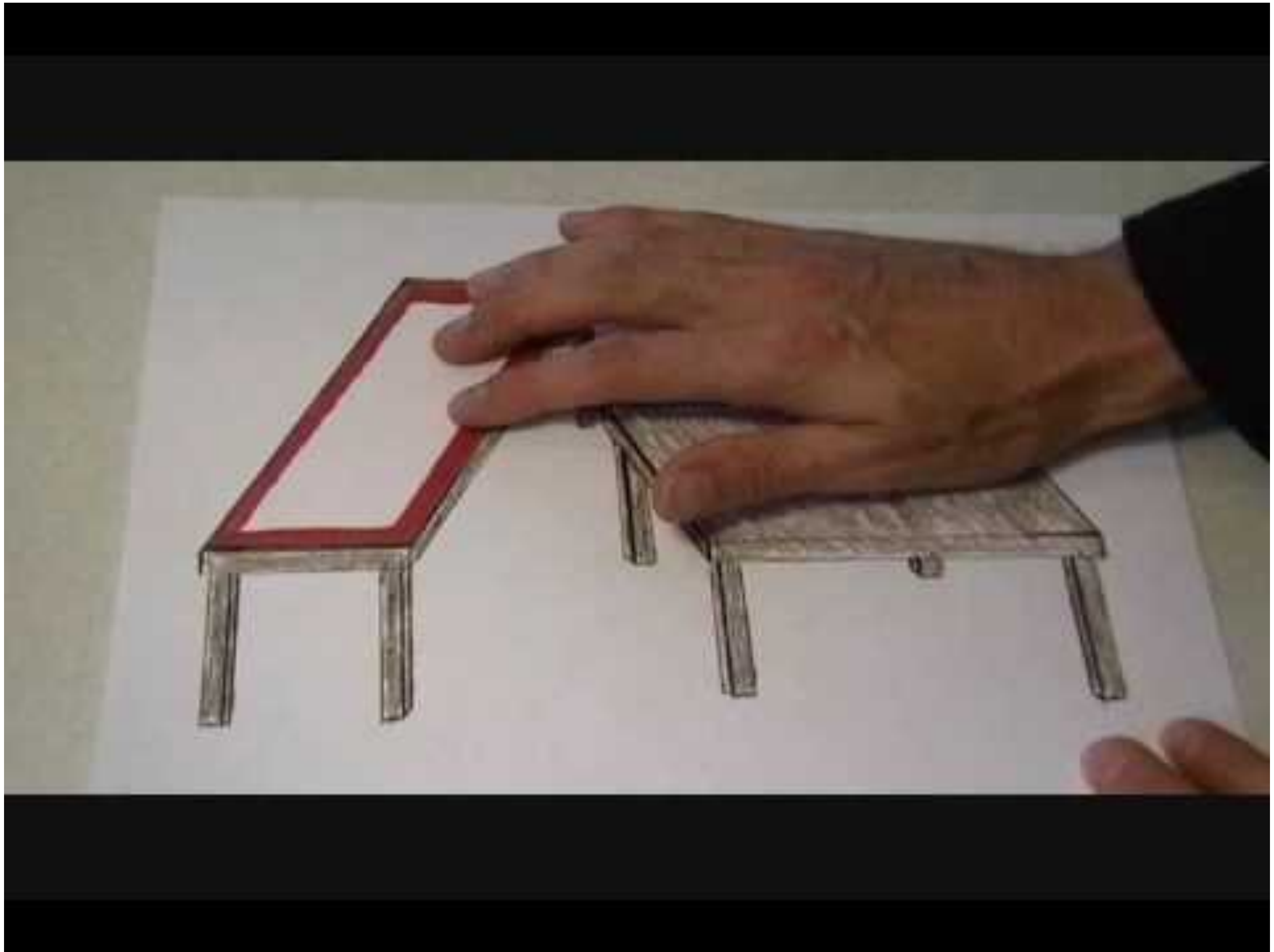
Measure light or perceive nature



http://www.michaelbach.de/ot/lum_adelsonCheckShadow/index.html

Adelson checker illusion: http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

Measure light or perceive nature

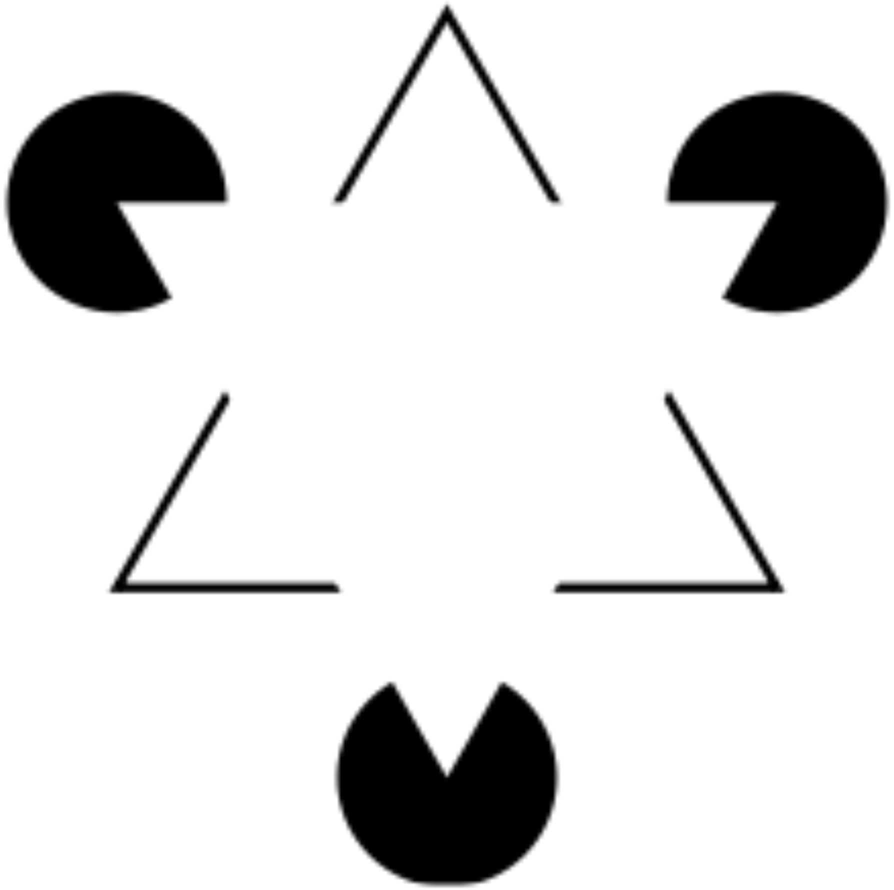


Turning tables (Roger Shepard)

The real world is in 3D world, not 2D

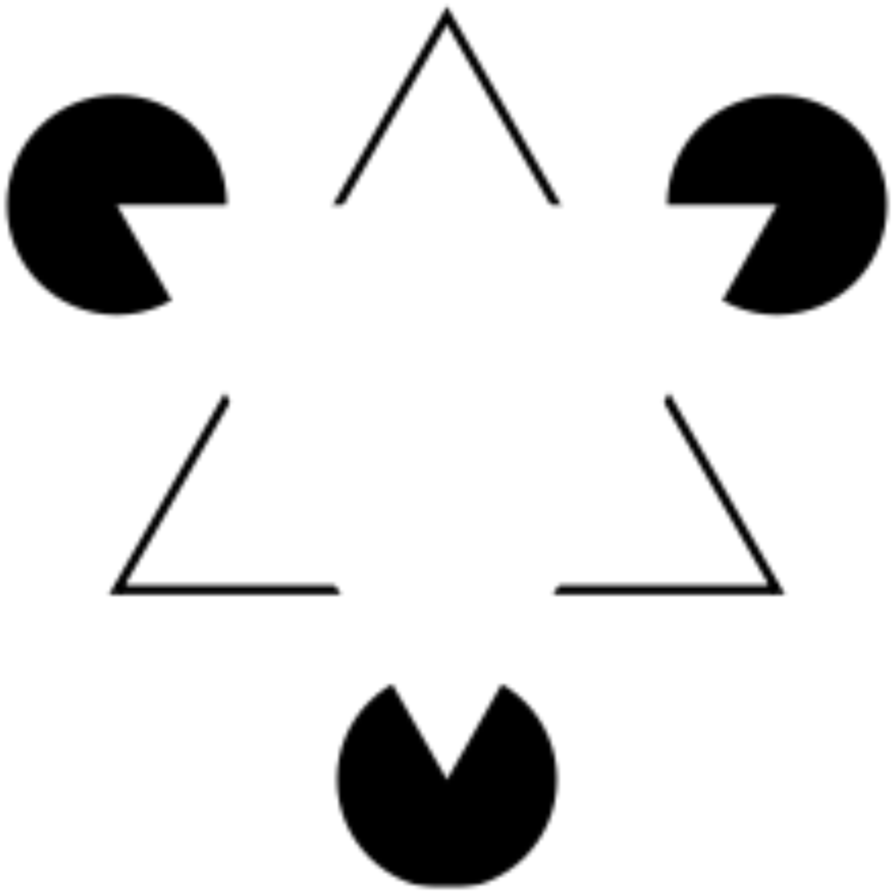
https://www.youtube.com/watch?v=PhRbsaSMMVU

Measure light or perceive nature

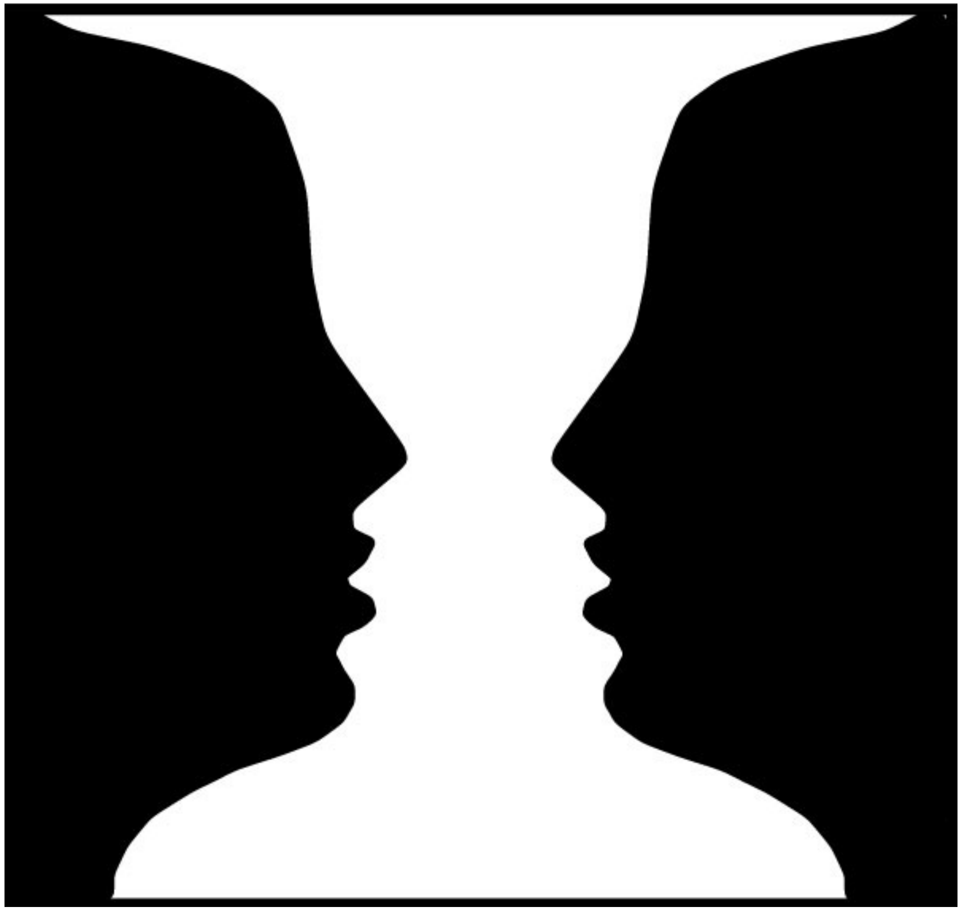


Kanizsa triangle

Measure light or perceive nature



Kanizsa triangle



Rubin’s vase

Sometimes we get it wrong

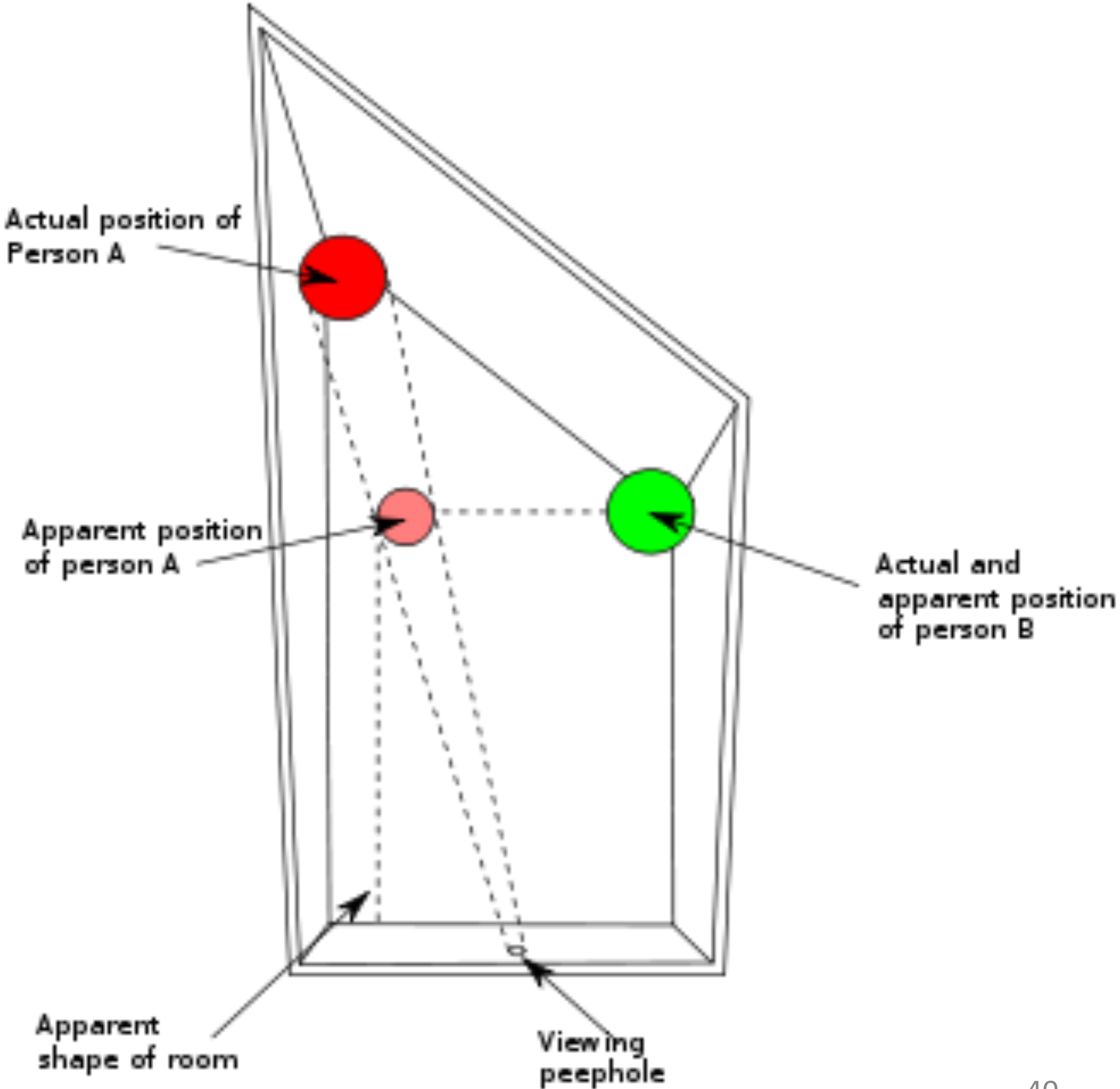
THE AMES ROOM

In the Ames room illusion, two people standing in a room appear to be of dramatically different sizes, even though they are the same size.

Sometimes we get it wrong

THE AMES ROOM

In the Ames room illusion, two people standing in a room appear to be of dramatically different sizes, even though they are the same size.

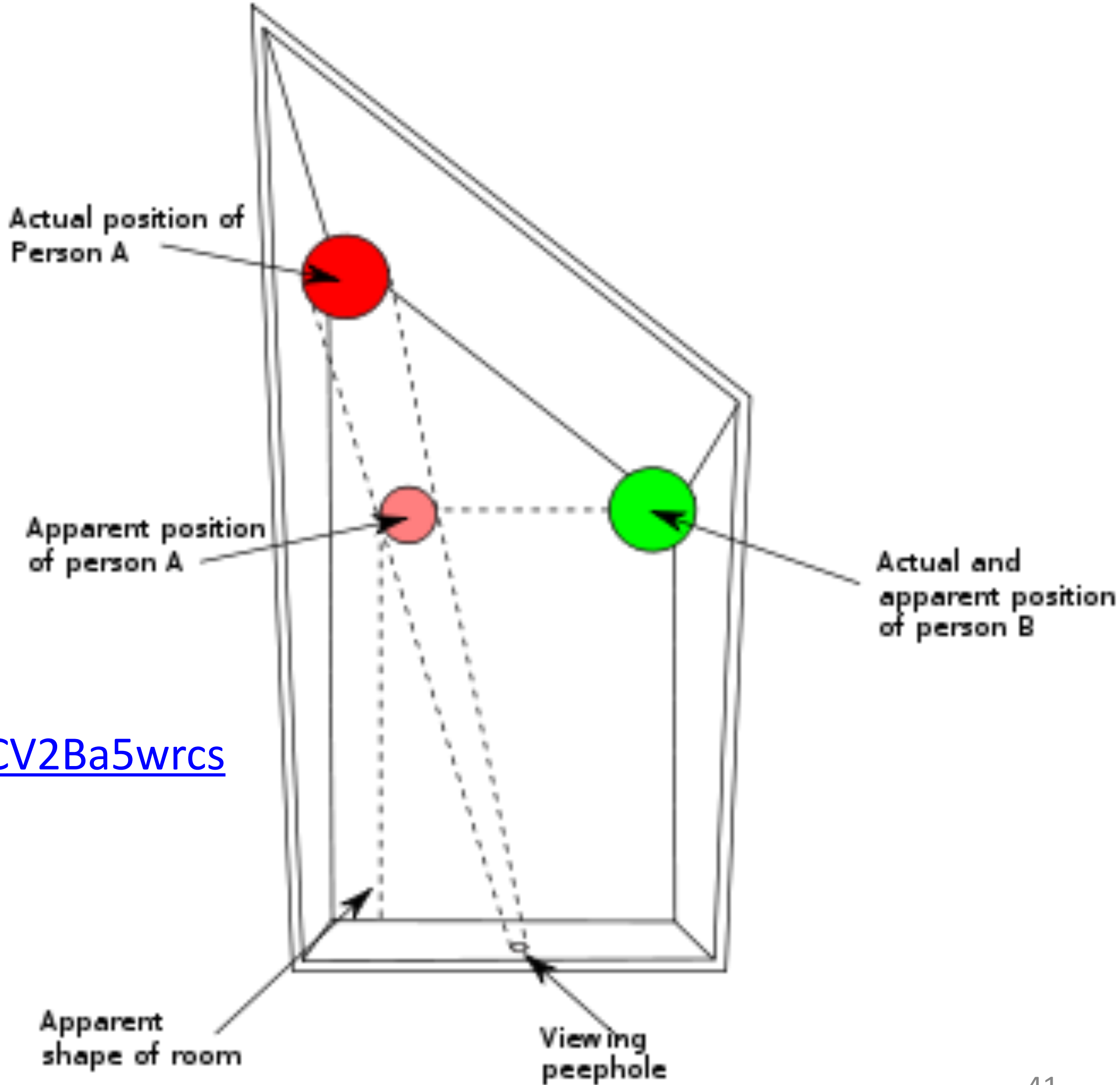


Sometimes we get it wrong

THE AMES ROOM

In the Ames room illusion, two people standing in a room appear to be of dramatically different sizes, even though they are the same size.

http://www.youtube.com/watch?v=hCV2Ba5wrcs

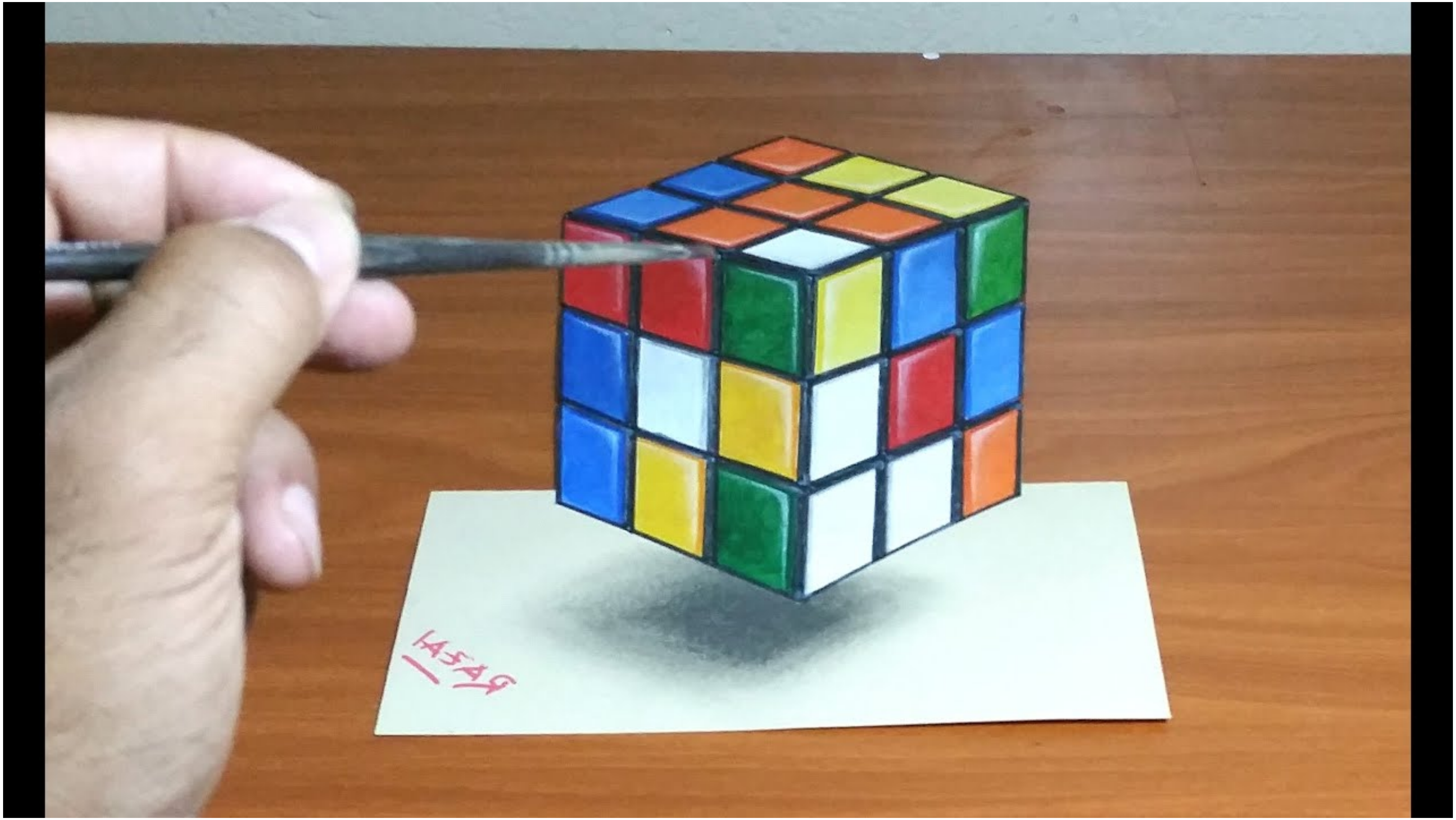


Sometimes we get it wrong

If you liked those illusions

- Check
  http://www.michaelbach.de/ot/
- Google “optical illusions”
- There are some in youtube as well:
  https://www.youtube.com/watch?v=hrhGTR54E5k
- And a competition on illusions !!
  http://illusionoftheyear.com

If you liked those illusions



Why is vision hard?

Vision has to solve an ill-posed problem

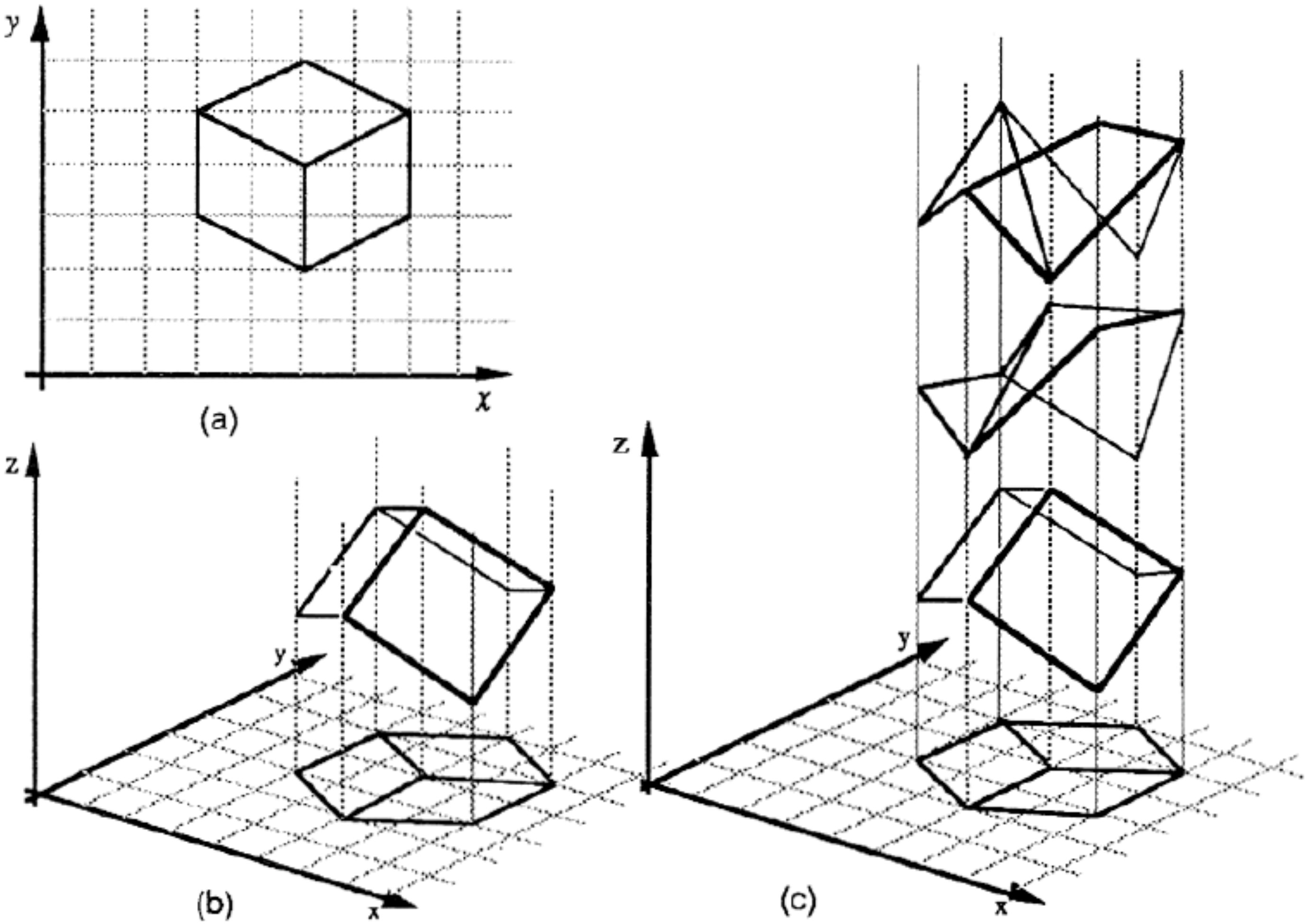
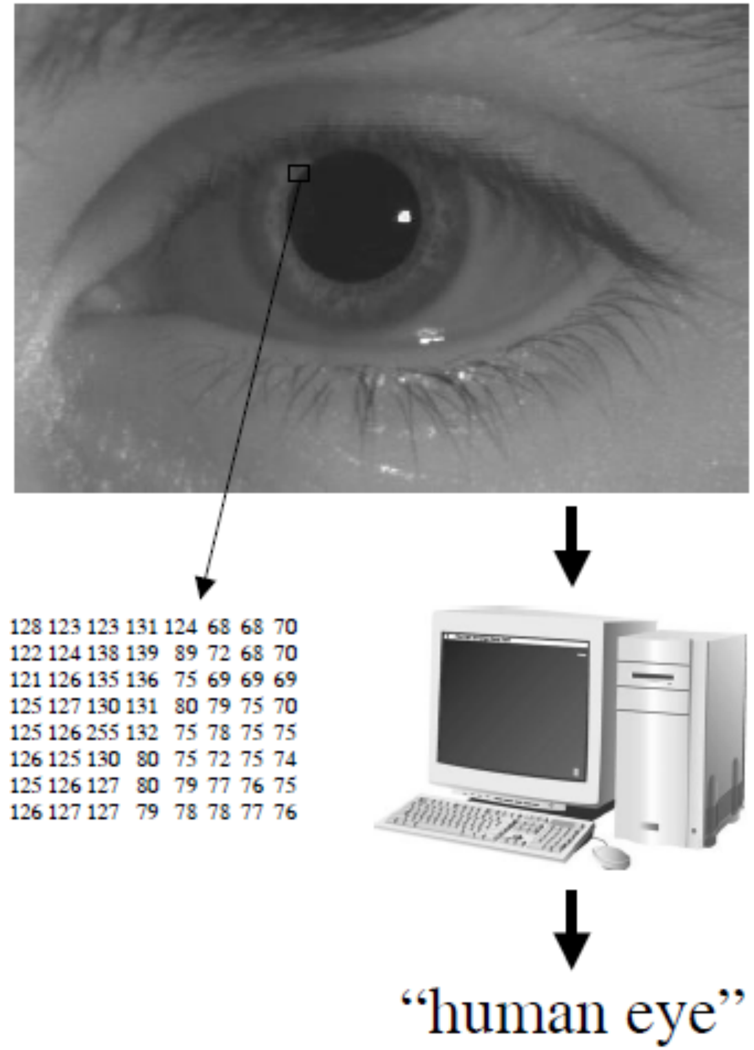


Figure 1. (a) A line drawing provides information only about the x, y coordinates of points lying along the object contours. (b) The human visual system is usually able to reconstruct an object in three dimensions given only a single 2D projection (c) Any planar line-drawing is geometrically consistent with infinitely many 3D structures.

Sihna and Adelson 1993

Why is computer vision hard?

- Desired information is not explicit in sensed data
  - Image data is a complex function of
    - Scene geometry
    - Scene surface microstructure
    - Scene illumination
  - Biological systems provide an existing proof
    - But appear to be complex
    - Are only partially understood



Source: Richard Wildes slides

Why is computer vision hard?

- Problems of computer vision
  - Computing properties of the world from one or more images
  - Properties of interest
    - Geometric: shape, position
    - Photometric: surface reflectance
    - Dynamic: velocity
- Tools of computer vision
  - Computers to interpret images
  - Specific hardware and software to acquire, store, process and communicate images
  - Algorithms for computing the desired information from available input

Source: Richard Wildes slides

Outline

- Why is vision hard?
- Applications
- Allied areas
- History

Some applications

- Face detection

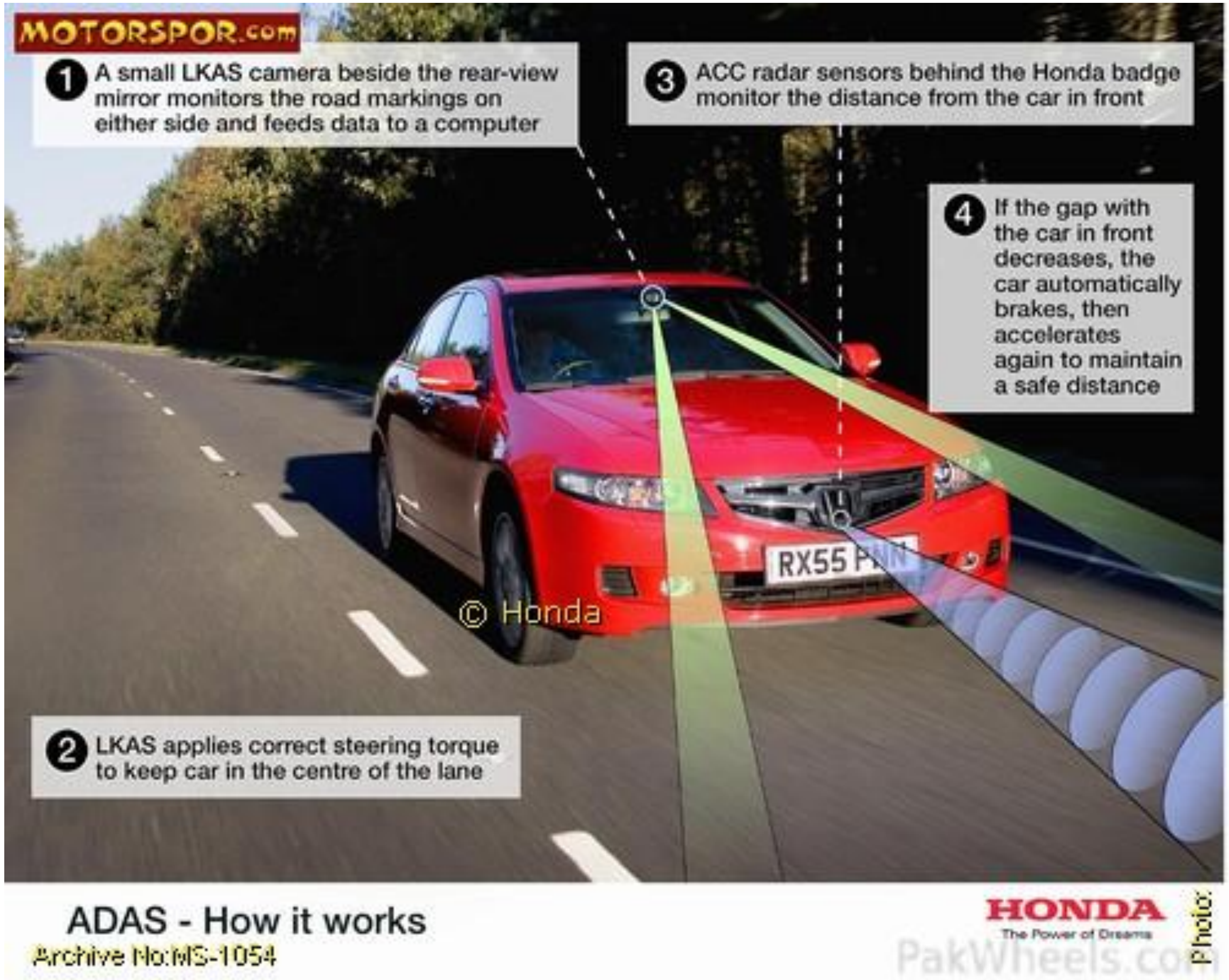
Some applications

- Face detection
  - Viola and Jones. ICCV 2001 (In OpenCV)

- It can process a 384x288 pixel image in 0.67 seconds (on a PIII @ 700MHz)

Some applications

- Assisted driving

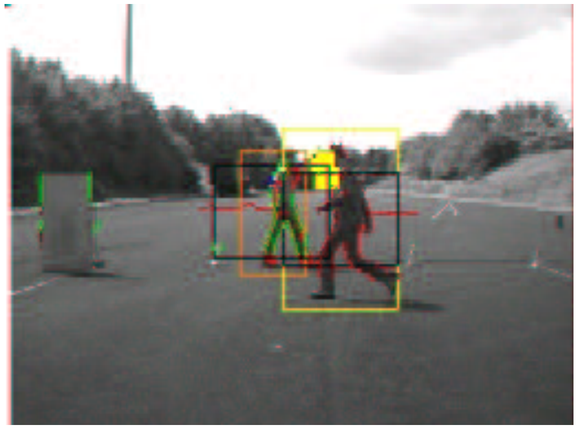


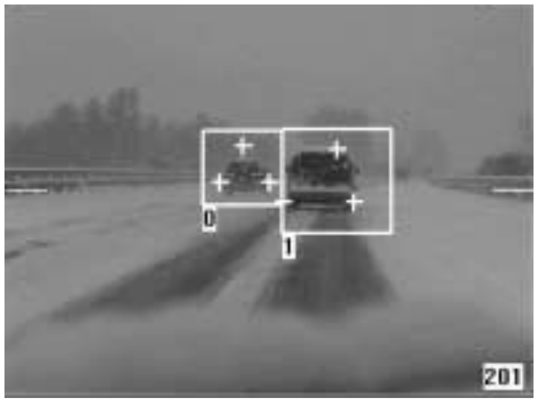
Some applications

- Assisted driving

Some applications

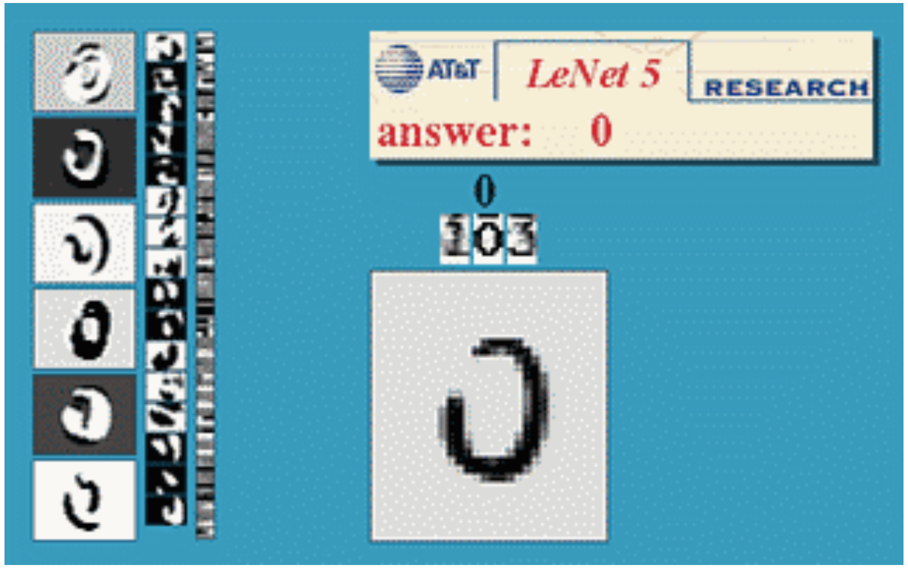
- Assisted driving
  - Pedestrian detection
    - Sotelo, Parra, Fernández & Naranjo 2006
  - Lane detection
    - Zheng, Tang, Chend, Li, Lai & Wang, 2004
  - Car detection in RT
    - Betke, Haritaoglu & Davis, 2000



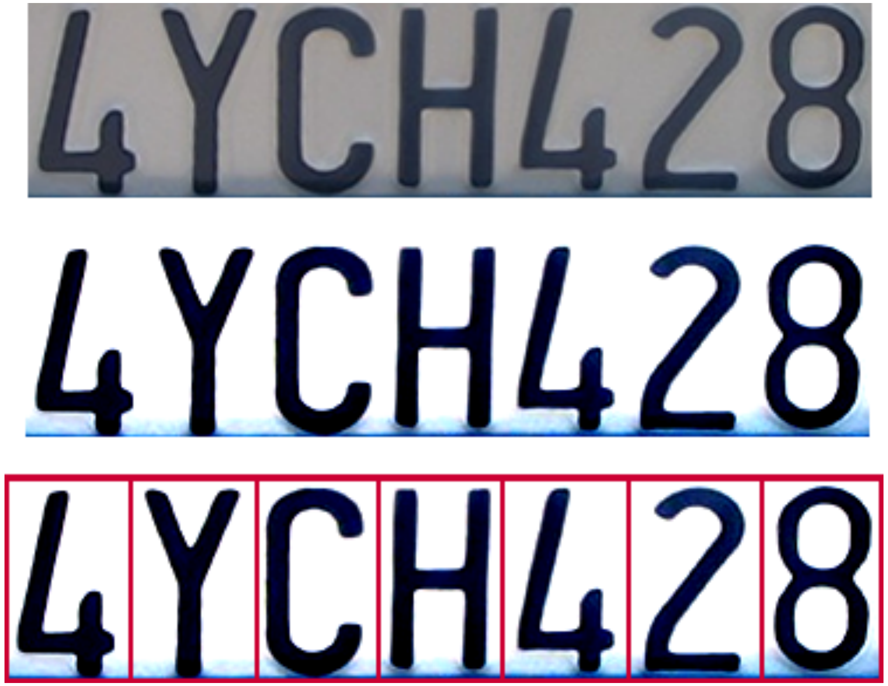


Some applications

- Reading license plates, zip codes



Digit recognition, AT&T labs

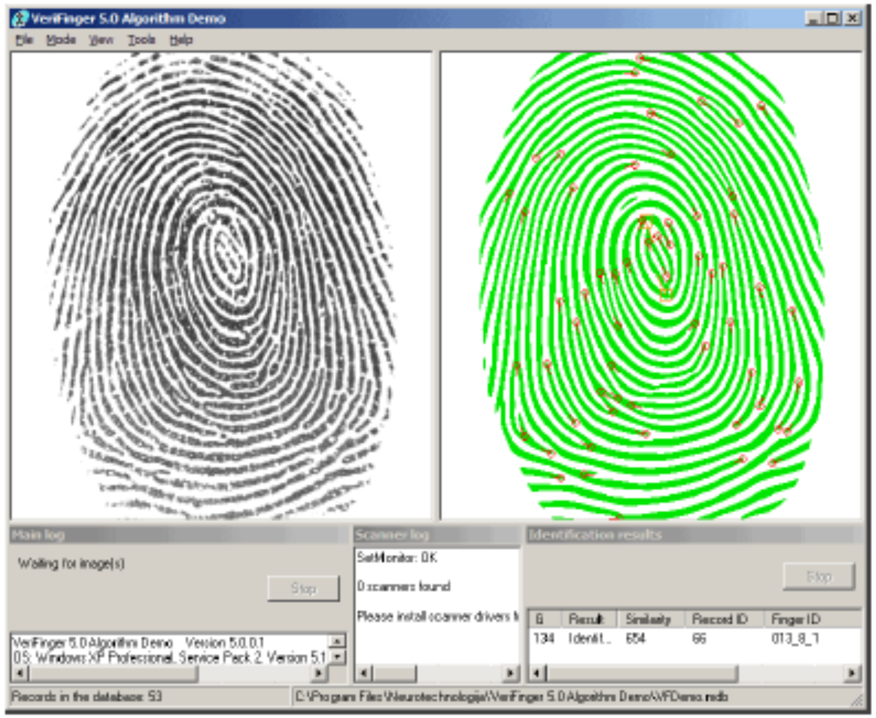


License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Source: S. Seitz

Some applications

- Reading license plates, zip codes
- Fingerprint recognition

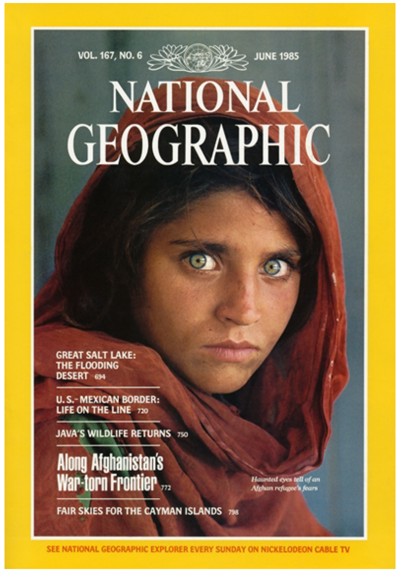


Some applications

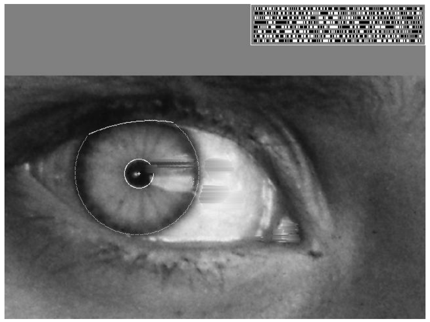
- Iris recognition

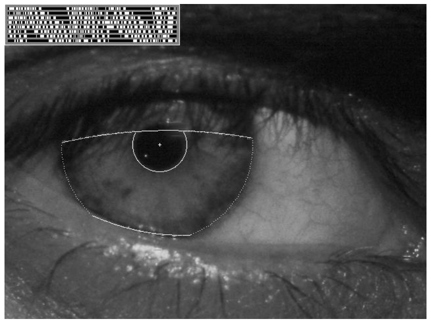
Some applications

- Iris recognition



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





Source: S. Seitz

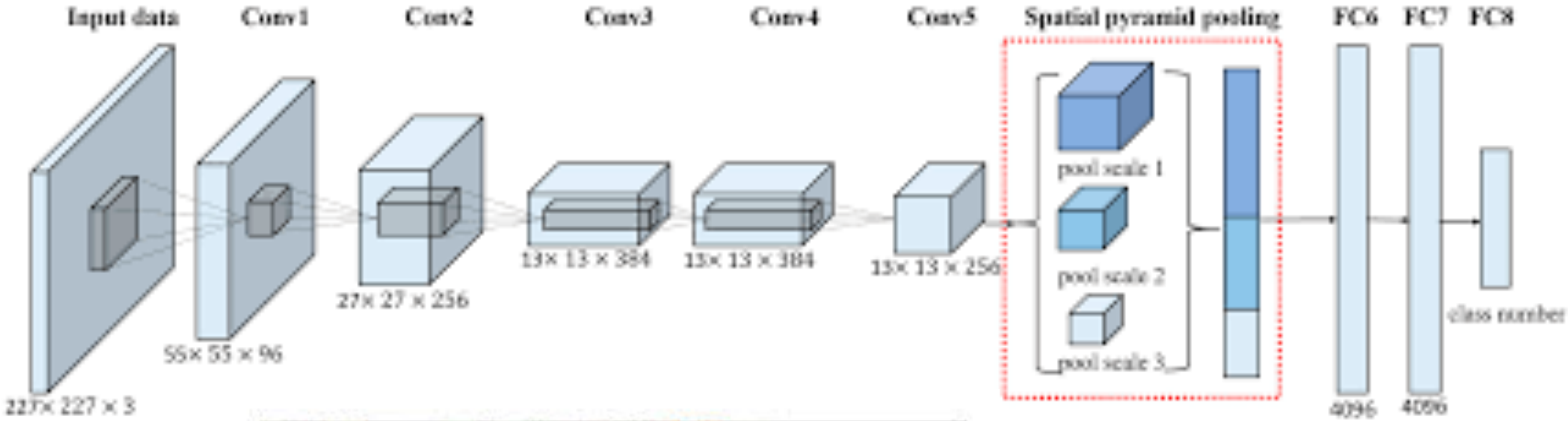
Some applications

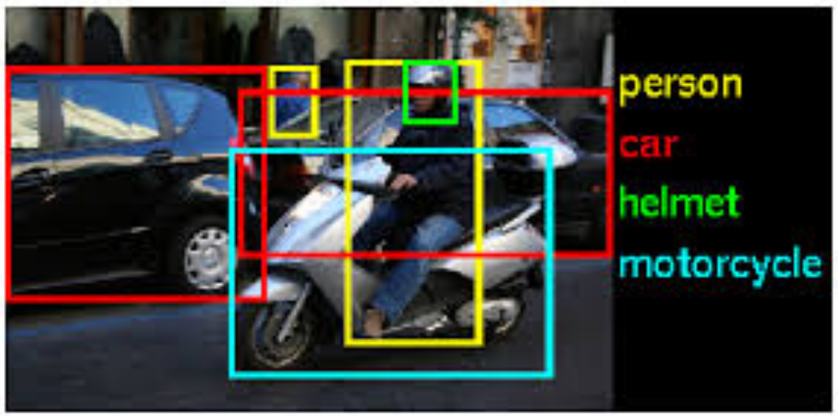
- Object recognition
  - David Lowe's SIFT

Figure 12: The training images for two objects are shown on the left. These can be recognized in a cluttered image with extensive occlusion, shown in the middle. The results of recognition are shown on the right. A parallelogram is drawn around each recognized object showing the boundaries of the original training image under the affine transformation solved for during recognition. Smaller squares indicate the keypoints that were used for recognition.

Some applications

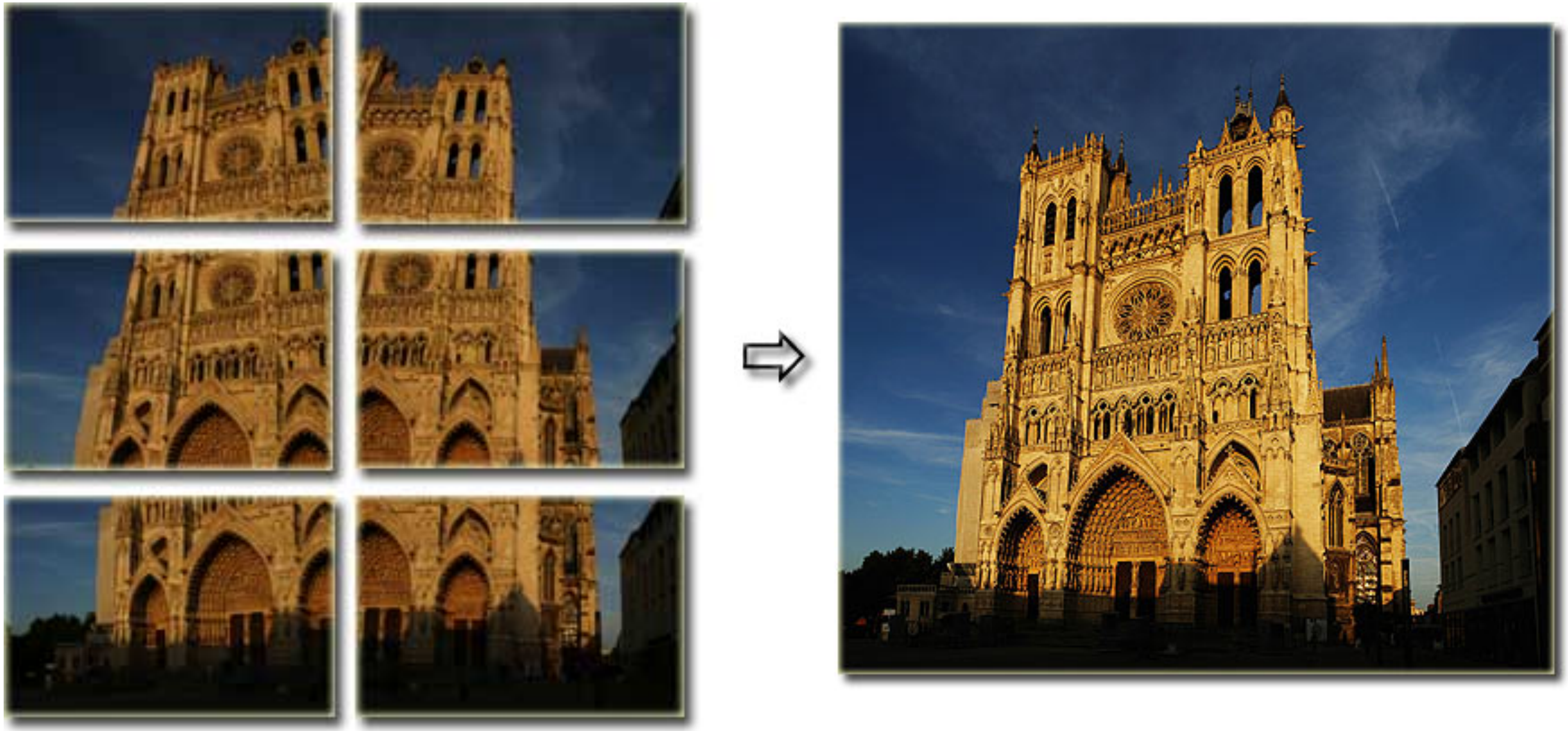
- Object recognition
  - GoogleNet





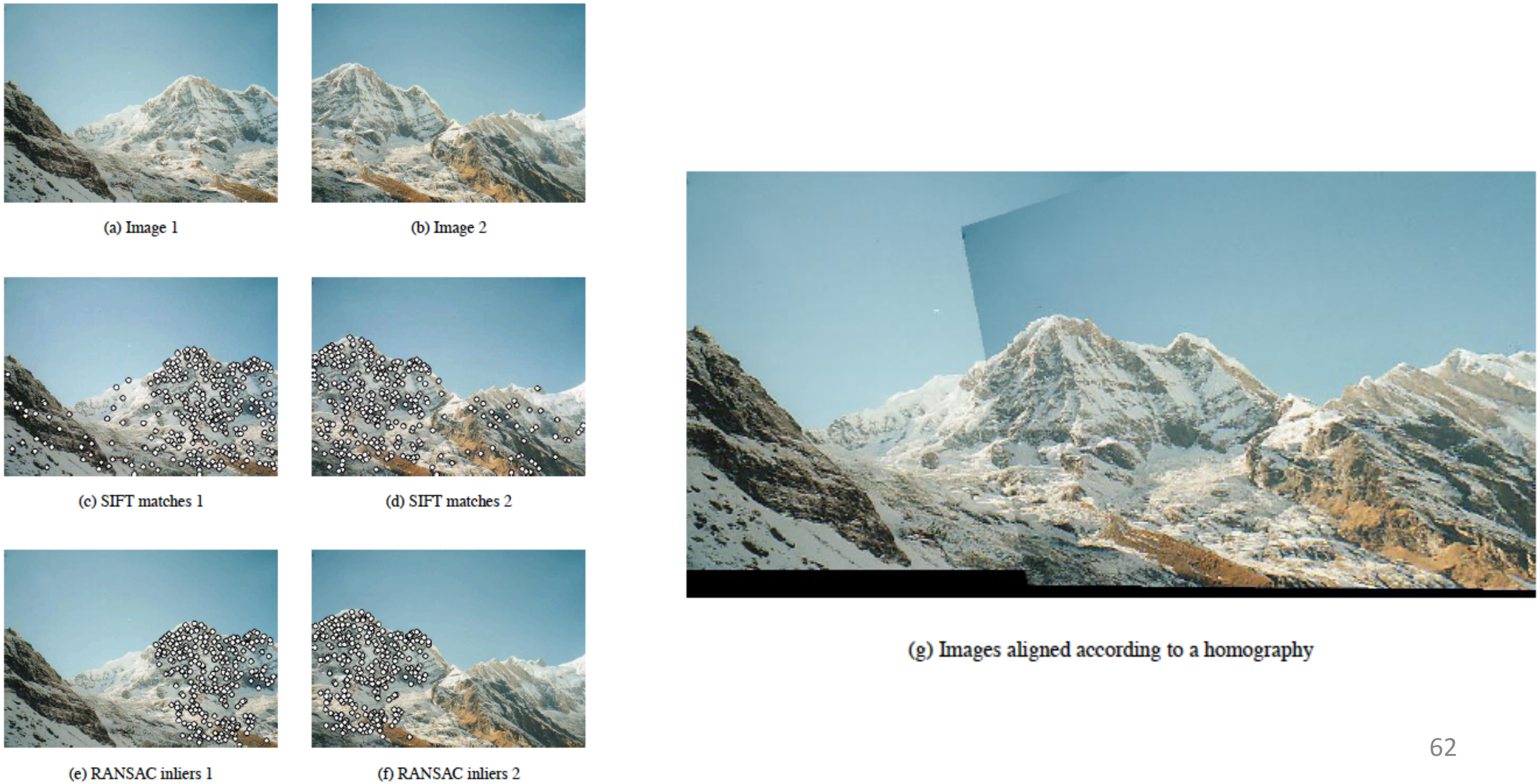
Some applications

- Image stitching



Some applications

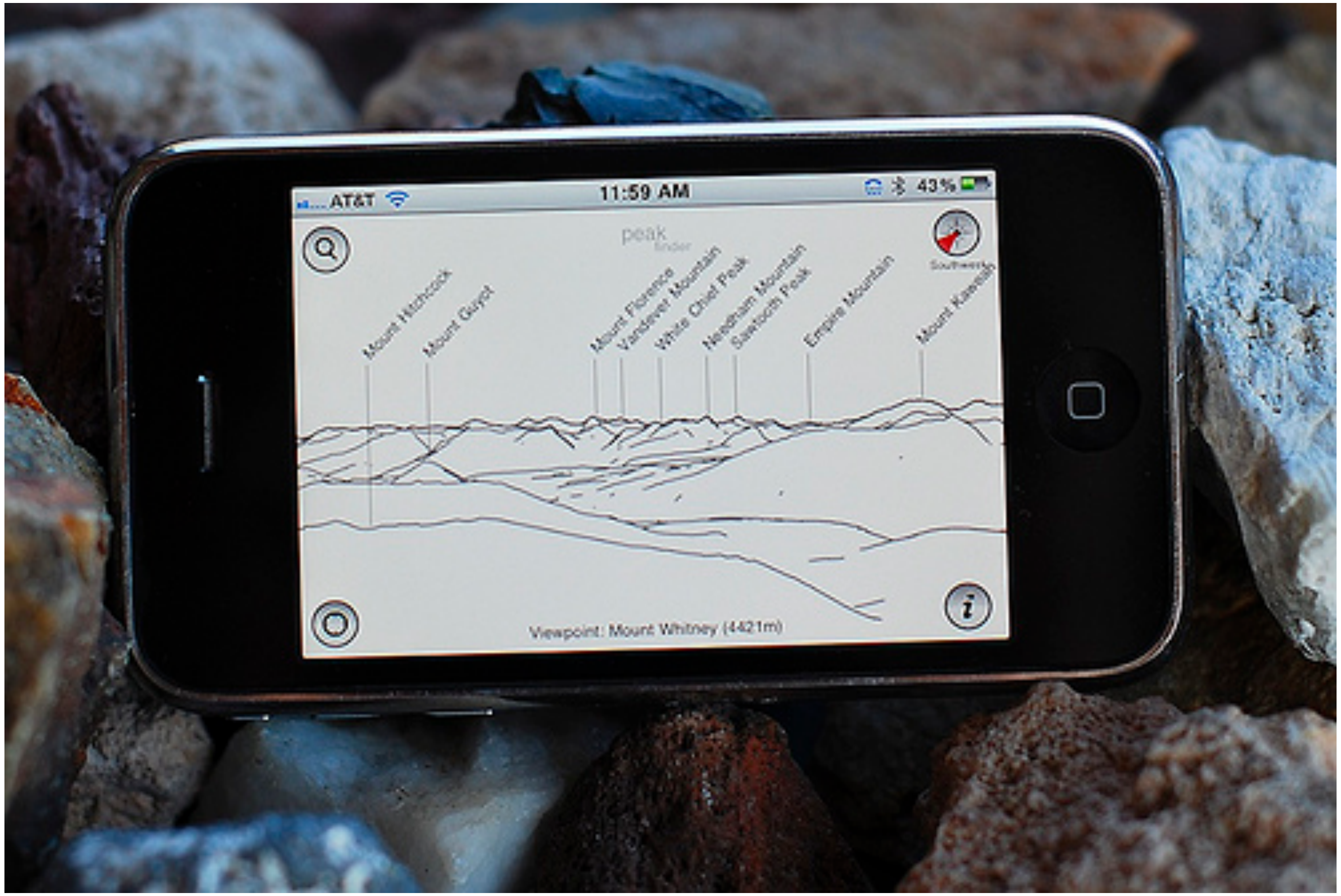
- Image stitching with SIFT (Brown and Lowe, 2007) and RANSAC (Fischler and Boiles, 1981)



(a) Image 1
(b) Image 2
(c) SIFT matches 1
(d) SIFT matches 2
(e) RANSAC inliers 1
(f) RANSAC inliers 2
(g) Images aligned according to a homography

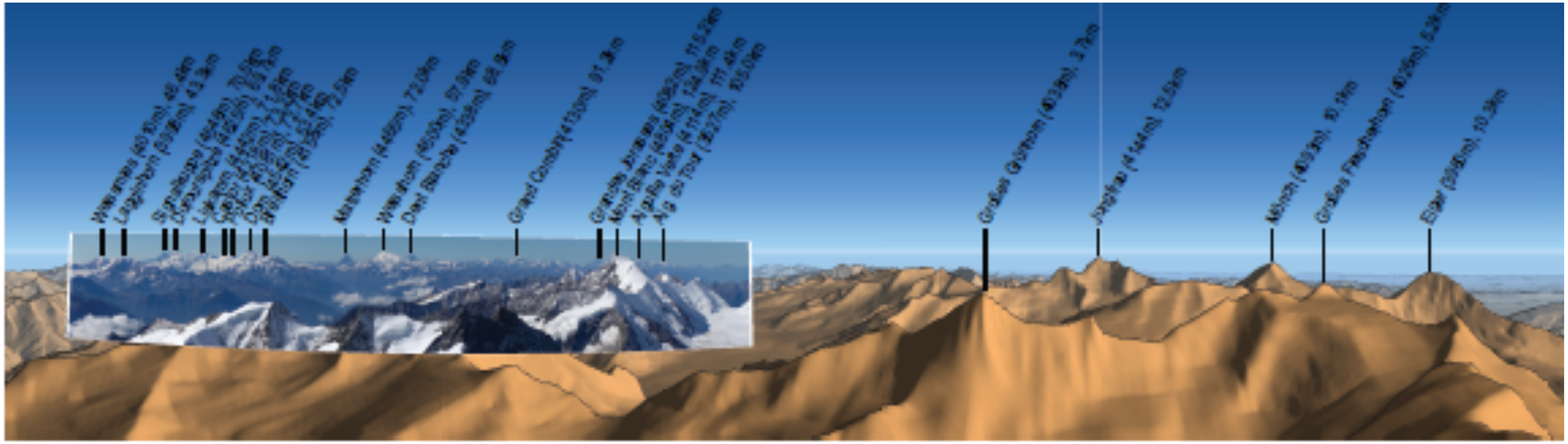
Some applications

- Object labeling (e.g. mountains)



Some applications

- Object labeling (e.g. mountains)
  – Baboud, Cadik, Eisemann & Seidel, 2011



Some applications

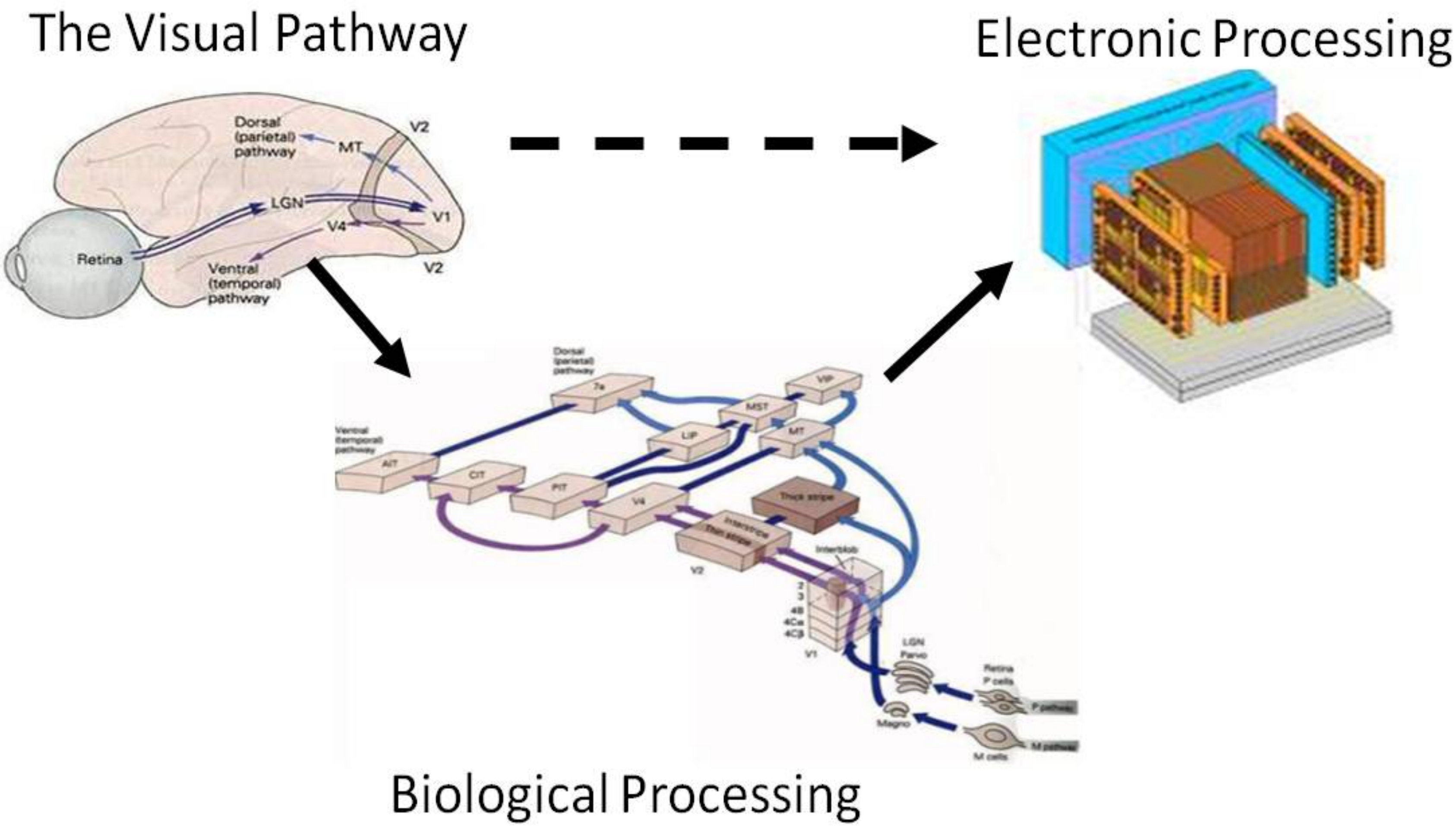
- Finding images in large collections
  - searching for pictures
  - browsing collections of pictures
  - E.g. Internet or photo collections
- Image based rendering
  - often very difficult to produce models that look like real objects
    - surface weathering, etc., create details that are hard to model
    - Solution: make new pictures from old

Some applications

- Surveillance
  - Warn me if there is a mugging in the grove
- HCI
  - Do what I show you
  - E.g. Intellact, http://www.intellact.eu/
- Military
  - Shoot this, not that
  - E.g. Neovision2

Some applications

- Darpa's Neovision2

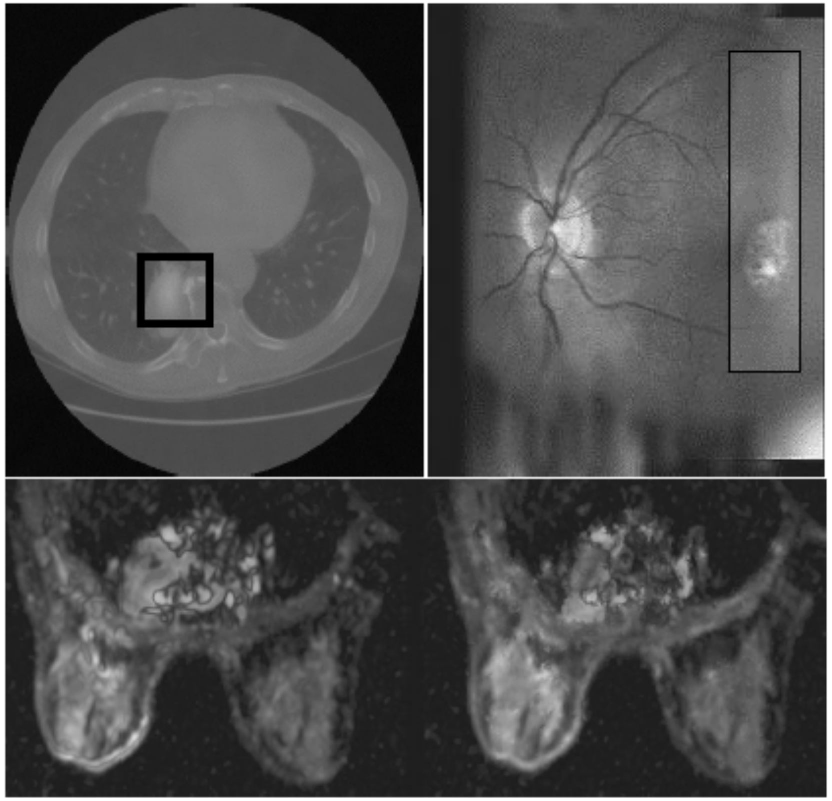


Some applications

- Analysis of aerial images and underwater images
  - E.g. Prof. Dudek's work: http://www.youtube.com/watch?v=fax1hERGMe4
- Exploring space
  - Advanced Space Vision System in the International Space Station
  - Mars Rover
  - Mars Perseverance

Some applications

- Medical image analysis
  – Aides to physicians in diagnosis of disease



Source: Richard Wildes slides

Outline

- Why is vision hard?
- Applications
- Allied areas
- History

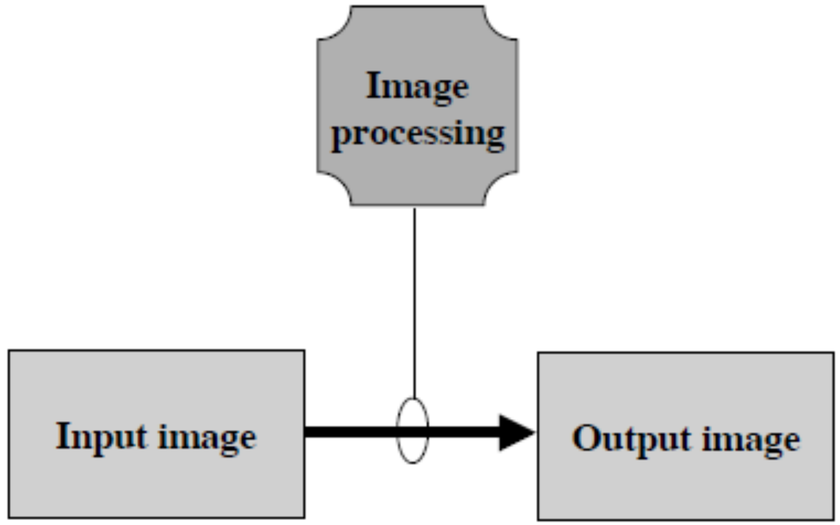
Areas

- Image feature detection
- Color analysis
- Texture analysis
- Camera calibration
- Shape representation
- Shape from X (Shape recovery from image cues)
- Stereopsis
- Motion analysis
- Object recognition and localization
- Active vision
- High-performance and real-time architectures

Source: Richard Wildes slides

Allied areas: Image processing

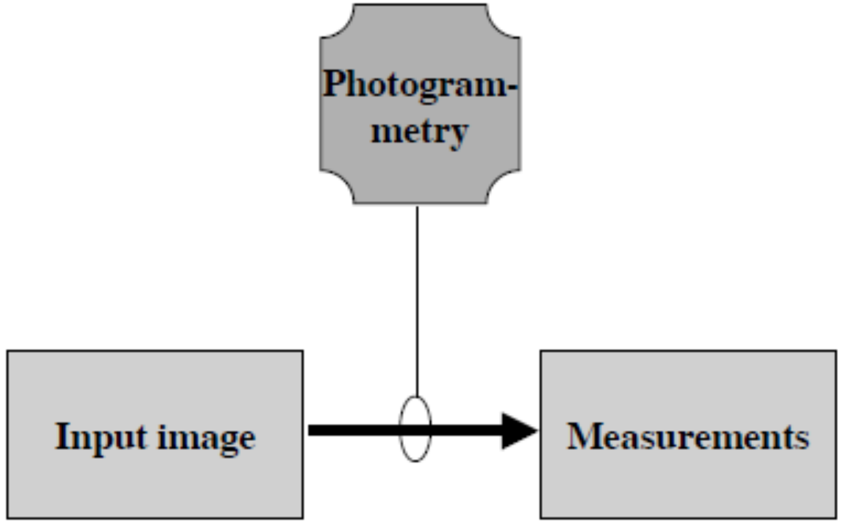
- Generation of new images from existing images
  - Noise suppression
  - Feature enhancement
  - Video stabilization
- Relationship to computer vision
  - Often serves to provide components to computer vision
  - Preprocessing of sensed data



Source: Richard Wildes slides

Allied areas: Photogrammetry

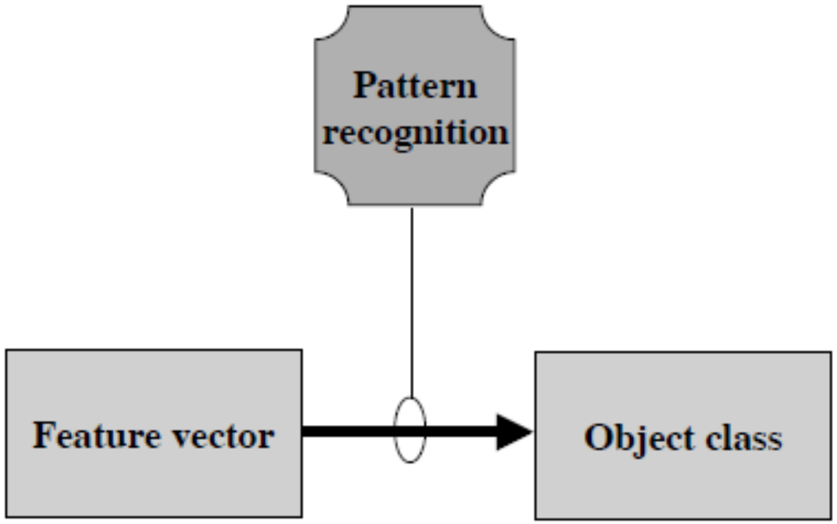
- Reliably accurate measurements
  - Aerial cartography
  - Surveying
- Relationship to computer vision
  - Seeks higher levels of precision than computer vision



Source: Richard Wildes slides

Allied areas: Pattern Recognition

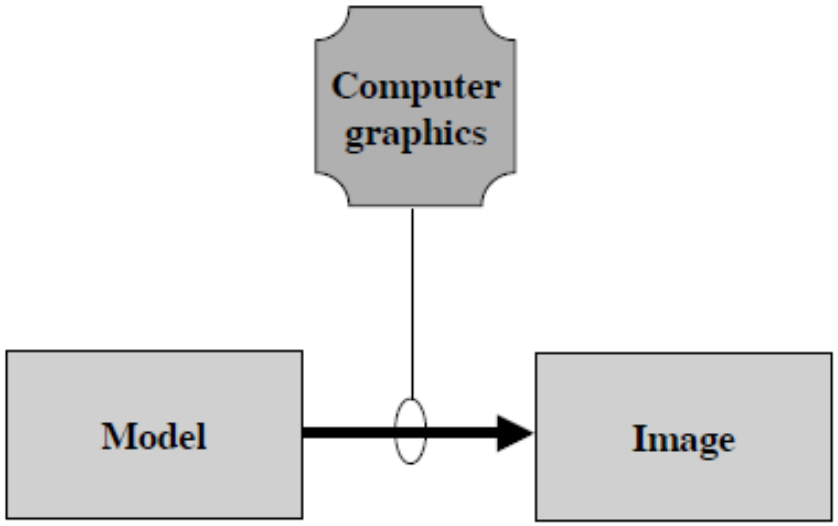
- Classification of patterns
  - Classification of chemical compositions
  - Disease from symptoms
  - Targets from visual features
- Relationship to computer vision
  - Can usefully be applied to the output of a computer vision system
  - Capable of assigning imaged objects to classes



Source: Richard Wildes slides

Allied areas: Computer graphics

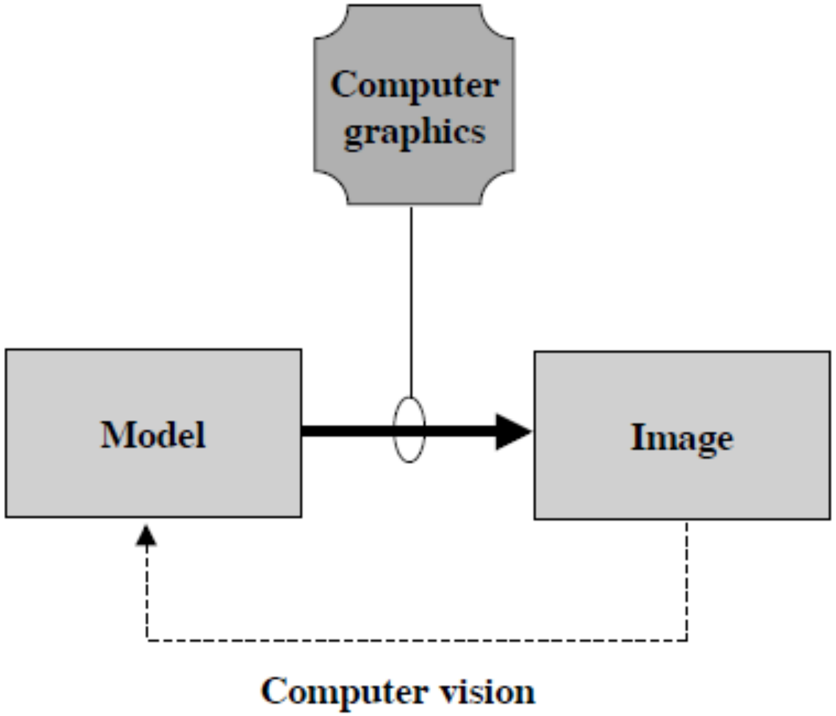
- Generation of images from models
  - Photorealistic rendering
  - Computer animation
  - Abstract design
- Relationship to computer vision
  - Akin to an inverse
  - Potential to combine forces



Source: Richard Wildes slides

Allied areas: Computer graphics

- Generation of images from models
  - Photorealistic rendering
  - Computer animation
  - Abstract design
- Relationship to computer vision
  - Akin to an inverse
  - Potential to combine forces



Source: Richard Wildes slides

Outline

- Why is vision hard?
- Applications
- Allied areas
- History

History

- Curiosity: Early Fourier analyzer

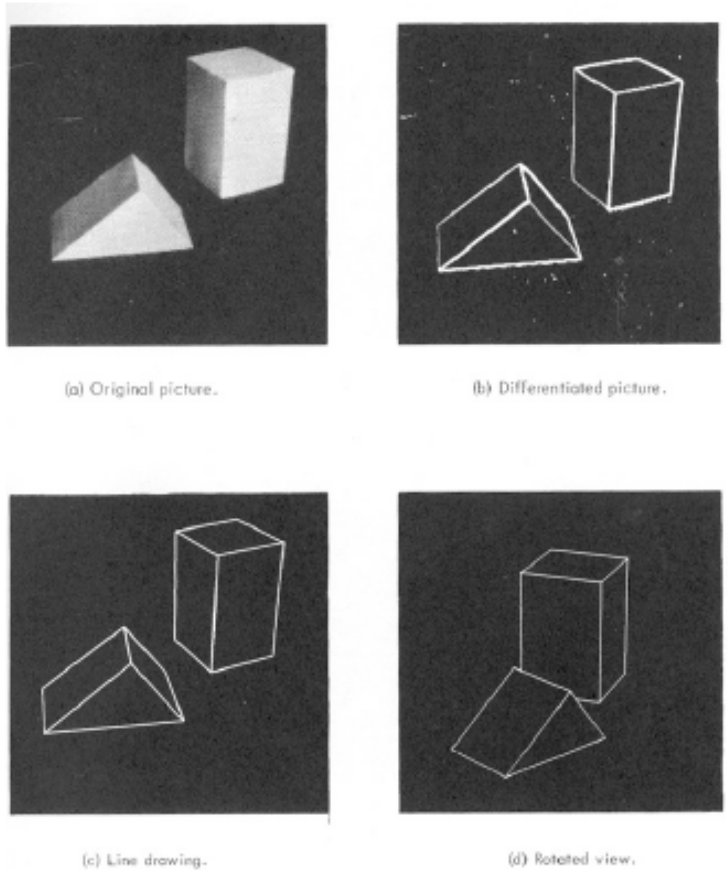
Present at the Center for Vision Research, York University

History

- There has been computer vision since 1960s
  - First it was called digital image processing
  - Appearance of journals and conferences
    - Pattern Recognition: 1968
    - ICPR (International Conference on Pattern Recognition): 1970
    - CVPR (Computer Vision and Pattern Recognition): 1978
    - IEEE TPAMI (Transactions on Pattern Analysis and Machine Intelligence): 1979
    - ICCV (International Conference on Computer Vision): 1987
    - And so on: Vision Research, Journal of Vision, etc.

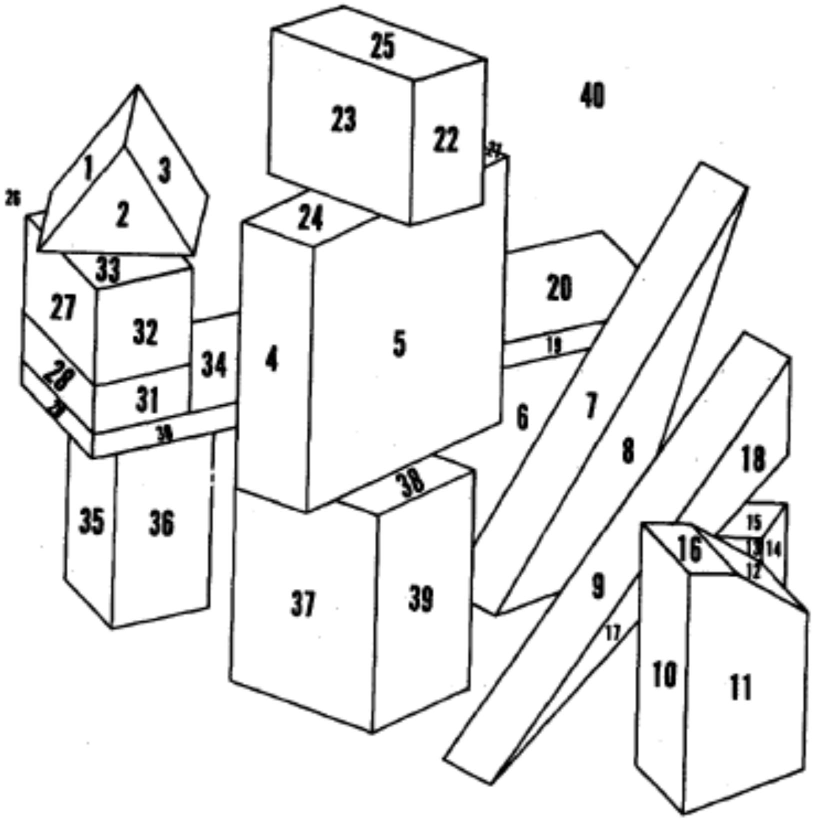
History

- There has been computer vision since 1960s
  - Early years
    - Roberts (1965): Machine perception of 3D solids (PhD thesis)
      - Line drawing of block scenes



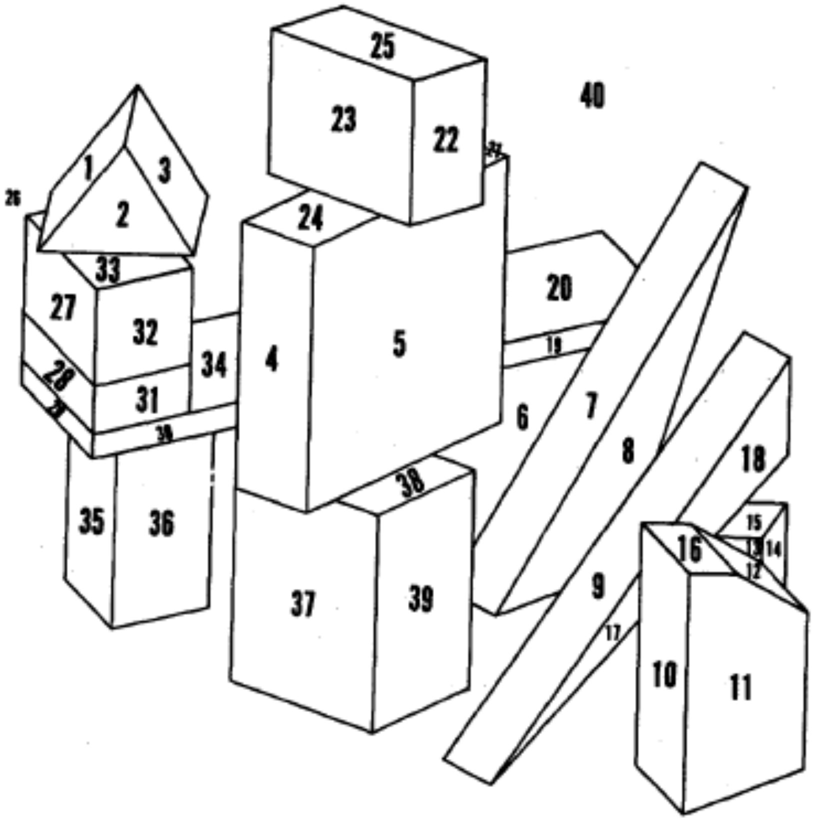
History

- There has been computer vision since 1960s
  - Early years
    - Guzmán (1967): Decomposition of a visual scene into three-dimensional bodies (PhD thesis)



History

- There has been computer vision since 1960s
  - Early years
    - Guzmán (1967): Decomposition of a visual scene into three-dimensional bodies (PhD thesis)

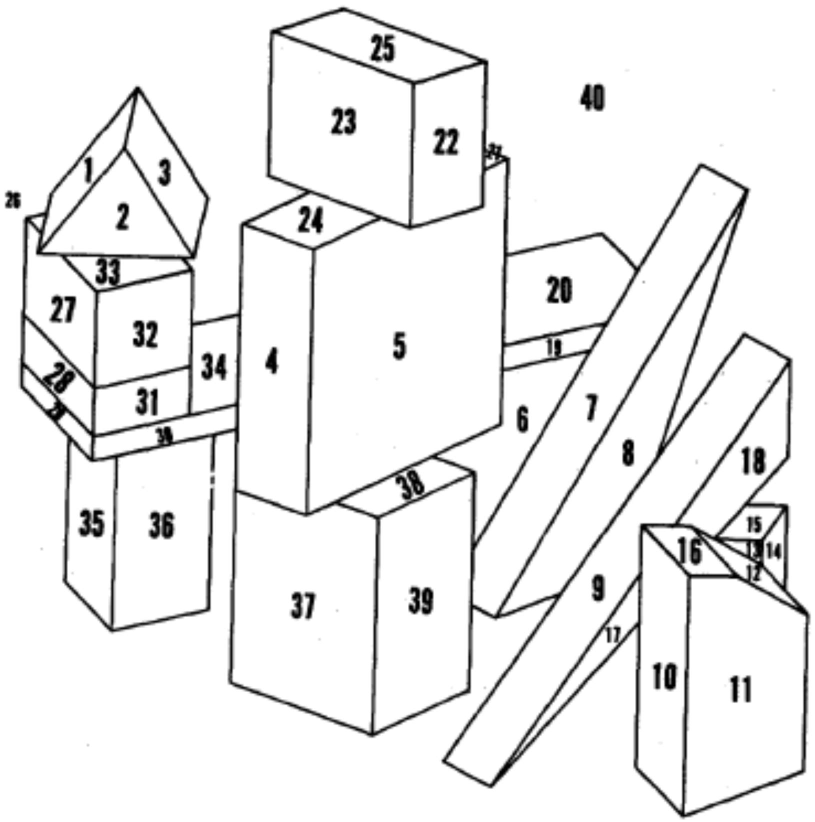


(LOCAL ASSUMES (:17) (:9) SAME BODY)
(LOCAL ASSUMES (:9 :17) (:18) SAME BODY)
(BODY 1. IS :3 :2 :1)
(BODY 2. IS :32 :33 :27 :26)
(BODY 3. IS :28 :31)
(BODY 4. IS :19 :20 :34 :30 :29)
(BODY 5. IS :36 :35)
(BODY 6. IS :24 :5 :21 :4)
(BODY 7. IS :25 :23 :22)
(BODY 8. IS :14 :13 :15)
(BODY 9. IS :10 :16 :11 :12)
(BODY 10. IS :18 :9 :17)
(BODY 11 IS :7 :8)
(BODY 12. IS :38 :37 :39)

Results for 'MOMO'

History

- There has been computer vision since 1960s
  - Early years
    - Guzmán (1967): Decomposition of a visual scene into three-dimensional bodies (PhD thesis)



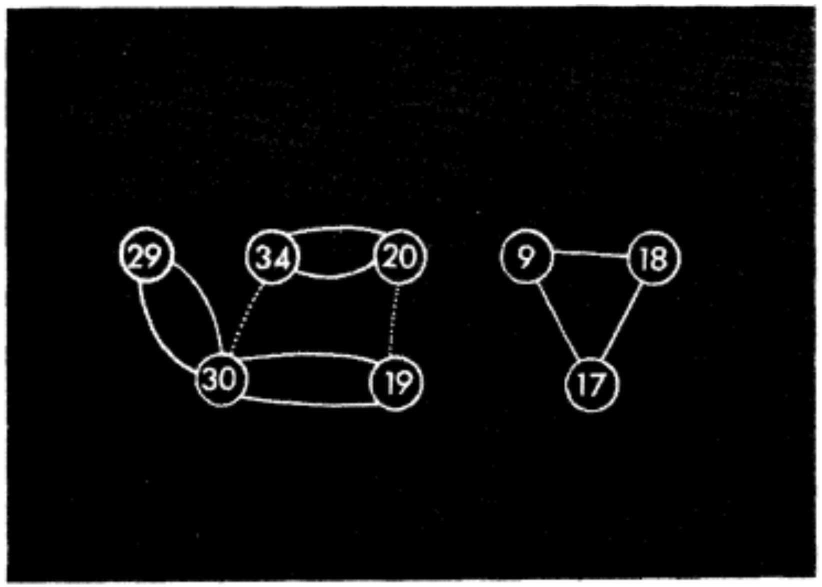


FIGURE 16—MOMO—LINKS'. Some links of figure 15 'MOMO'

History

- There has been computer vision since 1960s
  - Early years
    - Blum (1967): Transformation for extracting descriptors of shape
      - MAF: Medial Axis Function

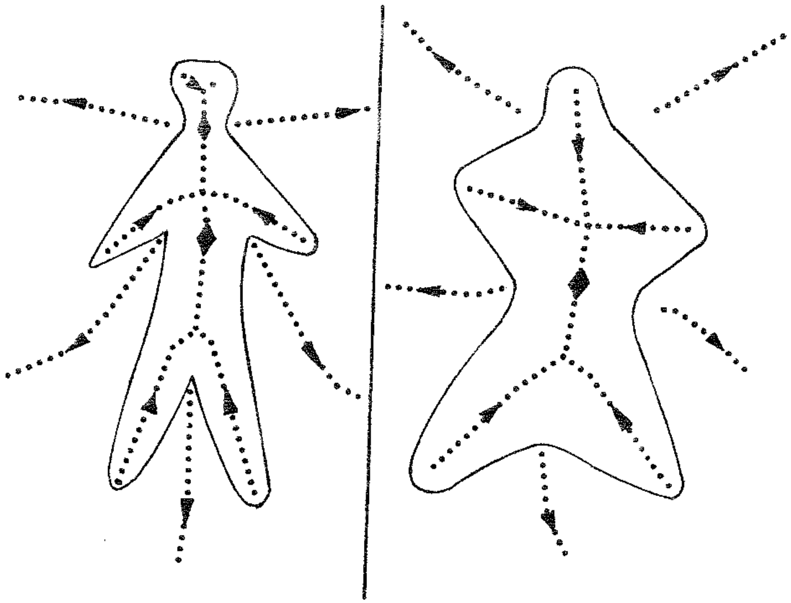


Fig. 6. Two anthropomorphs and their MAFs.

History

- There has been computer vision since 1960s
  - Early years
    - Blum (1967): Transformation for extracting descriptors of shape
      - MAF: Medial Axis Function

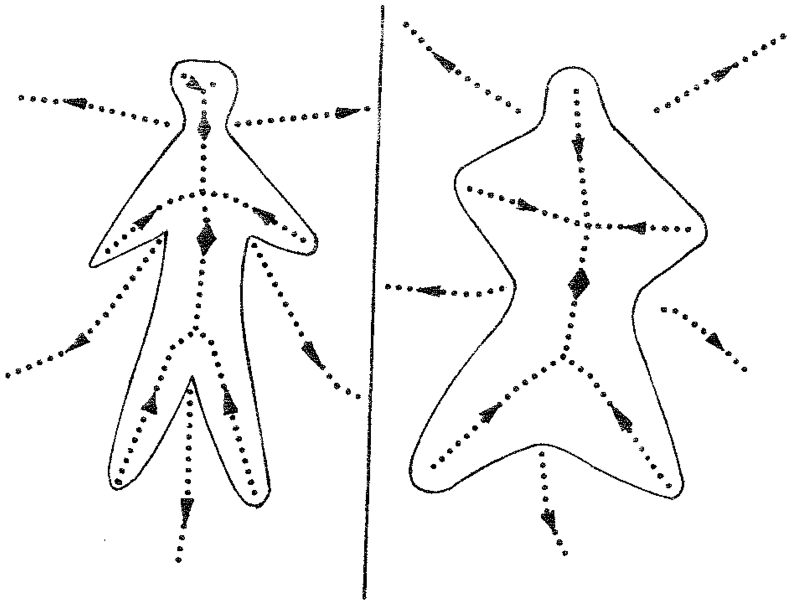


Fig. 6. Two anthropomorphs and their MAFs.

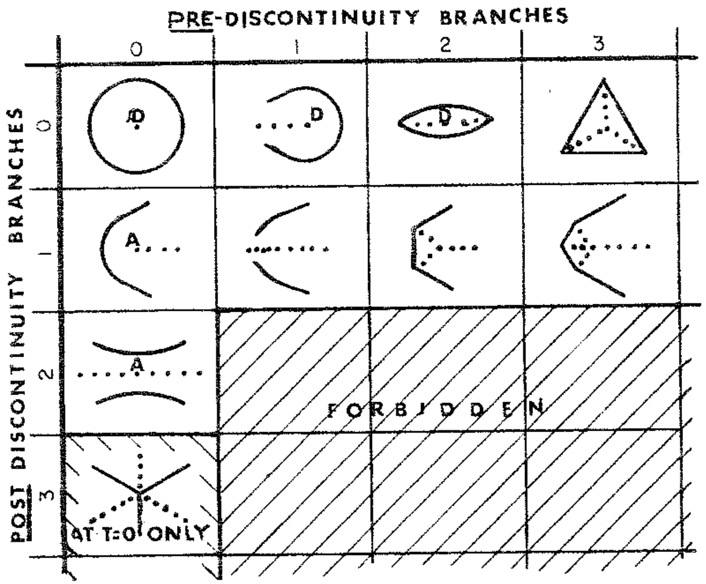


Fig. 13. Shape equivalents of discontinuities on MAF.

History

- There has been computer vision since 1960s
  - Early years
    - Blum (1967): Transformation for extracting descriptors of shape
      - MAF: Medial Axis Function

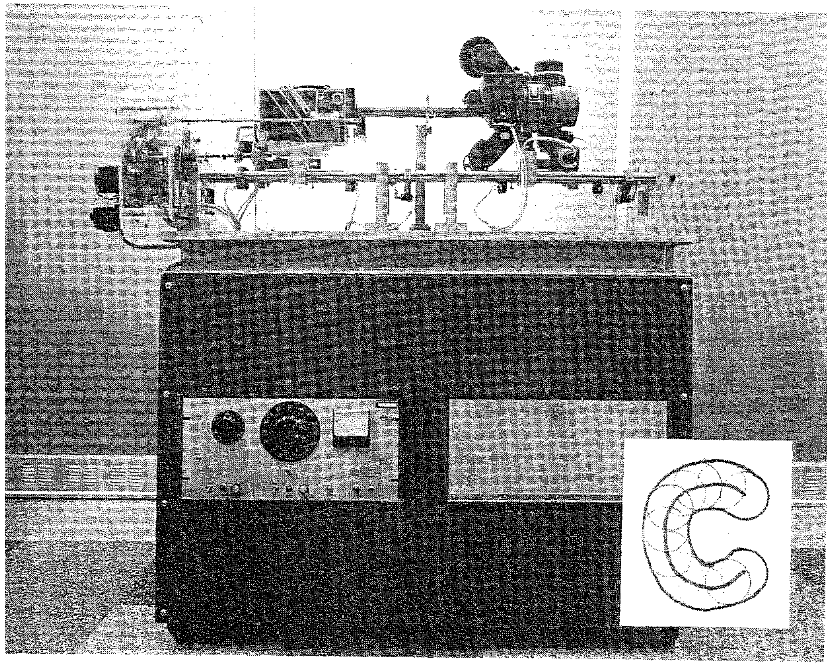


Fig. 19. An early opticomechanical device for obtaining grassfire propagation using photographic defocussing.

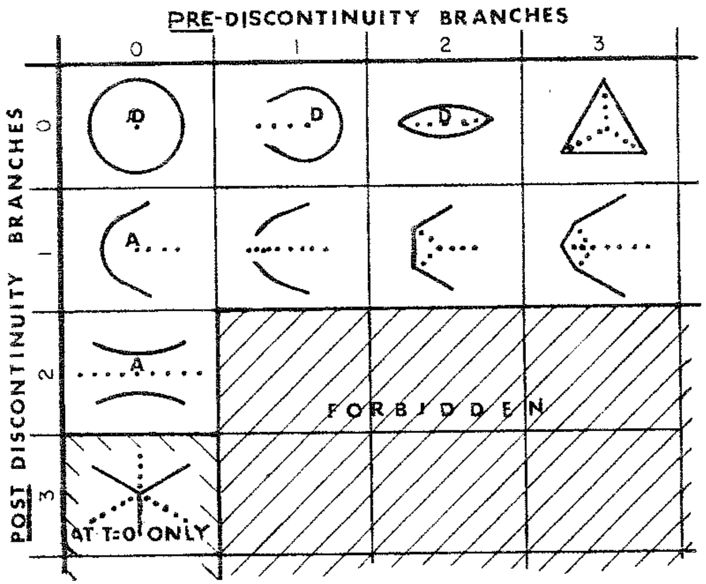
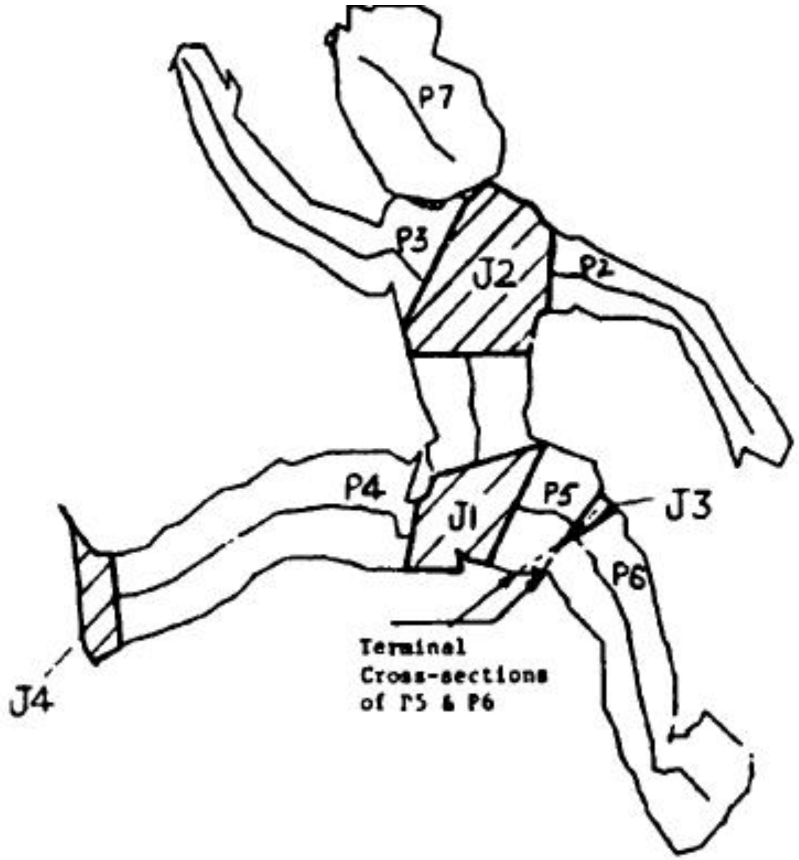


Fig. 13. Shape equivalents of discontinuities on MAF.

History

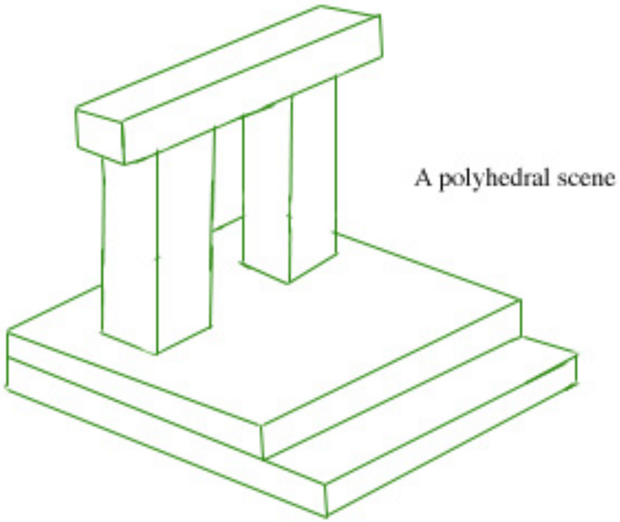
- 1970s: Generalized cylinders
  - Binford, Agin, Nevatia, Marr (1971-1977)



Nevatia and Binford (1977)

History

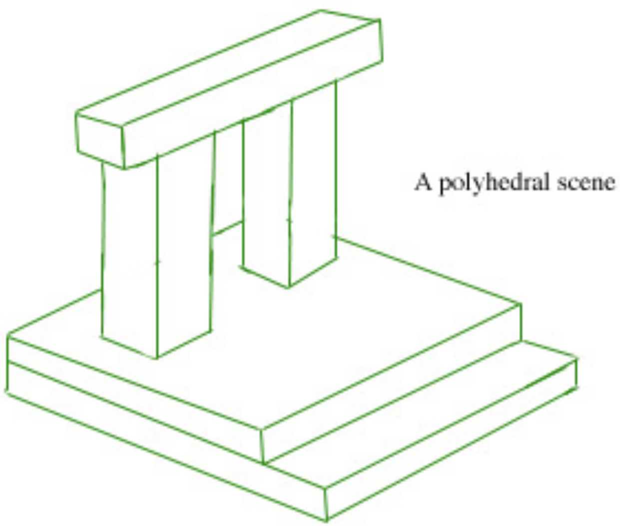
- 1970s: Label line drawings of 3D objects
  - Huffman, Clowes, Waltz (1971-1978)

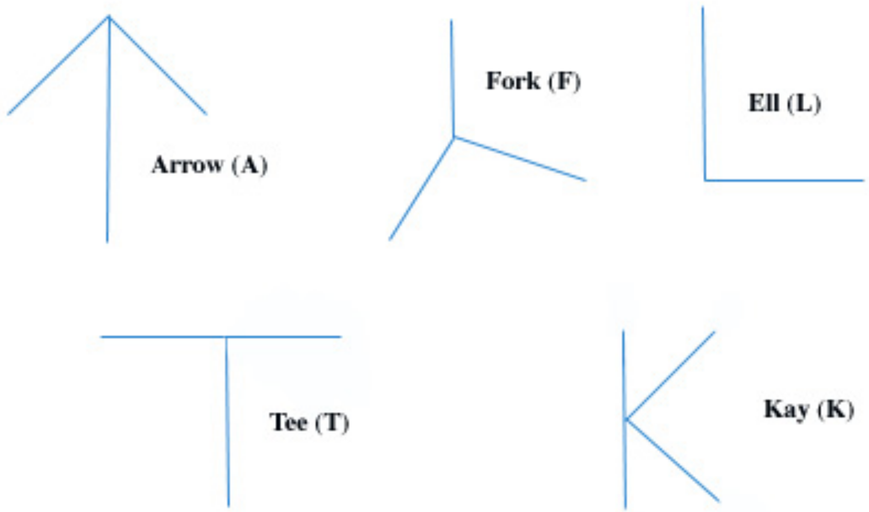


Huffman and Clowes (1977)

History

- 1970s: Label line drawings of 3D objects
  - Huffman, Clowes, Waltz (1971-1978)

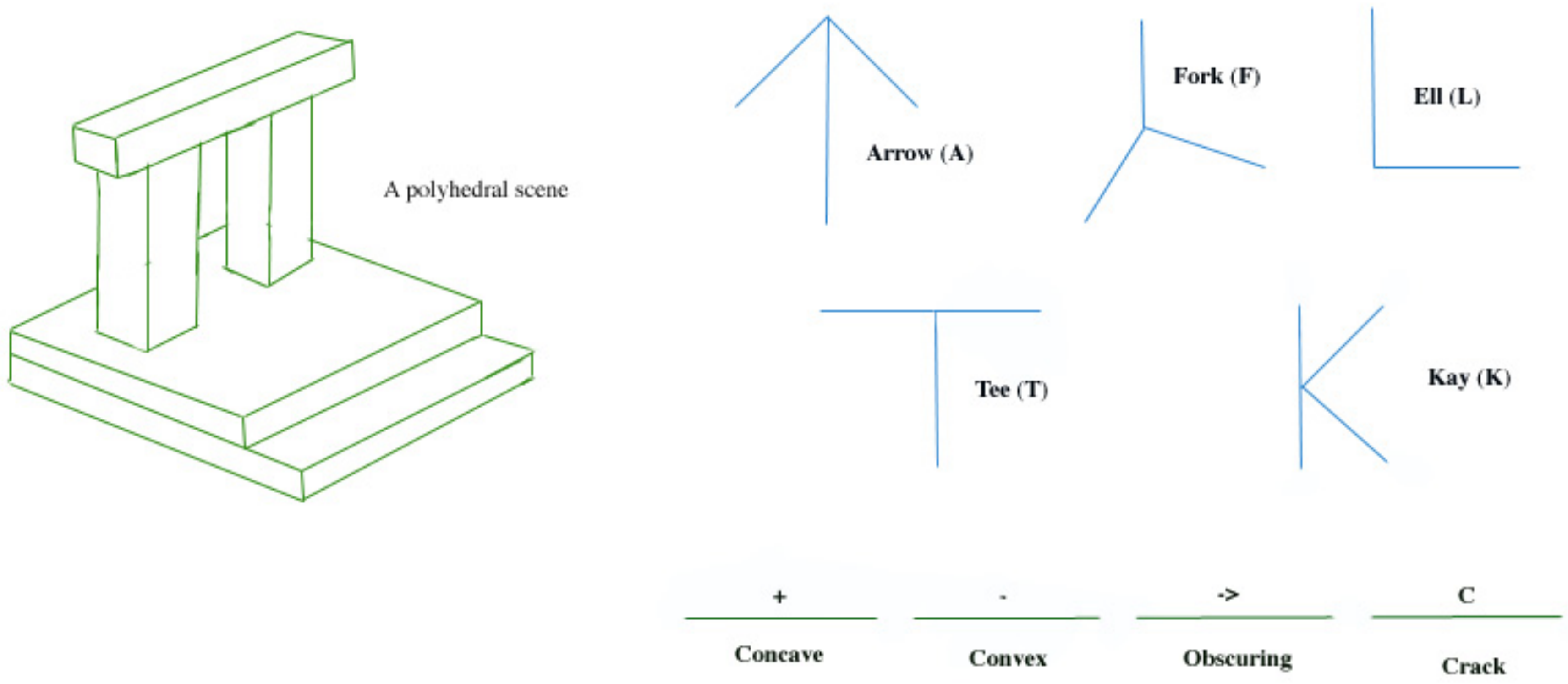




Huffman and Clowes (1977)

History

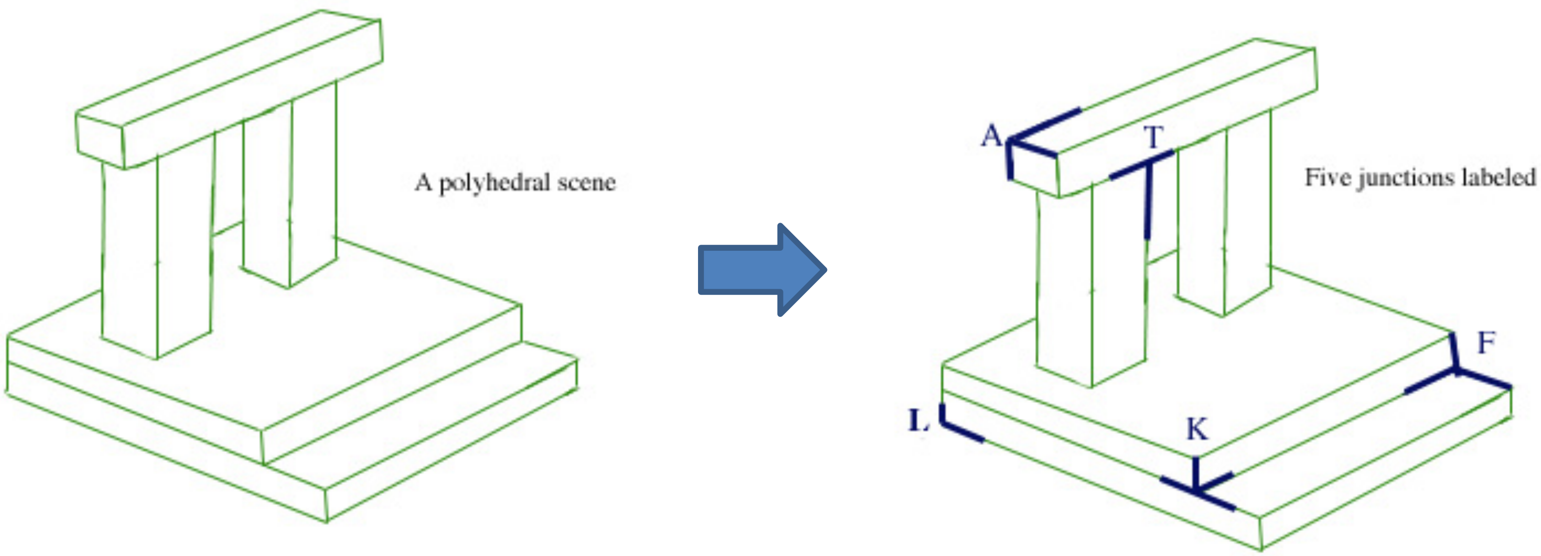
- 1970s: Label line drawings of 3D objects
  – Huffman, Clowes, Waltz (1971-1978)



Huffman and Clowes (1977)

History

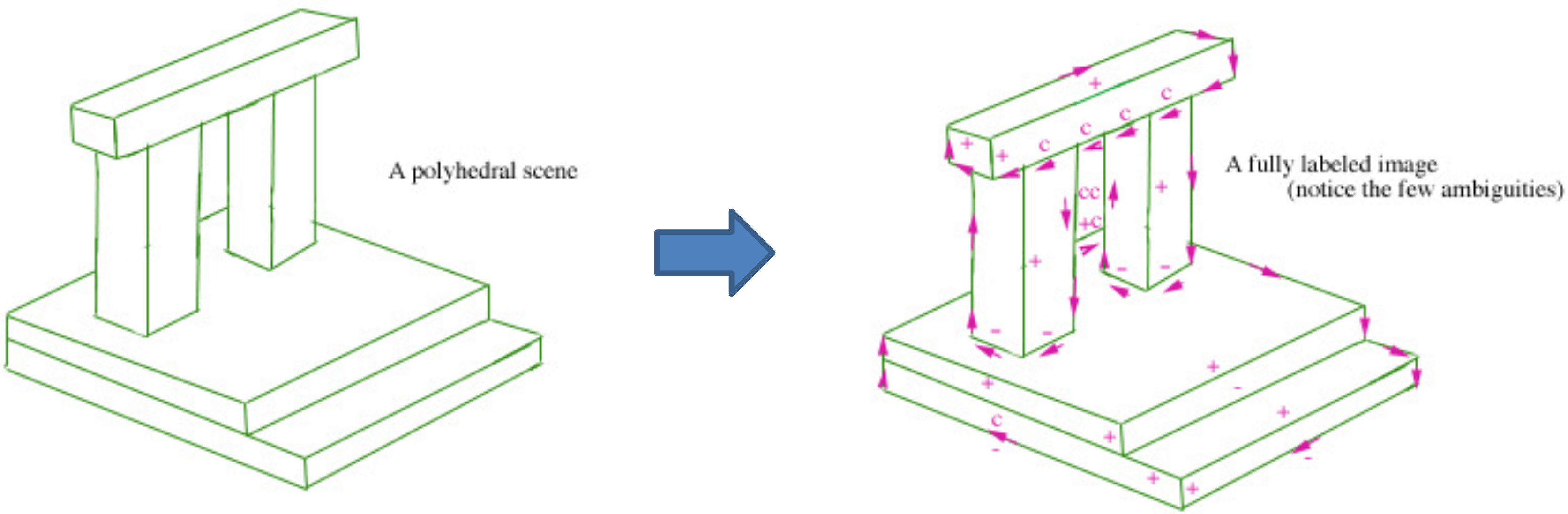
- 1970s: Label line drawings of 3D objects
  - Huffman, Clowes, Waltz (1971-1978)



Huffman and Clowes (1977)

History

- 1970s: Label line drawings of 3D objects
  - Huffman, Clowes, Waltz (1971-1978)



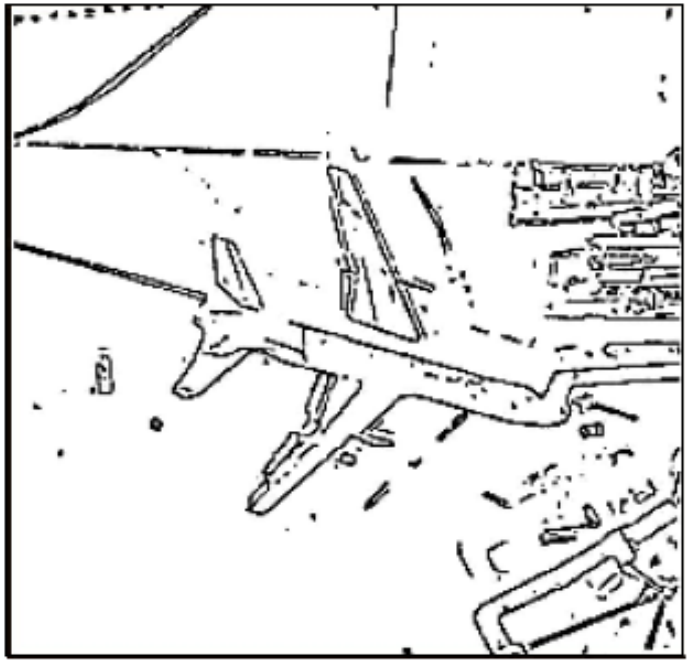
Huffman and Clowes (1977)

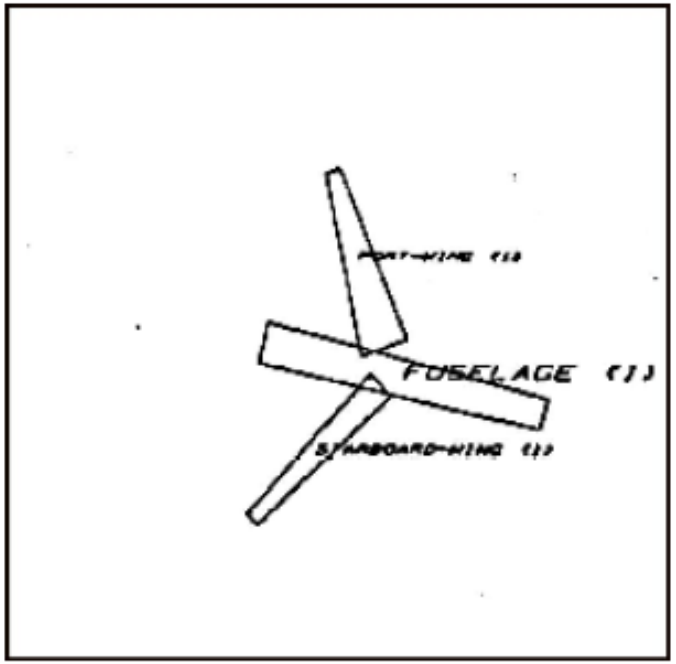
History

- 1970s:
  - MSYS (Barrow and Tanenbaum, 1975)
  - Labeling by Relaxation (Rosenfeld, Hummel, Zucker, 1976)
  - VISIONS schema (Hanson, Riseman, 1974 – 1978)

History

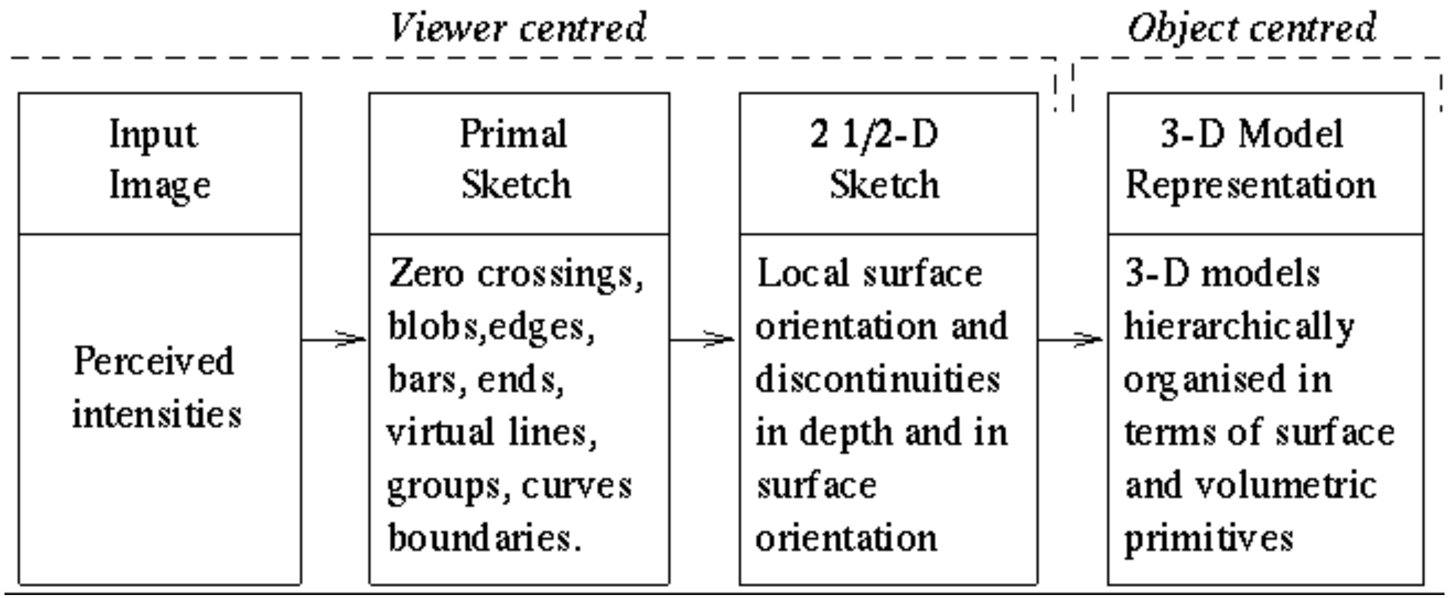
- 1980s:
  - ACRONYM (Brooks and Binford, 1981)





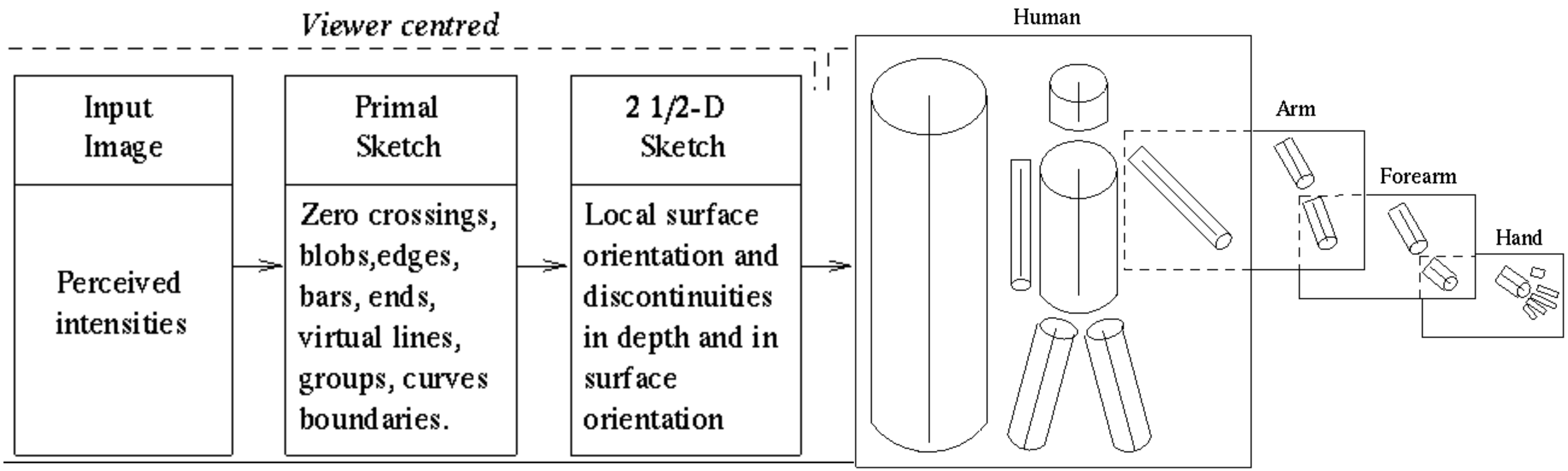
History

- 1980s:
  – Vision (Marr, 1982)



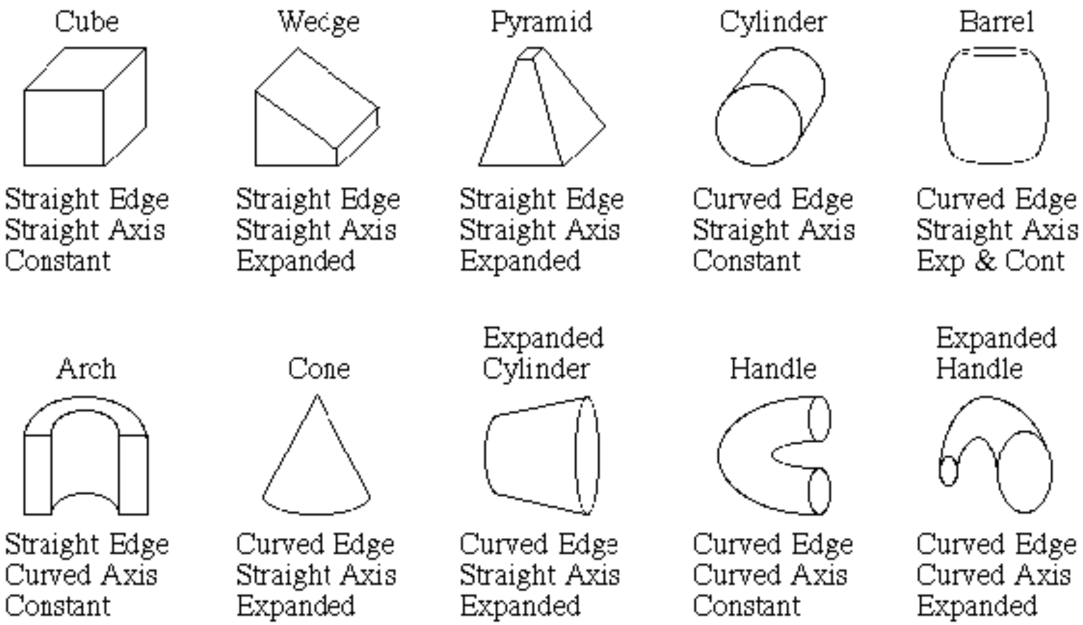
History

- 1980s:
  - Vision (Marr, 1982)



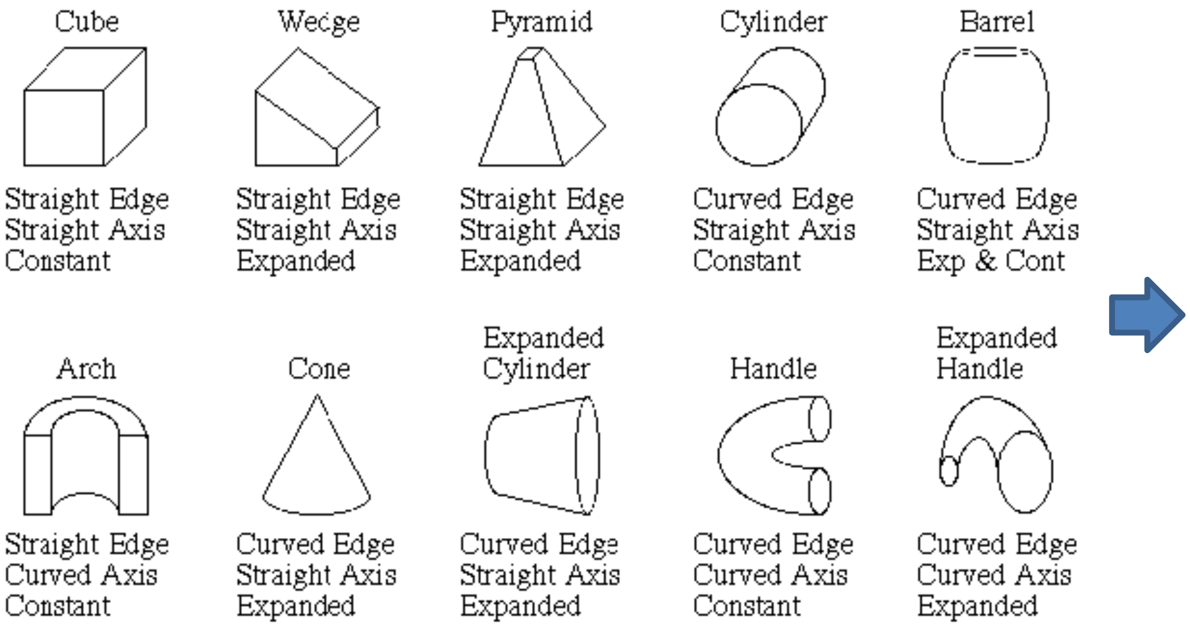
History

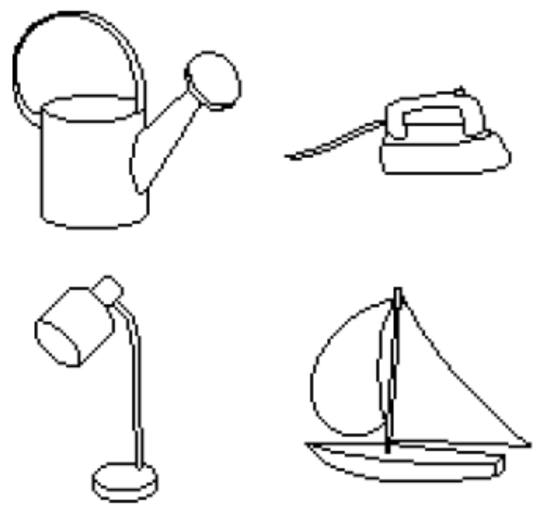
- 1980s: Recognition by Components
  - Geons (Biederman, 1987)



History

- 1980s: Recognition by Components
  - Geons (Biederman, 1987)





History

- 1990s:
  - Early 90s: Invariants, appearance-based methods
  - Mid 90s: Sliding window approaches
  - Late 90s: Feature-based methods, SIFT
  - Early 2000s: Parts and shape models
- 2003 – 2012: Bags of features
  - Present trends: Combination of local and global methods, modeling context, integrating recognition and segmentation
- Since 2012
  - Neural Networks: CNNs, RNNs, LSTMs, capsules, etc.
- More on this in this course

Summary

- Why is vision hard?
- Applications
- Allied areas
- History