COMPUTER VISION

Adhi Harmoko Saputro

UNIVERSITAS INDONESIA
Veritas, Probitas, Iustitia

ABOUT COMPUTER VISION

- Deals with the formation, analysis and Interpretation Images
- Is an integral part of Artificial Intelligence (AI)
- Is interdisciplinary subject area
- Is Practical and useful
- Is modern, challenging and continuously evolving

COMPUTER VISION DIFFICULTIES

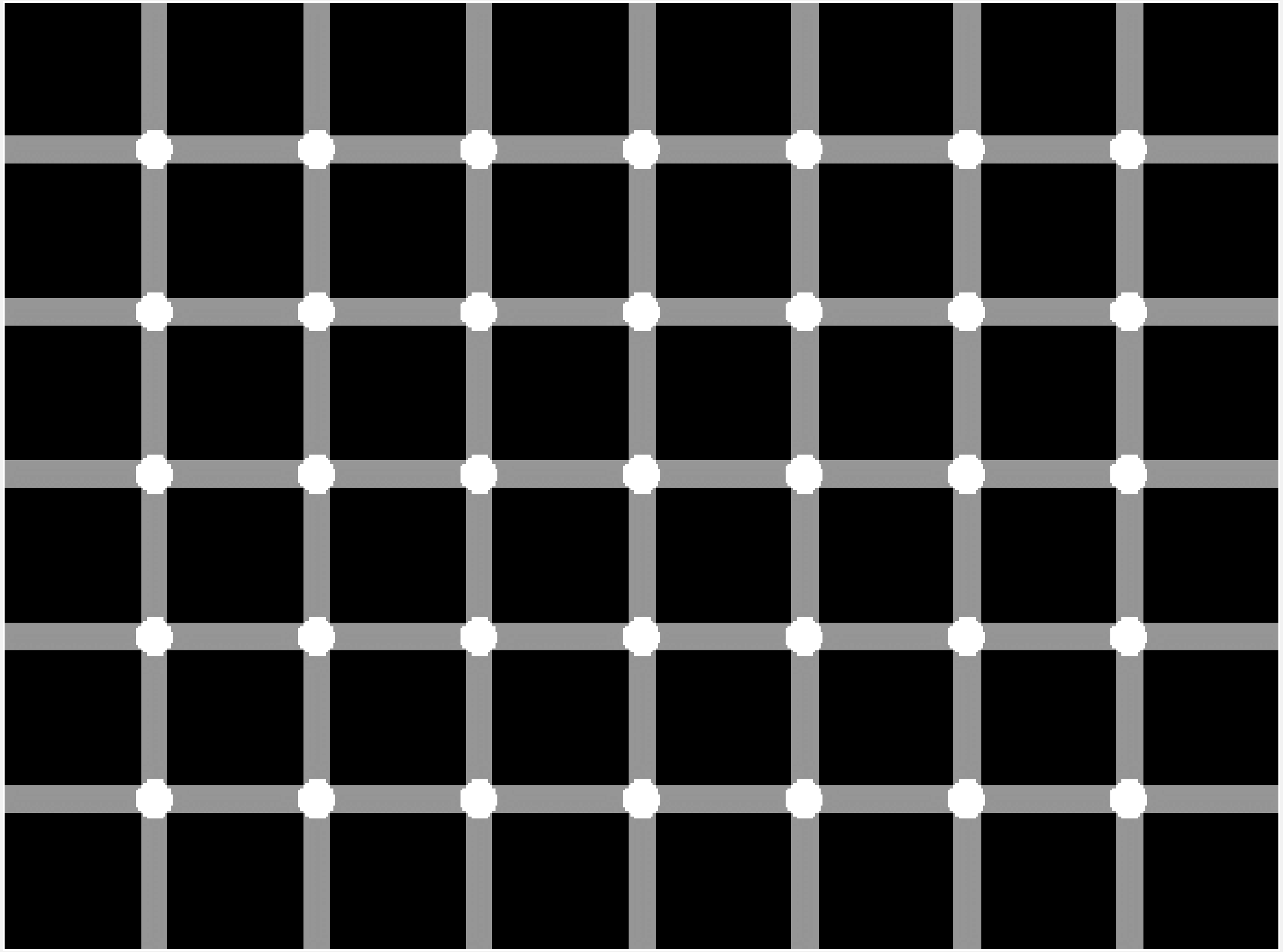
- Images are ambiguous : Projective
- Images are affected by many factors–Sensor model
  - Illumination
  - Shape of object(s)
  - Color of object(s)
  - Texture of object(s)
- There is no “Universal Solution” to the Vision problem
- There are many theories
- We do not understand how the human Visual System works
- Computer technology changes continuously, HenceComputer Vision is evolving.

WHY IS COMPUTER VISION DIFFICULT?

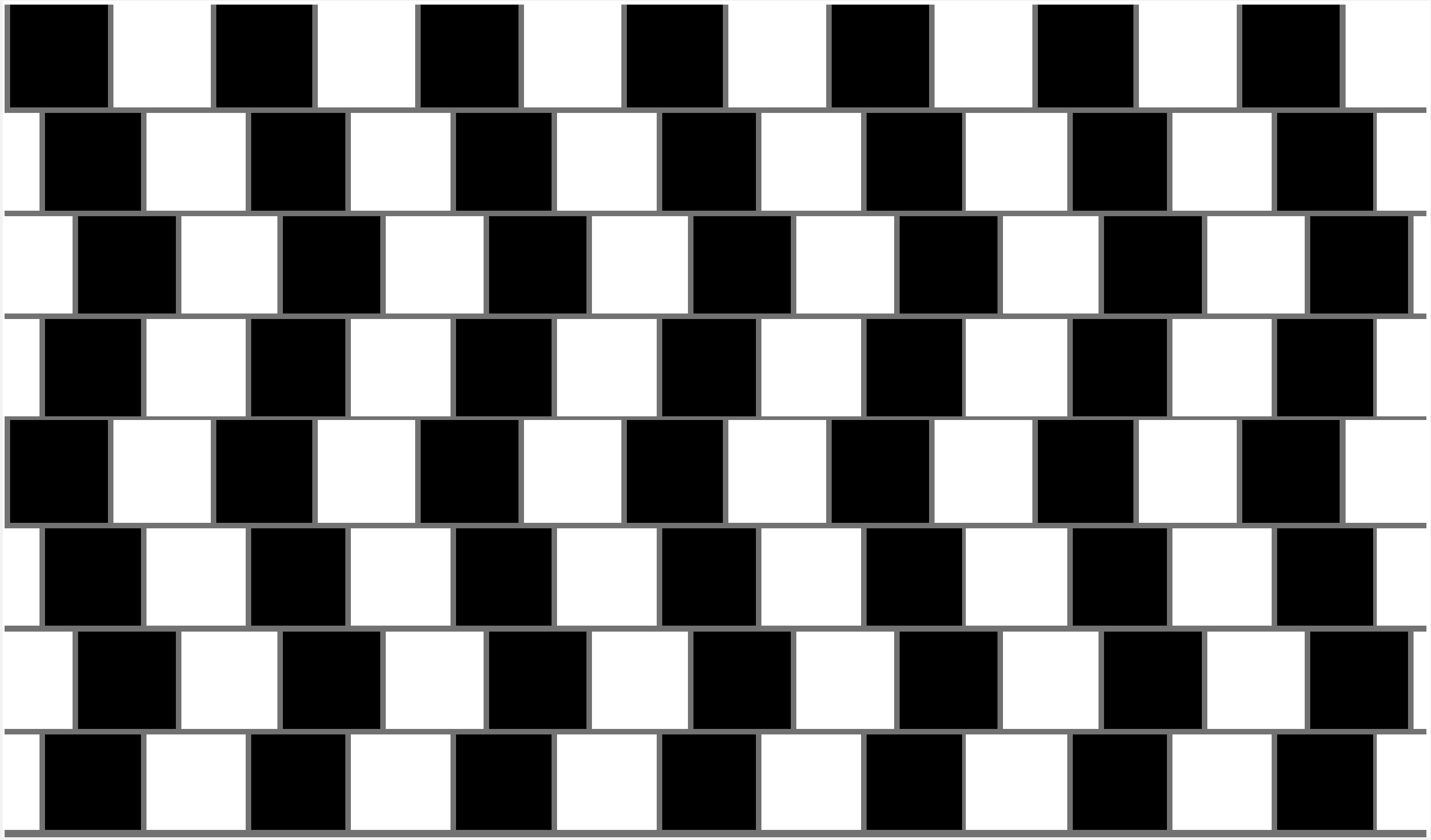
- It is a many-to-one mapping
  - A variety of surfaces with different *material* and *geometrical* properties, possibly under different *lighting* conditions, could lead to identical images
  - Inverse mapping is under-constrained–non-unique solution (a lot of information is lost in the transformation from the 3D world to the 2D image)

- It is computationally intensive
- We still do not understand the recognition problem

ILLUSIONS



ILLUSIONS



WHAT DO HUMANS SEE?

Nomenclature

- Somewhat interchangeable names, with somewhat different implications:
  - **Computer Vision**
    - Most general term
  - **Computational Vision**
    - Includes modeling of biological vision
  - **Image Understanding or Scene Understanding**
    - Automated scene analysis (e.g., satellite images, robot navigation)
  - **Machine Vision**
    - Industrial, factory-floor systems for inspection, measurements, part placement, etc.

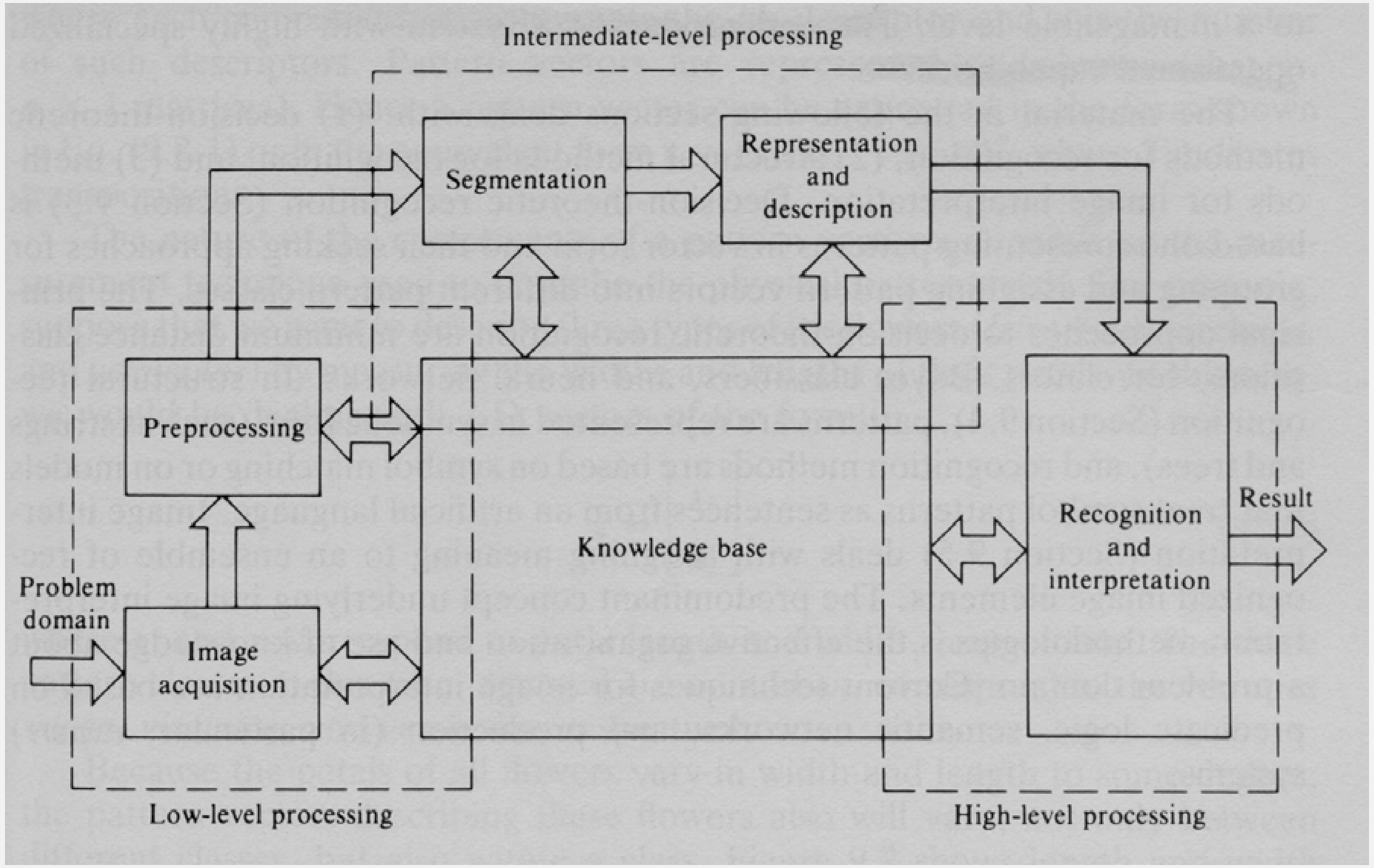
RELATED FIELDS

- Largely built upon
  - Image Processing
  - Statistical Pattern Recognition, Machine Learning
  - Artificial Intelligence

- Related areas
  - Robotics
  - Biological vision
  - Medical imaging
  - Computer graphics
  - Human-computer interaction

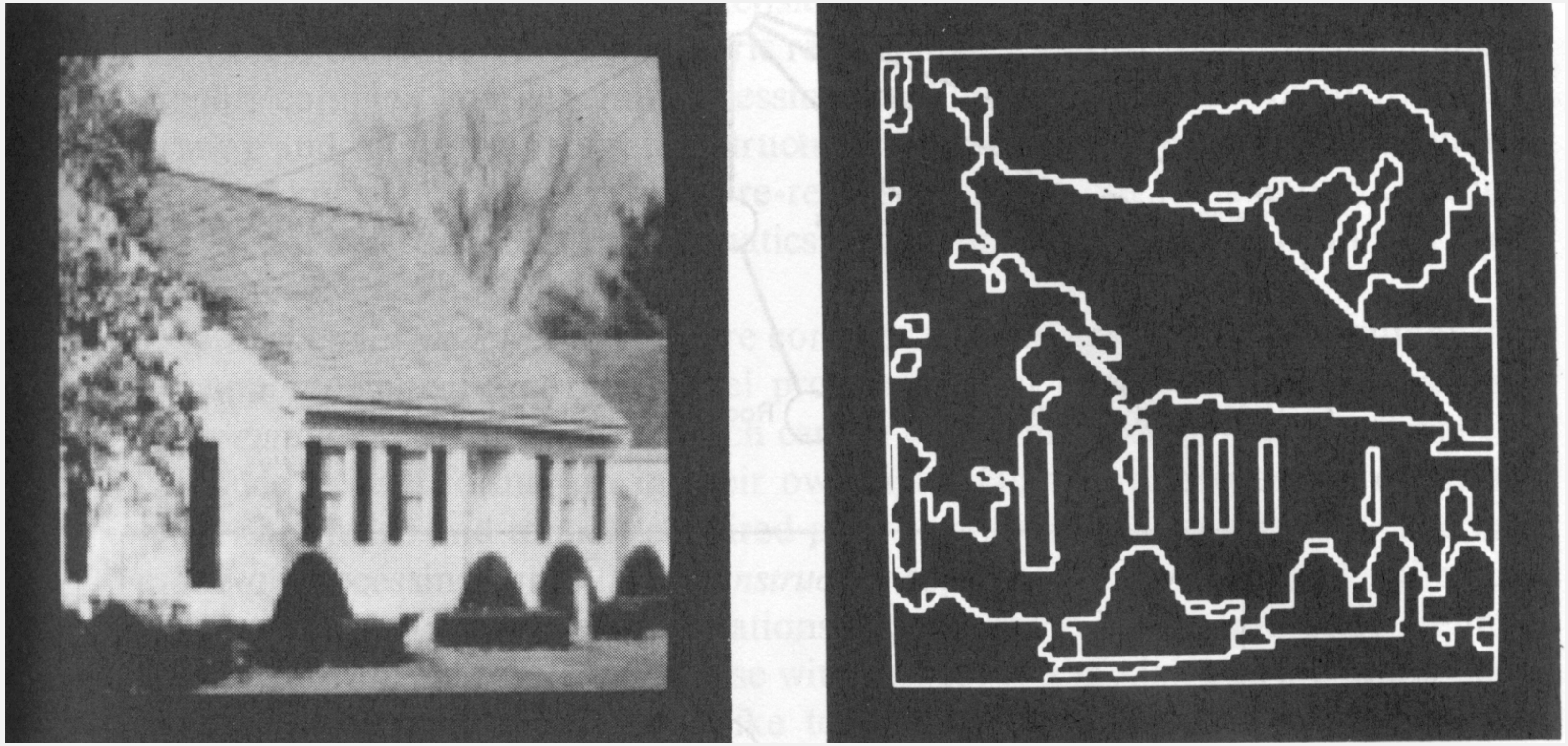
THE THREE PROCESSING LEVELS

- Low-level processing
  - Standard procedures are applied to improve image quality
  - No "intelligent" capabilities



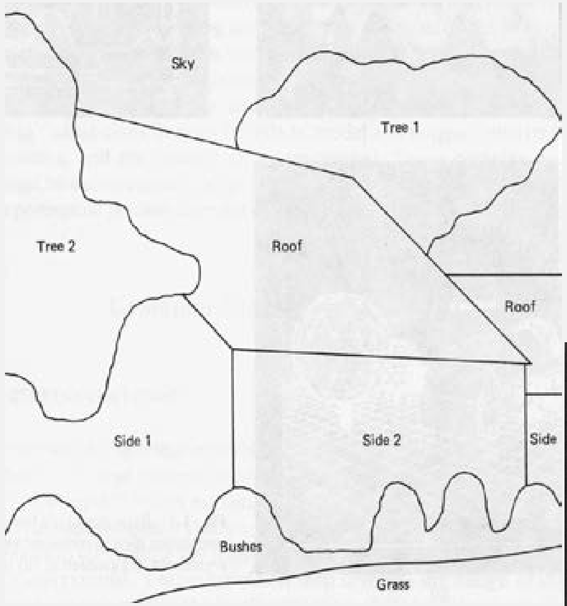
THE THREE PROCESSING LEVELS

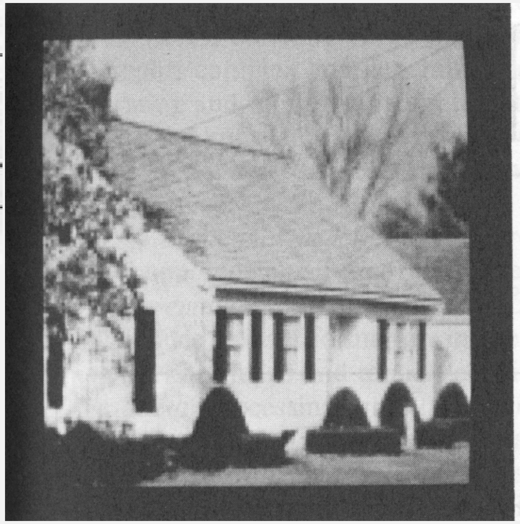
- Intermediate-level processing
  - Extract and characterize components in the image
  - Some intelligent capabilities are required

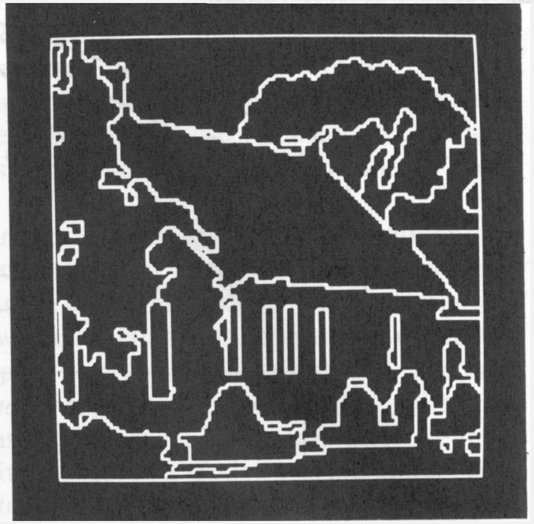


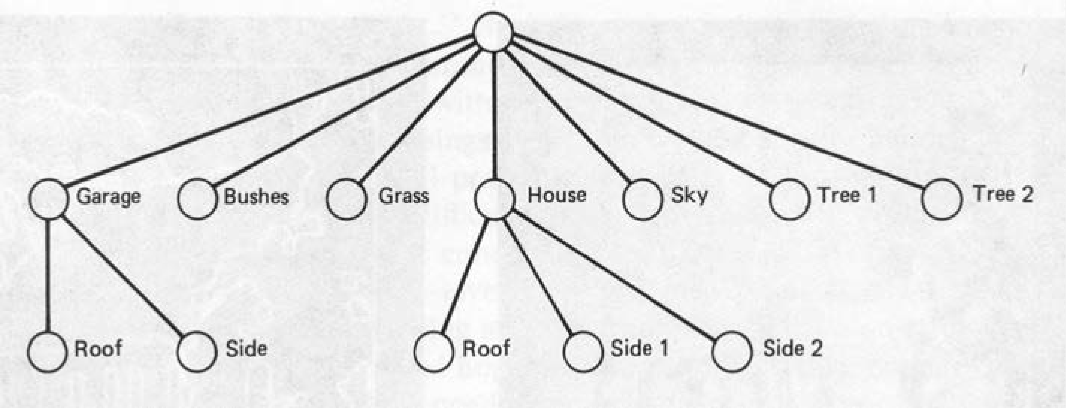
THE THREE PROCESSING LEVELS

- High-level processing
  - Recognition and interpretation
  - Procedures require high intelligent capabilities









COMPUTER VISION

- Some applications
  - **Robotics**
    - Navigation, object manipulation, interaction with humans…
  - Inspection, measurement
  - **Medical imaging**
  - Graphics and animation, special effects
  - Multimedia database indexing and retrieval
  - Human-computer interaction
  - **Surveillance and security**

WHY STUDY COMPUTER VISION?

- Images and movies are everywhere
  - Mobile phone, cheaper cameras enabling the imaging capture

- Fast-growing collection of useful applications–Face finding, recognition, analysis
  - building representations of the 3D world from pictures
  - automated surveillance (who’s doing what)
  - movie post-processing

- Various deep and attractive scientific mysteries–how does object recognition work?

- Greater understanding of human vision

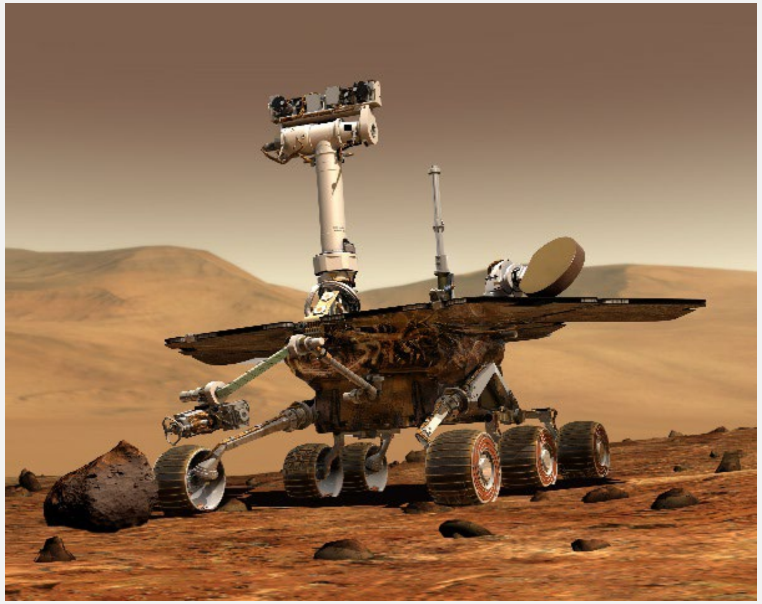
EVERY PICTURE TELLS A STORY

- Goal of computer vision is to write computer programs that can interpret images

EXAMPLES OF COMMERCIAL VISION SYSTEMS

High Precision Inspection System

Par Systems: Vision Guided Robotics for manufacturing and assembly lines

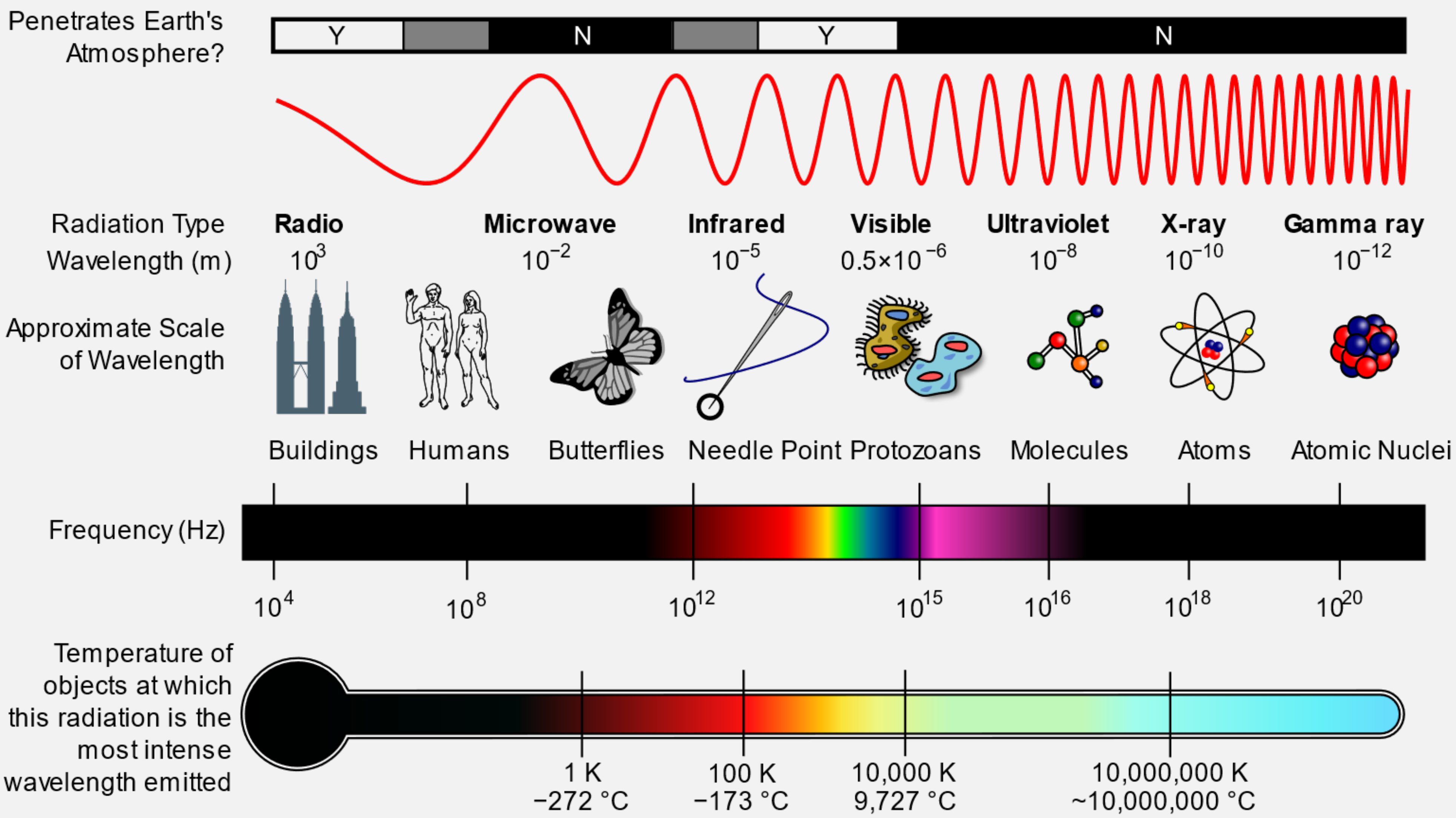


NASA Rover

TYPES OF IMAGES

- Infra-red
- Ultra-violet
- Radio-waves (radio astronomy)
- Visible light
- Micro-waves (radar)
- Roentgen (tomography)
- Sound-waves (echoscopy, sonar)
- Electrons (microscopy)
- Positron emission (PET-scan)
- Magnetic resonance (NMR)

ELECTROMAGNETIC (EM) SPECTRUM



IMAGING IN DIFFERENT WAVELENGTHS

PART I: THE PHYSICS OF IMAGING

- How images are formed

  - Cameras
    - How a camera creates an image
    - How to tell where the camera was
  - Light
    - How to measure light
    - What light does at surfaces
    - How the brightness values we see in cameras are determined
  - Color
    - The underlying mechanisms of color
    - How to describe it and measure it

PART II: EARLY VISION IN ONE IMAGE

- Representing small patches of image
  - For three reasons
    - We wish to establish correspondence between (say) points in different images, so we need to describe the neighborhood of the points
    - Sharp changes are important in practice ---known as “edges”
    - Representing texture by giving some statistics of the different kinds of small patch present in the texture.
      - Tigers have lots of bars, few spots
      - Leopards are the other way

REPRESENTING AN IMAGE PATCH

- Filter outputs
  - essentially form a dot-product between a pattern and an image, while shifting the pattern across the image
  - strong response -> image locally looks like the pattern
  - e.g. derivatives measured by filtering with a kernel that looks like a big derivative (bright bar next to dark bar)

TEXTURE

- Many objects are distinguished by their texture
  - Tigers, cheetahs, grass, trees

- We represent texture with statistics of filter outputs
  - For tigers, bar filters at a coarse scale respond strongly
  - For cheetahs, spots at the same scale
  - For grass, long narrow bars
  - For the leaves of trees, extended spots

- Objects with different textures can be segmented
- The variation in textures is a cue to shape

PART III: EARLY VISION IN MULTIPLE IMAGES

- The geometry of multiple views
  - Where could it appear in camera 2 (3, etc.) given it was here in 1 (1 and 2, etc.)?

- Stereopsis
  - What we know about the world from having 2 eyes

- Structure from motion
  - What we know about the world from having many eyes
    - or, more commonly, our eyes moving.

Part IV: Mid-Level Vision

- Finding coherent structure so as to break the image or movie into big units
    - Segmentation:
        - Breaking images and videos into useful pieces
        - E.g. finding image components that are coherent in internal appearance
    - Tracking:
        - Keeping track of a moving object through a long sequence of views

Part V: High Level Vision (Geometry)

- The relations between object geometry and image geometry
  - Model based vision
    - find the position and orientation of known objects
  - Smooth surfaces and outlines
    - how the outline of a curved object is formed, and what it looks like
  - Aspect graphs
    - how the outline of a curved object moves around as you view it from different directions
  - Range data

PART VI: HIGH LEVEL VISION (PROBABILISTIC)

- Using classifiers and probability to recognize objects
    - Templates and classifiers
        - how to find objects that look the same from view to view with a classifier
    - Relations
        - break up objects into big, simple parts, find the parts with a classifier, and then reason about the relationships between the parts to find the object.
    - Geometric templates from spatial relations
        - extend this trick so that templates are formed from relations between much smaller parts

CONCLUSION

- Computer Vision is an exciting and challenging field
- Applied in many real world solutions
- Tremendous progress in the last decades
- Still have long way to go …

TERIMA KASIH

Adhi Harmoko Saputro

UNIVERSITAS INDONESIA
Veritas, Probitas, Iustitia