HA NOI UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

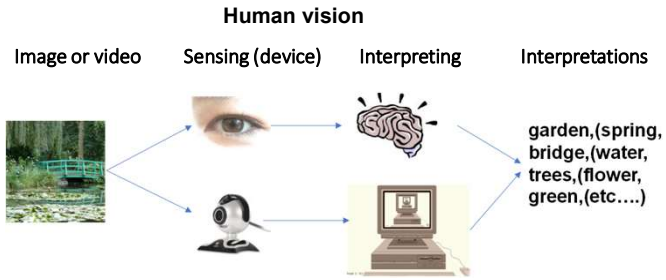
Computer Vision

Chapter 1: Introduction

Plan

- What is computer vision?
  - Concepts and definitions
  - Levels of vision (Low level vision, Middle level vision, High level vision)
- Related fields
- Applications areas

What is computer vision?



From CS131 course “computer vision”,
Prof. Fei-Fei Li, Stanford 'Vision'Lab

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

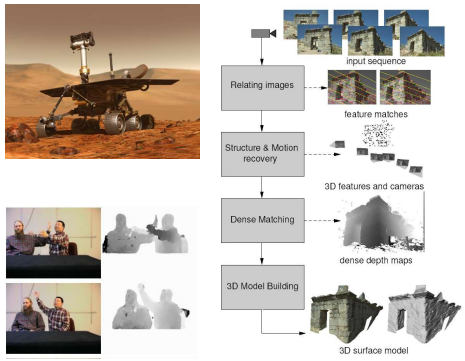
What a computer sees

Source: S. Narasimhan

What kind of information can we extract from an image?

- Metric 3D information
- Semantic information

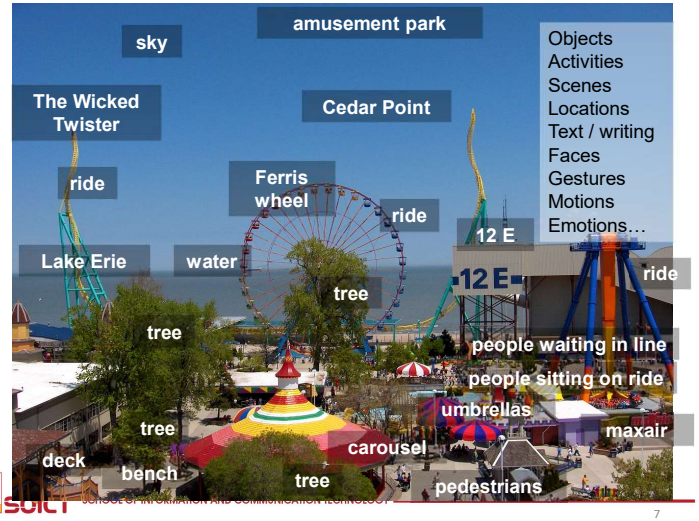
Vision as measurement device



Pollefeys et al.

Goesele et al.

Vision as a source of semantic information



Slide credit: Kristen Grauman

What is computer vision?

- Computer vision
  - Is an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos.
  - From the perspective of engineering, it seeks to automate tasks that the human visual system can do.
- Computer vision tasks include
  - methods for acquiring, processing, analyzing and understanding digital images,
  - and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions. (*Wikipedia*).

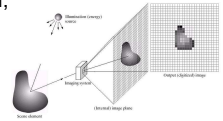
What is computer vision?

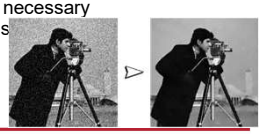
- The two definitions of CV can be defined as a scientific field that extracts information out of digital images.
- Another way to define CV is through its applications.
  - Computer vision is building algorithms that can understand the content of images and use it for other application [3].

What kind of scene?
Where are the cars?
How far is the building?
…

What is computer vision?
Levels of vision

- **Low-level Vision**: Image Formation, Acquisition, Image Processing
  - **Image formation** studies the forward process of producing images and videos
  - **Image acquisition**:
    - A digital image is produced by several image sensors.
    - Depending on the type of sensor, the resulting ***image data*** is an ordinary 2D image, a 3D volume, or an image sequence.
  - **Image processing** focuses on 2D image data processing using point operators such as contrast enhancement, filtering (local operations), noise reduction, image transforms. Image processing is considered as pre-processing that is usually necessary to process the image data for CV applications
    - Work with image as a matrix
    - Input: image ➔ output: image

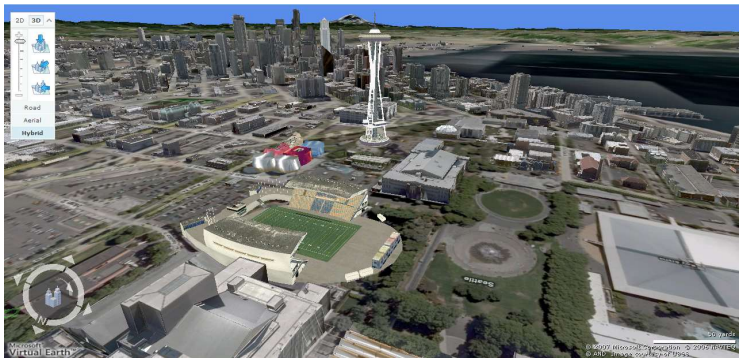




What is computer vision?
Levels of vision

- **Middle-level Vision:** Feature, Image matching
  - Feature extraction: Image features at various levels of complexity are extracted from the image data. Examples of such features: Edges, ridges, lines, texture, shape …
  - Image matching
  - Image segmentation

3D urban modeling

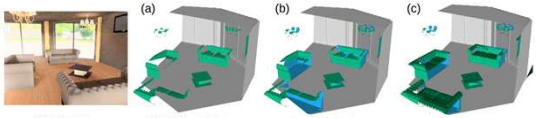


Bing maps, Google Streetview

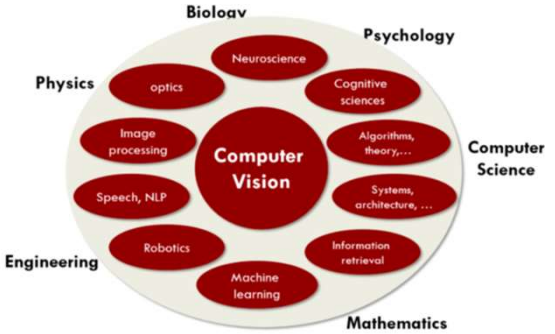
Source: S. Seitz

What is computer vision?
Levels of vision

- **High-level Vision**: High-level vision is to infer the semantics, for example, object recognition and scene understanding.
- Several application topics:
  - Object recognition (classification), Identifica
  - Detection
  - Motion analysis
  - Scene reconstruction; 3D reconstruction
  - Image-base

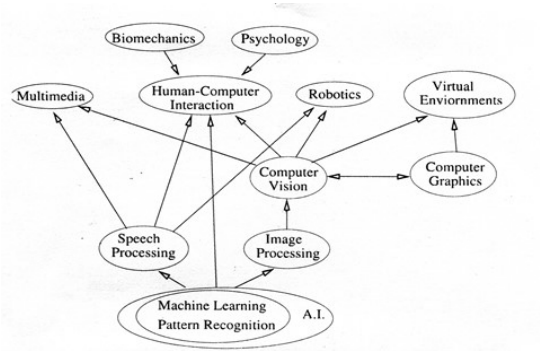


Related fields



Computer vision at the intersection of multiple scientific fields [3]

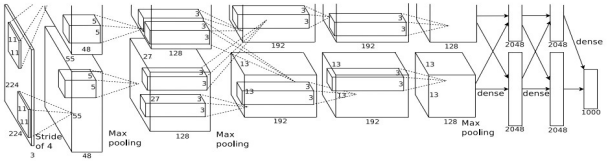
Related fields



Source: From EECS 432-Advanced Computer Vision, Northwestern University

Related fields

- Machine Learning: “The field of study that gives computers the ability to learn without being explicitly programmed.” – *Arthur Samuel*
- Artificial intelligence and computer vision share other topics such as pattern recognition and learning techniques.
- Computer vision - Deep learning: Artificial Neural Networks with many layers (CNN: Convolutional Neural Network)



Applications areas

Robotics Application

- Localization-determine robot location automatically
- Navigation
- Obstacles avoidance
- Assembly peg − in − hole, welding, painting
- Manipulation e. g. PUMA robot manipulator
- Human Robot Interaction HRI: Intelligent robotics to interact with and serve people

Applications areas

Security Application

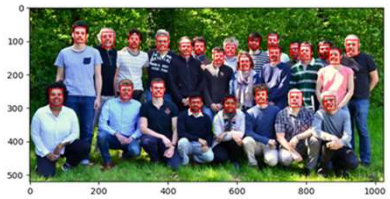
- Biometrics iris, fingerprint, face recognition
- Surveillance-detecting certain suspicious activities or behaviors
- …

Fingerprint scanners

Face recognition systems

Source: from S. Seitz

Examples
Face Detection



Source: from S. Seitz

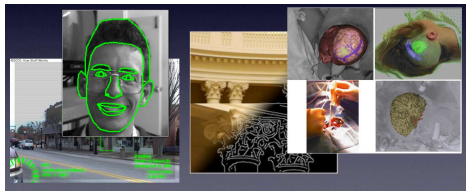
Examples of Computer Vision

Slide from Vicente Ordonez

Applications areas

Medicine Application

- Classification and detection e. g.
- 2D/3D segmentation
- 3D human organ reconstruction MRI or ultrasound
- Vision-guided robotics surgery
- …



Slide from Jason Lawrence

Applications areas

Industrial Automation Application

- Industrial inspection defect detection
- Barcode and package label reading
- Object sorting
- Document understanding e. g. OCR
- …

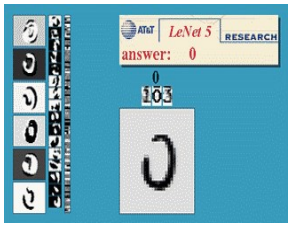
Transportation Application

- Autonomous vehicle
- Safety, e.g., driver vigilance monitoring
- …

Facebook's suggestion

Examples
Toys and Robots

Examples
Optical character recognition (OCR)



Digit recognition,(AT&T labs)

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

Source: from S. Seitz