What is computer vision?

Computer vision is concerned with modeling and replicating human vision using computer software and hardware. It combines knowledge in computer science, electrical engineering, mathematics, physiology, biology, and cognitive science. It needs knowledge from all these fields in order to understand and simulate the operation of the human vision system.
What is computer vision? (cont’d)

Computer vision (image understanding) is a discipline that studies how to reconstruct, interpret and understand a 3D scene from its 2D images in terms of the properties of the structures present in the scene.
Typical hardware components of a computer vision system

We focus on computer vision algorithms and their software implementation.
Computer Vision Hierarchy

• Low-level vision: process image for feature extraction (edge, corner, or optical flow).

• Intermediate-level vision: object recognition and 3D scene interpretation using features obtained from the low-level vision.

• High-level vision: interpretation of the evolving information provided by the intermediate level vision as well as directing what intermediate and low level vision tasks should be performed. Interpretation may include conceptual description of a scene like activity, intention and behavior.
Computer Vision Hierarchy (cont’d)
Why Is Computer Vision Difficult?

- The problem is *ill-posed inverse* problem.
- Noisy image data or data with uncertainties.
Related Fields

Computer vision overlaps significantly with the following fields: image processing, pattern recognition, and photogrammetry.

Image processing focuses on image manipulation to enhance image quality, to restore an image or to compress/decompress an image. Most computer vision algorithms usually assume a significant amount of image processing has taken place to improve image quality.

Pattern recognition studies various techniques (such as statistical techniques, neural network, support vector machine, etc..) to recognize/classify different patterns.
Pattern recognition techniques are widely used in computer vision.

Photogrammetry is concerned with obtaining accurate and reliable measurements from images. It focuses on accurate mensuration. Camera calibration and 3D reconstruction are two areas of interest to both computer vision and photogrammetry researchers.
Computer Vision v.s. Image Processing

Image processing studies *image-to-image transformation*. The input and output of image processing are both images. Typical image processing operations include

- image compression
- image restoration
- image enhancement
Computer Vision v.s. Image Processing (cont’d)

Computer vision is the construction of explicit, meaningful descriptions of physical objects from their images. The output of computer vision are a description or an interpretation or some quantitative measurements of the structures in the 3D scene. Image processing and pattern recognition are among many techniques computer vision employs to achieve its goals.
What is robot vision?

Robot vision applies computer vision techniques to robotics applications. Specifically, it studies the machine vision in the context of robot control and navigation.
Tasks of Robot Vision

Given one or a sequence of images of a 3D scene, the tasks of robot vision include

- Avoid static and moving obstacles.
- Build models of objects and places in order to recognize and locate them.
- Characterize its own position and motion and those of the moving objects.
Tasks of Robot Vision (cont’d)

- determine the identities of the objects in the scene (object recognition).
- determine the pose of the objects.
- determine the motion of the objects.
- reconstruct the 3D geometry of objects from their 2D images for measurements.
- build maps of the environments for robot’s navigation.
Issues involved with Robot Vision

- What information should be extracted from the output of the visual sensors?
- How is the information extracted?
- How is the extracted information represented?
- How must the information be used to allow robotics system to perform its tasks?
What do we emphasize?

- Emphasize the geometry of objects and the use of prior knowledge about the geometry of objects for 3D object recognition and reconstruction, and for improving the accuracy of the estimated parameters.

- Emphasize representation and propagation of uncertainties with image data.
Example Applications

- Robotics
- Medicine
- Security
- Transportation
- Industrial automation
- Image database
- Human Computer Interface
Robotic Applications

- Localization—determine robot location automatically
- Obstacles avoidance
- Navigation and visual servoing
- Assembly (peg-in-hole, welding, painting)
- Manipulation (e.g. PUMA robot manipulator)
- Intelligent robotics to interact with and serve people
Figure 1: NASA rover for planetary surface exploration
Figure 2: A vision-guided welding machine
Figure 3: Real time visual servoing for robot grasping
Industrial Automation

- Industrial inspection (defect detection and mensuration)
- Assembly
- Barcode and package label reading
- Object sorting
- Document understanding (e.g. OCR)
Figure 11.3: Results of feature extraction and 3D reconstruction for part 3.
Medicine

- Classification and detection (e.g. lesion or cells classification and tumor detection)
- 2D/3D segmentation
- 3D human organ reconstruction (MRI or ultrasound)
- Vision-guided robotics surgery
Medical Imaging

Figure 1.4: Magnetic resonance image (left) where brightness relates to material movement and binary image (right) resulting from changing all pixels with value 208 or more to 255 and those below 208 to 0.
Medical Imaging (cont’d)

Figure 1.7: (Top left) Binary microscope image of red blood cells; (top right) cleaner image resulting from removal of tiny dark regions inside light regions or visa versa; (bottom) templates showing how pixel neighborhoods can be cleaned.
Security

- Biometrics (iris, fingerprint, face recognition)
- Surveillance-detecting certain suspicious activities or behaviors
Transportation

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

It is mainly used for image retrieval based on image content.

Figure 1.2: Image query by example: query image (left) and two most similar images produced by an image database system (from the MSc thesis of Aditya Vallaya).
Human Computer Interface

- Gaze estimation
- Face expression recognition
- Head and hand gesture recognition
Head Pose and Gaze
Computer Vision Literature

1. Journals
   - IEEE transactions on Pattern Recognition and Machine Intelligence (PAMI)
   - International Journal of Computer Vision
   - Computer vision and image understanding
   - Machine vision and application
   - Image and vision computing
   - Pattern recognition

2. Conferences
   - International conference on computer vision
(ICCV)

- IEEE conference on computer vision and pattern recognition (CVPR)

- International conference on image processing (ICIP)

- International conference on pattern recognition (ICPR)

- IEEE conference on robotics and automation
Computer Vision Resources

Computer vision homepage:
http://www.cs.cmu.edu/~cil/vision.html

Additional links for computer vision may be found
http://www.cns.nyu.edu/~eero/vision-links.html

- Publications
- News groups
- Software
- Vision groups
- Test images
Image Processing Resources

- image processing newsgroup:
  sci.image.processing

- Fundamentals of Image Processing
  http://
  www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip.html

- An Image Processing Tutorial
  http://www.utu.fi/ml/kartografia/hypermedia/HTML/glossary.htm
Mathematical Background Needed

- Numerical analysis
- Statistics
- Linear and non-linear optimization and regression
- Computational geometry
- Projective geometry
- Programming skills
- Digital signal processing
Outcomes

• understand the fundamental computer vision theories

• have the ability to design and implement computer vision techniques

• have the capability of applying computer vision technologies to applications of interest.
References to Computer Vision Terminologies
