Introduction to Computer Vision

What is computer vision?

Computer vision is the automated extraction of information from images and video.

- Measurement: deals with problems of measurement of real-world data.
- Modeling visual perception and understanding: algorithms to recognize people, objects, scenes, and activities.
- Retrieval of visual data: algorithms to mine, search and organize large data sets.

It is called machine vision for industrial vision applications where the tasks are more specific and usually coupled with robotic applications.

Related fields of computer vision

- Image acquisition
- Image processing
- Computer graphics
- Machine learning/deep learning
- Cognitive science
- Algorithms

The challenge

"…However, despite all of these advances, the dream of having a computer interpret an image at the same level as a two-year old (for example, counting all of the animals in a picture) remains elusive. Why is vision so difficult? In part, it is because vision is an inverse problem, in which we seek to recover some unknowns given insufficient information to fully specify the solution. We must therefore resort to physics-based and probabilistic models to disambiguate between potential solutions..." [Computer Vision: Algorithms and Applications (2010) by Richard Szeliski]

A general approach

"… it is my conviction that a careful analysis of the problem specification and known constraints from image formation and priors (the scientific and statistical approaches) must be married with efficient and robust algorithms (the engineering approach) to design successful vision algorithms…" [Szeliski]

The goal

"The ultimate goal of computer vision is to mimic human visual perception. Therefore, in the broadest sense, robustness of a computer vision algorithm is judged against the performance of a human observer performing an equivalent task." [Robust Techniques for Computer Vision by Peter Meer in Emerging Topics in Computer Vision (2005)]

Typical CV tasks:

- Object Classification: What broad category of object is in this photograph?
- Object Identification: Which type of a given object is in this photograph?
- Object Verification: Is the object in the photograph?
- Object Detection: Where are the objects in the photograph?
- Object Landmark Detection: What are the key points for the object in the photograph?
- Object Segmentation: What pixels belong to the object in the image?
- Object Recognition: What objects are in this photograph and where are they?

[Brownlee; https://machinelearningmastery.com/what-is-computer-vision/]

Some computer vision applications

• Optical character recognition (OCR)
  ○ Scanned text
  ○ Handwritten postal codes on letters
  ○ Optical music recognition (OMR)
  ○ Automatic number plate recognition in live traffic
	

• Biometrics and fingerprint recognition
  ○ Face detection
  ○ 3D face recognition
  ○ Automatic access authentication
  ○ Forensic applications
	 

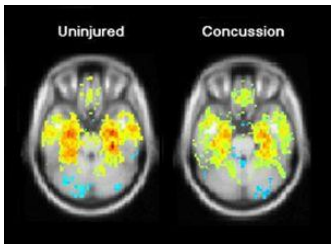
• Surveillance
  ○ Occupancy monitoring
  ○ Intruder or fire detection
  ○ Traffic analysis
  ○ Unattended luggage detection
	

- Mobile applications
  - QR codes
  - Computational photography (Lens Blur, face detection)
  - Facial expression detection
  - Snapchat filters (face tracking)

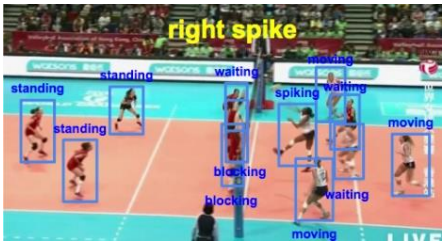
  - Panorama construction
  - Image stitching (Google maps)

- Satellite imagery
  - Crop yield prediction
  - Google's project sunroof
  - Ship tracking
  - General surveillance

- Medical imaging
  - Assisted diagnosis from x-rays, MRI and ultrasound images.
  - Pathological analysis
  - Studying evolution of brain morphology

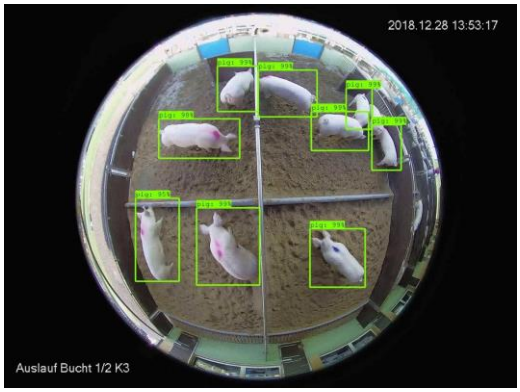


- Quality assurance in production
  - Parts inspection on production lines
  - High-resolution analysis of precision parts
  - Defects in steel castings using X-ray vision

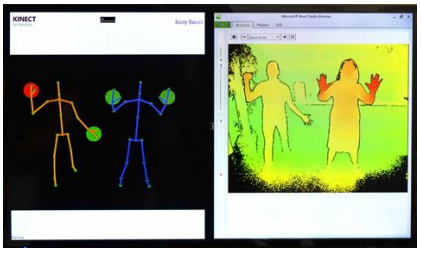
• General recognition
  ○ Image retrieval based on content
  ○ Face recognition, object recognition, scene recognition
  ○ Image captioning
	

• Self-driving cars and automotive safety
  ○ Scene understanding and recognizing all elements of a scene (pedestrians, cars, bicycles, road, curb)
  ○ Road sign recognition
  ○ Automatic braking
	

• 3D model building
  ○ 3D model reconstruction from stereo images/video
  ○ Automated construction of 3D models from aerial photographs
  ○ Live action and computer-generated imagery (CGI)
  ○ Camera viewpoint calculation for CGI
  ○ Augmented reality (AR): recognizing the environment to generate superimposed images
  ○ Virtual sports replay (reconstruction), real-time graphics augmentaion
	

• Computer Vision in Agriculture
  ○ Computer Vision Systems in Livestock Farming
  ○ Computer Vision Systems in Poultry Farming
  ○ Fish Farming with Computer Vision
  ○ Yield Estimation with Fruit or Vegetable Counting
  ○ Security Monitoring for Remote Farms
  ○ Achieve Compliance with Animal Welfare Law
  ○ Drone-Based Crop Monitoring with AI
  ○ Weed Detection and Management
  ○ Soil Health Assessment
	
	https://viso.ai/applications/computer-vision-in-agriculture/

• Robotic applications
  ○ Robot vision
  ○ Ecommerce product shipment
	

• Motion capture (mocap)
  ○ Capturing skeletal motion using reflective markers on an actor viewed from multiple cameras
  ○ Xbox: Kinect, full body tracking of skeleton, gesture recognition
	

Conferences in the field:
CVPR (Conference on Computer Vision and Pattern Recognition)
ECCV (European Conference on Computer Vision)
ICCV (International Conference on Computer Vision)
SIGGRAPH (ACM annual conference; Special Interest Group on Computer GRAPHics and Interactive Techniques)

Stages of a typical computer vision system

Image acquisition	Read image sensors
Pre-processing	Re-sampling, noise reduction, contrast enhancement
Feature extraction	Find lines, edges, corners, blobs
Segmentation	Separation of a scene into background, objects etc.
Decision/Recognition	Making final decisions based on competing evidence

Scope of the course
Computer vision is a vast topic, has a long history, and is an active area of research. This course will provide an introduction to the fundamental concepts with a hands-on approach. In this course we will concentrate on algorithms for analyzing visual information and in particular recognizing objects, scenes, and activities.