Computer Vision

——Introduction

Dr. Wu Xiaojun

September 9, 2020

Outline

Time, Lecturers & Teaching assistant

- Wens. 18:30-20:15, Fri. 18:30-20:15.
- Location: T6207
- Lecturers
  - Dr. Xiaojun Wu
  - Email: wu.xj@hit.edu.cn
  - Mobile: 18682202093
  - Office: C305C

- Teaching Assistant:
  - Name: Haohao Qu
  - Mobile:13713521217
  - email:1063510670@qq.com
  - office:C305C

Textbook and References

1. Computer Vision—A modern approach, David A.Forsyth, Jean Ponce, Pearson Education.
2. References
   - Computer Vision: Algorithms and Applications, Richard Szeliski, Springe, 2012.
   - Multiple View Geometry in Computer Vision. Richard Hartley, Andrew Zisserman. Cambridge University Press (July 31, 2000).
   - Computer Vision, Sunde Ma, Zhengyou Zhang (In Chinese)
3. Internet—the richest source of information
   - Web link http://www.cs.cmu.edu/~cil/vision.html
   - Web link http://www.cs.cmu.edu/~cil/v-source.html
   - Web link http://homepages.inf.ed.ac.uk/rbf/CVonline/
   - Web link http://www.eecs.berkeley.edu/Research/Projects/CS/vision/
     Web link http://mi.eng.cam.ac.uk/research/vision/

Class Objectives

1. Mastering the fundamental theory and techniques on CV.
2. Independent Research, Presentation & Collaborative capabilities.
3. Lectures emphasize on knowledge and fundamentals.
4. Training hands on skills.

Prerequisites

- Prerequisites
  - Linear Algebra (LA)
  - Programming skills----C , C++, etc.
  - Data structure and algorithms

- Submit the homework or code to TA.
- Get the lecture notes from TA.

Grading

- Homework (Exercises and programming)
    - Homework, class attendance and performance 30%
- Examination
    - Final exam. 30%
- Course project
    - Final project 40% (oral presentation, report)

Details

- Attendance and class performance

  TA or I will keep a roster to take a roll call at the beginning or after of class

  Keep track of who is involved with the questions and discussions

  Power off your mobile phone and keep good class order

- Homework

  Assigned homework according to content of class will be collected by the TA the following week Should cover the basics, and encourage research

- Participation and presentation in class

  Most important for graduate students

  Everyone should actively participate

- Projects

  One final project after one weeks, (after drop period) everyone should choose one partner to be a group, and report your members to the TA. In principle, each team has two members.

Syllabus

- Introduction
- Camera, projection
- Camera calibration
- Lighting and shading
- Image filtering
- Edge detection
- Image segmentation
- Color image and processing
- Features and feature detection
- Eigenface and PCA
- Optical Flow
- Stereo vision
- Structure from motion
- Final Exam. and projects

Summarize

1. Should not be hard to pass
2. Can take a lot of time
    - If you are interested in the subject For
    - research topics and projects
3. Lecture participation
    - intelligent questions and participate in class discussions can improve your attendance scores

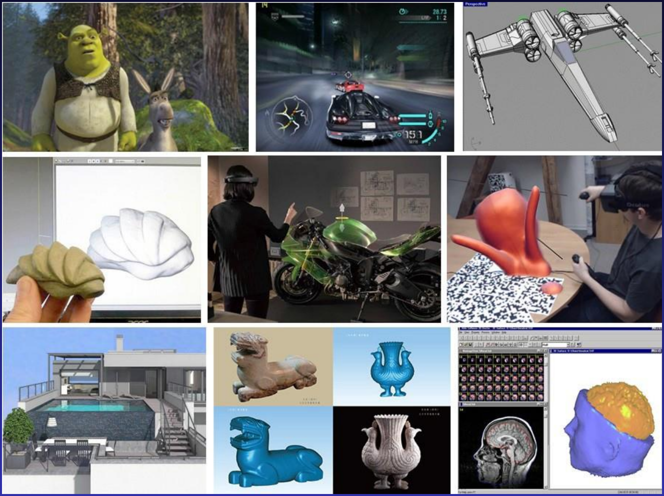
- Goal
    - Understand the fundamentals
    - Can locate resources on the Internet, and understand the cutting edge of the current academic research topics
    - Acquire some programming skills to move into further research and projects, and to be able to make use in real practice
    - Collect some useful tools for future research work

Previous Research Focus

- Surface reconstruction from cloud points.
- Image based 3D modeling.
- Machine vision algorithms and system.

Previous Research Focus

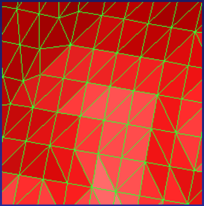
1. 3D models.



Previous research focus

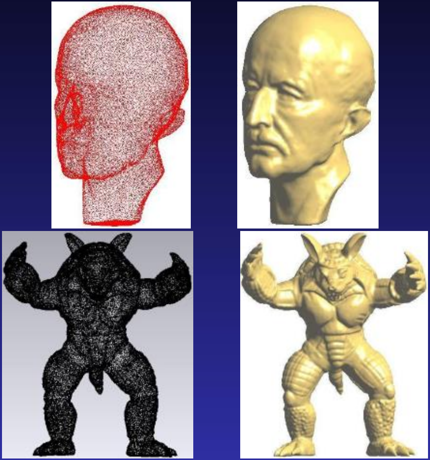
1. 3D models.





Previous research focus

1. Surface reconstruction from cloud points.

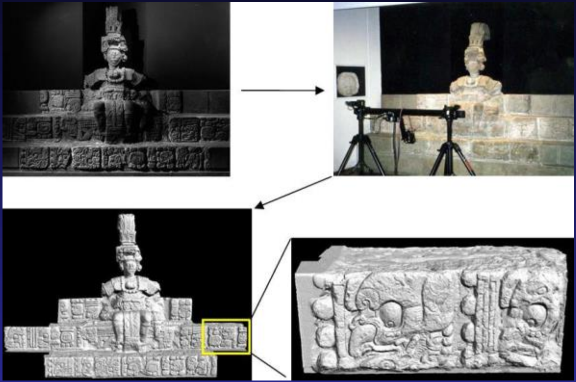


Previous research focus

1. Applications.

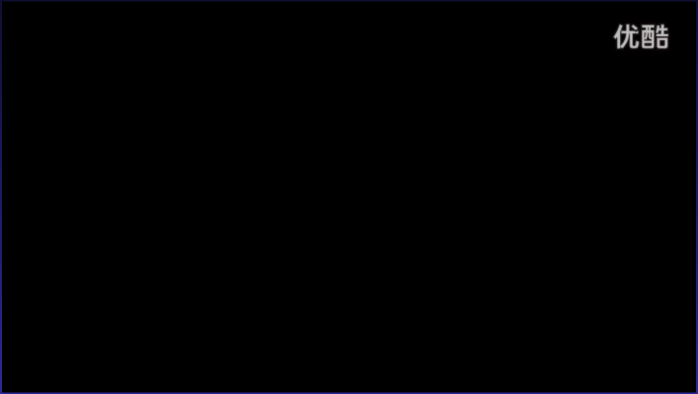
Previous research focus

1. Applications.



Previous research focus

1. Image based 3D modeling



Previous research focus

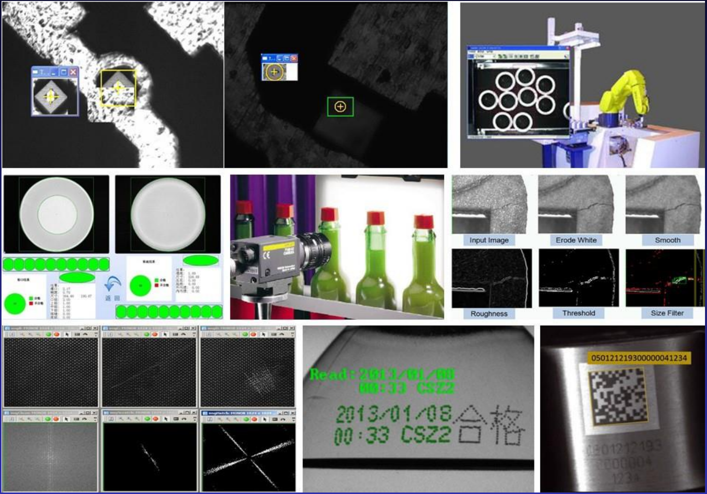
1. Image based 3D modeling.

Previous research focus

- Image based 3D modeling—more examples.

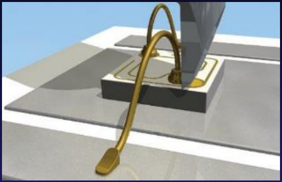
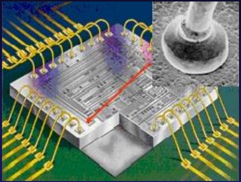
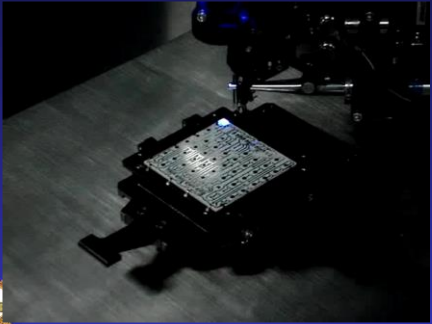
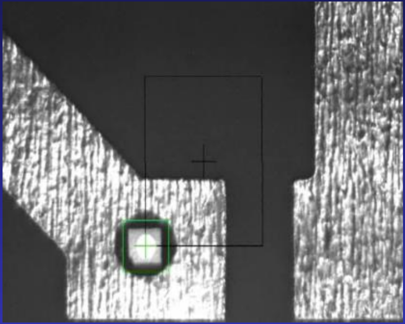
Previous research focus

1. Machine vision algorithms and system.



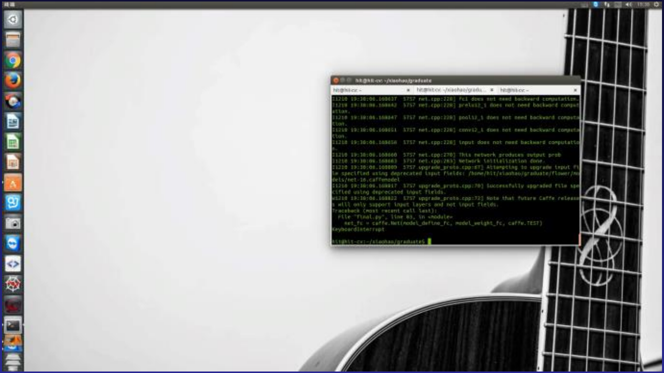
Previous research focus

Demo of machine vision systems

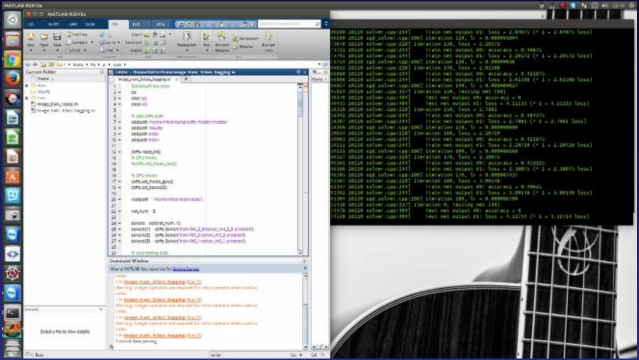
Previous research focus

Deep learning based object localization in low quality images.



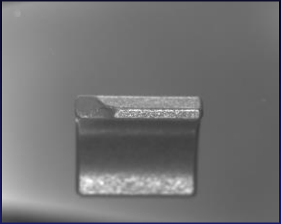
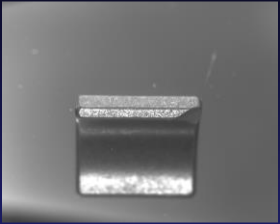
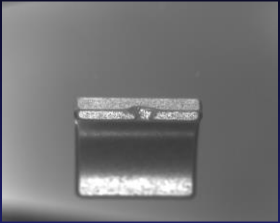
Previous research focus

Deep learning based OCR in low quality images.



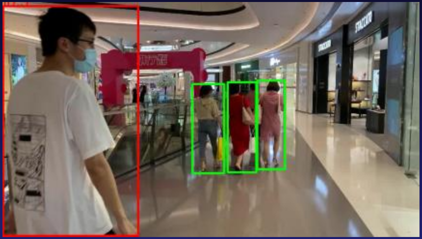
Previous research focus

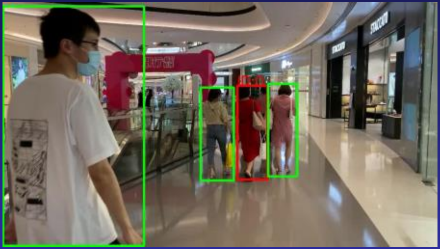
Deep learning based Surface Inspectiom.


Previous research focus

Deep learning based Pedestrian Detection and Tracking





Why study CV?

1. Computer vision's great trick is extracting useful information of the world from pictures or sequences of pictures.Taking pictures is usally non-destructive, easy and (now) cheap.
2. Fast-growing collection of useful applications

Why study CV?

1. Cutting edge technologies

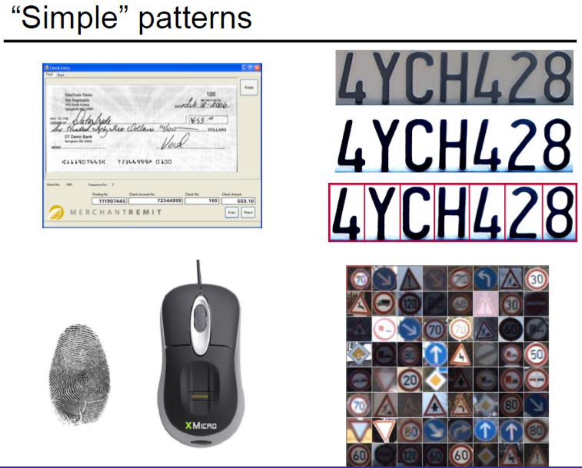
Self-service store

Why study CV?

- Vision is useful
- Vision is interesting
- Vision is difficult

Why study CV?

- Vision is useful



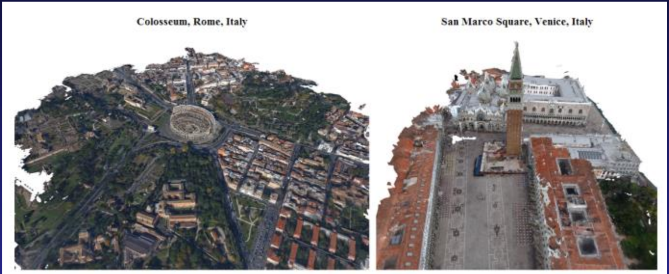
Why study CV?

- Vision is useful

Face++
Sensetimes
Tencent
Baidu
Hikivision

Why study CV?

- Vision is useful
- Reconstruction: 3D from photo collections



Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. Seitz, The Visual Turing Test for Scene Reconstruction, 3DV 2013

Why study CV?

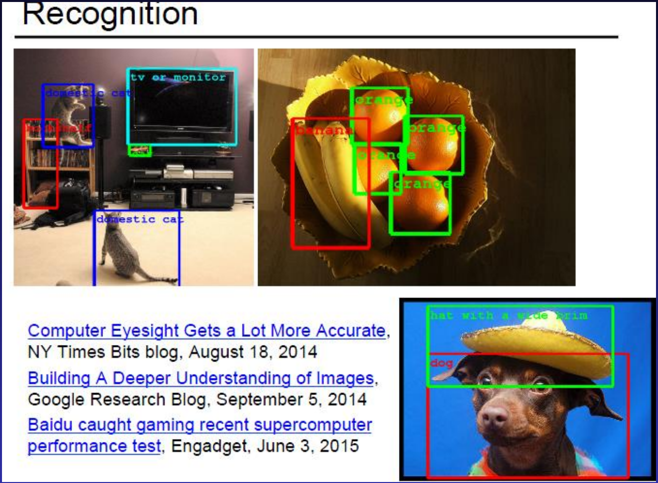
- Vision is useful
- Reconstruction: 4D from depth cameras

Figure 1: Real-time reconstructions of a moving scene with DynamicFusion; both the person and the camera are moving. The initially noisy and incomplete model is progressively denoised and completed over time (left to right).

R. Newcombe, D. Fox, and S. Seitz, DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time, CVPR 2015

Why study CV?

- Vision is useful



Why study CV?

- Vision is interesting
- Automatic age progression

I. Kemelmacher-Shlizerman, S. Suwajanakorn, and S. Seitz, Illumination-Aware Age Progression, CVPR 2014

Why study CV?

- Vision is interesting
- Digital puppetry

Figure 1. Model of Tom Hanks (bottom), derived from Internet Photos, is controlled by his own photos or videos of other celebrities (top). The Tom Hanks model captures his appearance and behavior, while mimicking the pose and expression of the controllers.

S. Suwajanakorn, S. Seitz, and I. Kemelmacher-Shlizerman, What Makes Tom Hanks Look Like Tom Hanks, ICCV 2015

Why study CV?

- Vision is useful and difficult

Self-driving cars

TECHNOLOGY

The New York Times

For Now, Self-Driving Cars Still Need Humans

By JOHN MARKOFF JAN. 17, 2016

http://www.nytimes.com/2016/01/18/technology/driverless-cars-limits-include-human-nature.html

Possible applications

From functionality of CV.

1. For inference and recognition
2. For measurement

From application domains of CV.

1. Industrial inspection
2. Image and movie editing software
3. Surveillance and automatic monitoring
4. Unmanned vehicles
5. Robotics
6. Medical imaging
7. Character, facial recognition
8. 3D model reconstruction

Application O—Machine vision

1. Object localization
2. Working with robot
3. Precise measurement
4. Inspection

Application I—Vehicle tracking

1. Edge detection
2. Change detection and motion detection
3. Bounding box approximation
4. Correspondence

Application II—Facial recognition

1. Facial recognition and tracking
2. www.cs.nott.ac.uk/~lls

Application III—3D reconstruction

1. From multiple views
2. From shading
3. From structural models
4. Rigid geometry

Application III—3D reconstruction

1. Scene reconstruction from single image—CMU.

Application —Large Scale Reconstruction From Internet Image Collections

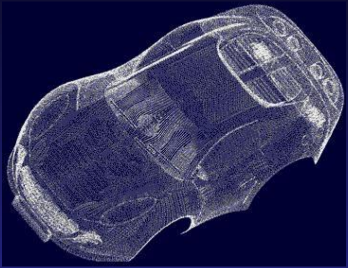
1. Build Rome in a day.

Application III—3D reconstruction

1. Laser scanner.

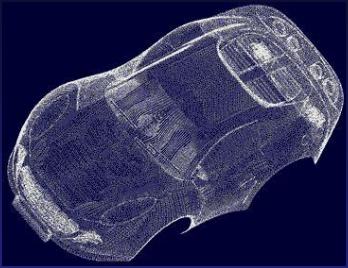
Application III—3D reconstruction

1. 3D reconstruction from cloud points.

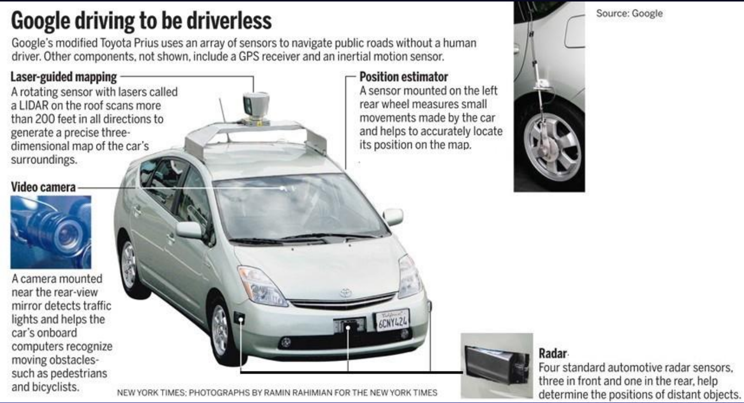


Application III—3D reconstruction

1. 3D reconstruction from cloud points.



Application IV—Automotive Machine Vision



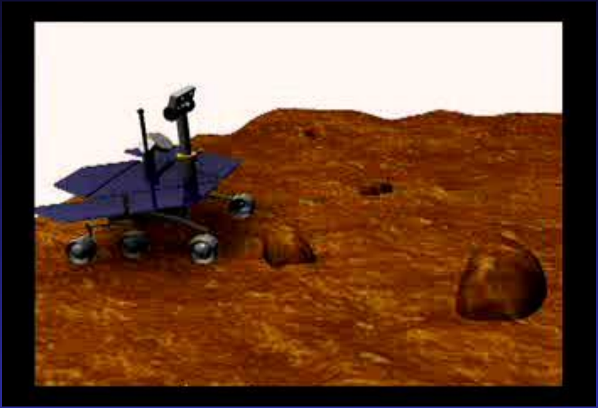
1. Google's driverless car uses vision technology (LIDAR* and video camera), radar and position sensors.
2. Driverless cars are now legal in Nevada, Florida and California.

Application IV—unmanned vehicle

1. Mar rover.

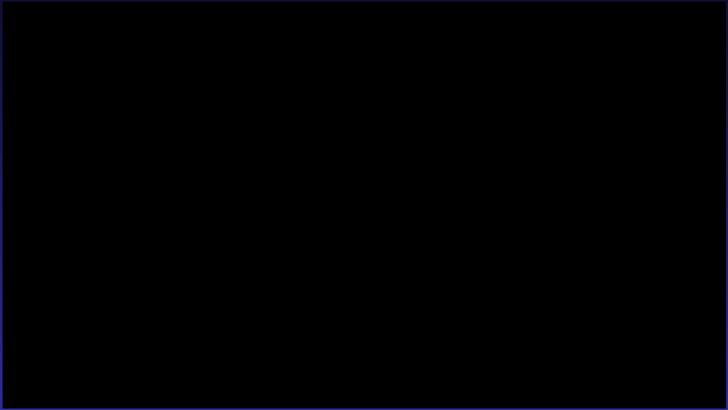
Application IV—unmanned vehicle

1. Mar rover.



Application V — Motion Capture

1. Performance Capture.



Application V — Motion Capture

1. Motion Capture.

Application VI—Motion Estimation

1. Motion Estimation.

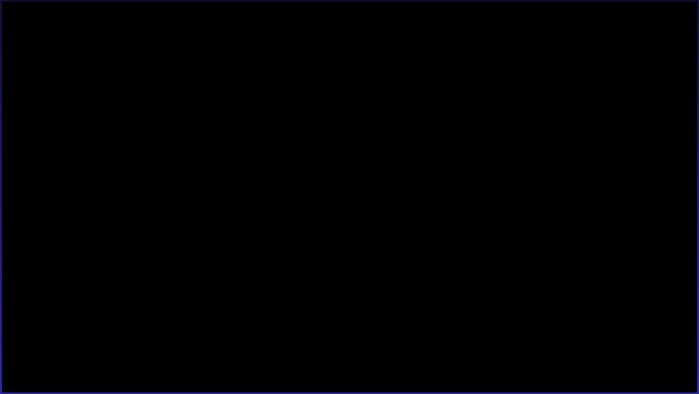
Application VII—Kinect

With the Kinect Sensor, put down the controller and play Xbox One games using just your body, voice, and gestures. Command your TV and even make Skype.

Application 1 of kinect.

Application VII—Kinect

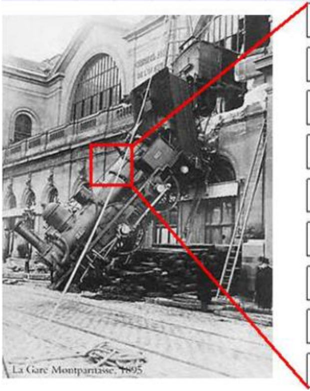
Application 2 of kinect.



Application VII—Kinect

The secret behind the Kinect.

Why is Computer Vision Difficult?



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 we see

What a computer sees

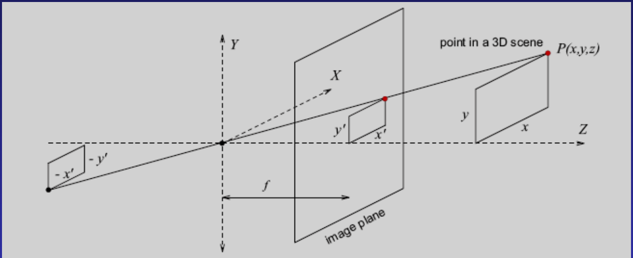
Srinivasa Narasimhan's Slide

Why is Computer Vision Difficult?

(1) It is a many-to-one mapping

   (1) A variety of surfaces having different material and geometrical properties, possibly under different lighting conditions, could lead to similar images.

   (2) Inverse mapping has non-unique solution; a lot of information is lost in the transformation from the 3D world to the 2D image.



Why is Computer Vision Difficult?

(2) It is computationally intensive

(1) - A typical video is 30 frames / sec

Main Challenges

(1) Viewpoint variations
(2) Illumination changes
(3) Scale changes
(4) Deformation
(5) Occlusions
(6) Background clutter
(7) Motion
(8) Intra/Inter-class variations

Why is Computer Vision Difficult?

Main Challenges

Viewpoint variations

Illumination changes

Why is Computer Vision Difficult?

Main Challenges

Scale changes

Deformations

Why is Computer Vision Difficult?

Main Challenges

Object intra-class variation

Why is Computer Vision Difficult?

Main Challenges

Occlusions


Background clutter

Why is Computer Vision Difficult?

Main Challenges

Motion blur

Ambiguity

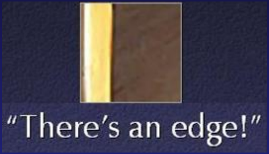
Steps of CV

1. Image processing
   - Single image or image sequence
   - Filtering
   - Edge detection
   - Segmentation
2. Clue finding
   - Multiple images, stereo
   - Bounding boxes, contours and junctions
   - Features and textures
   - Correspondence
   - Template Matching
   - Using properties of geometry models, rigid bodies
3. Action

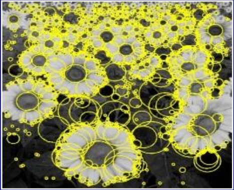
   Isolation, tracking, searching, prediction, avoidance, targeting, maneuvering

Marr's Framework of CV

1 Low-level or "early" vision —Considers local properties of an image



Edge detection

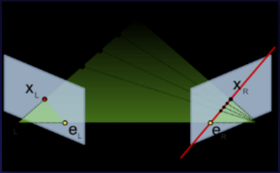


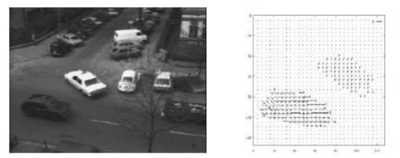
Corner and blob detection

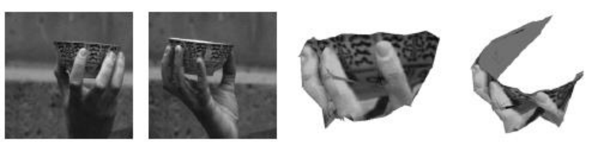
Marr's Framework of CV

2 Mid-level vision — Grouping and segmentation


"There's an object and a background!"

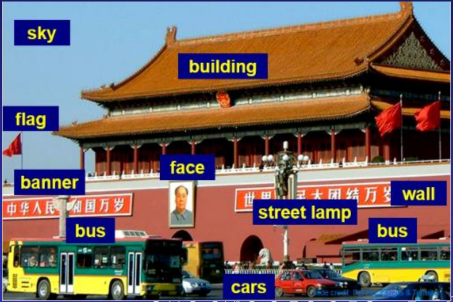






Marr's Framework of CV

3 High-level vision — Recognition



Marr’s Framework of CV

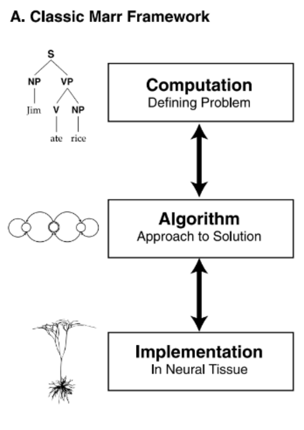


Fig. 1. Marr’s Three Levels Panel A), contrasted

❖信息处理有三个研究层次(three layers)

➢第一个层次是信息处理的计算理论(theory)

✓研究的是对什么信息进行计算和为什么要进行这些计算

➢第二个层次是算法(algorithm)

✓如何进行所要求的计算，也就是要设计特定的算法;

➢第三个层次是实现算法的机制或(implementation)

✓研究完成某一特定算法的计算机构。

Some related terms

- Image processing: the study of the operators that produce images from other images.
- Machine vision: main refer to industrial vision applications where (usually) a single camera is used to solve a structured inspection task.
- Pattern recognition: typically refers to the recognition of structures in 2D images (usually without reference to any underlying 3D information.
- Photogrammetry: the science of measurement through non-contact sensing, e.g. terrain maps from satellite images, 3D measurement by laser scanner. Usually it is more focused on accuracy issues than interpolation.

Computer vision

1. Computer vision
   - Vision: Latin, vi sio, to see something
2. Computer is far from being able to ”see” like human—the trick is extracting meaningful descriptions of the world from pictures or sequences of pictures.
   - Perceive, register in sensors
   - Distinguish, segment our features
   - Motion tracking, detect changes
   - Recognize and understanding
   - Prediction
3. Definition: allow computer devices to extract useful information of the surroundings from received signals

Computer Vision vs. Graphics

1. Computer Graphics
   - Produce "plausible" images
   - You choose the models, conditions, image parameters, etc physically or empirically.
2. Computer Vision
   - Given real images with noise, sampling artifacts . . . Estimate
   - physically quantities
   - Ill-posed — what is the minimum world knowledge we need?

Computer Vision vs. Image Processing

- Image Processing
  - Mostly concerned with image-to-image transformations

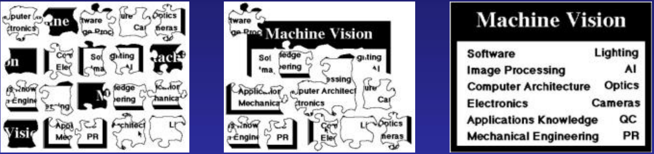
    Filtering

    Enhancement

    Comression
- Computer Vision
  - Wish to receive quantitative and qualitative information from visual data.
  - Much like the process of visual reasoning of human vision;
  - Distinguish between objects, classify them, sort them according to their size, and so forth.
  - Computer vision, like image processing, takes images as input. However, it returns another type of output, namely information on size, color, number, etc.
  - Image processing methods are harnessed for achieving tasks of computer vision.

Computer Vision vs. Machine Vision

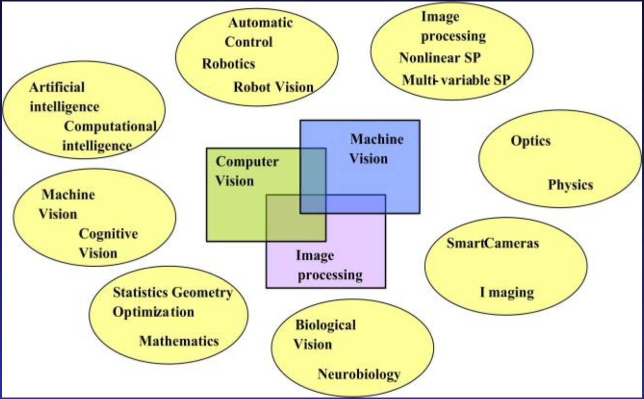
1. Machine vision is covered by the terms *computer vision* and *image processing*, but they are not synonymous. None is a subset of either of the others.
2. Computer vision is a branch of *computer science*, while machine vision is an area of specialisation within *System Engineering*.
3. Designing a machine vision system is like assembling a jigsaw.



4. Machine vision does not necessarily imply the use of a computer.

Sketch Illustration between CV and other Fields

- Relation Between Computer Vision and Various other Fields.



Still at early stage

- Most techniques are ad-hoc, highly experimental
- Difficult to segmentation, understanding, recovery from 2D image

  1. No one correct answer
  2. Heuristic processes

     Mathematical approximation Statistic and stochastic methods Pattern matching

     Neural networks

     Involves digital signal processing
     More Complex than voice, character recognition
     Finite set

- There is a lot can be done!
- In summery, CV is exciting and disorganised; there is often no reliable authority to appeal to.

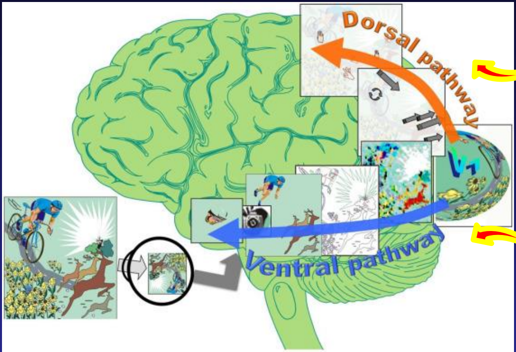
  1. many useful ideas have no theoretical grounding, and some theories are useless in practice.
  2. developed areas are widely scattered.

Deep Learning

Deep Learning is a new area of Machine Learning research, which has been introduced with the objective of moving Machine Learning closer to one of its original goals: Artificial Intelligence.

MV vs Deep Learning

北侧通路，空间的位移的通路，参与空间位置感知。



腹侧通路，主要功能是物体的识别。

MV vs Deep Learning

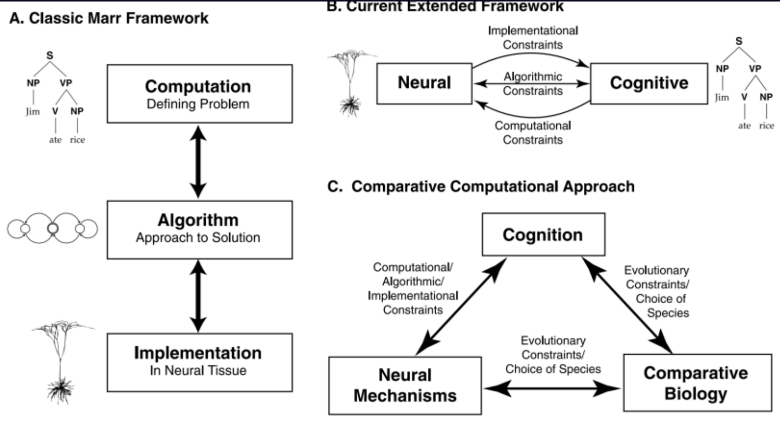
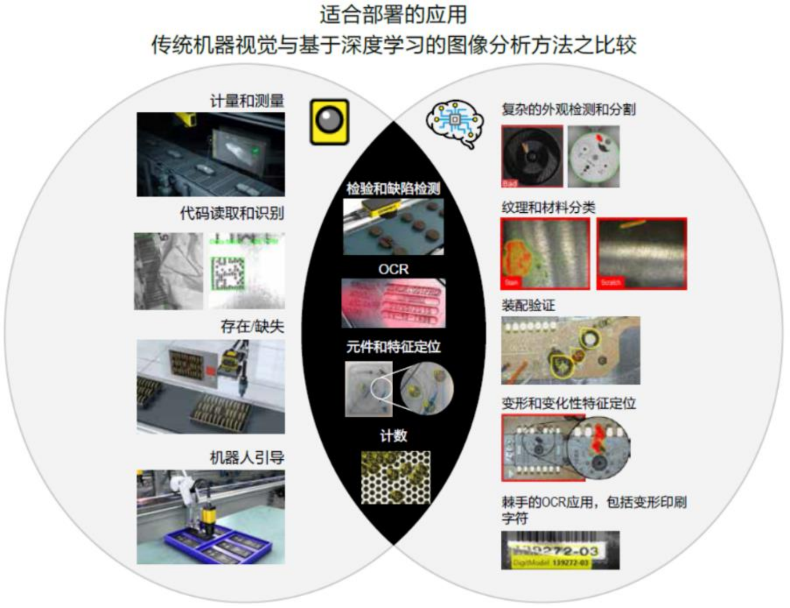


Fig. 1. Marr's Three Levels Panel A), contrasted with the approach developed in the current paper (B and C). See text for details.

MV vs Deep Learning



Summary

1. CV is useful.
2. CV is still in early age.
3. It will cost much effort to learn it well.

See You