Computer Vision

CS 600.361/600.461

Instructor: Greg Hager (hager@cs.jhu.edu)

https://cirl.lcsr.jhu.edu/Vision_Syllabus

Picture sources: slides from Nicolas Padoy, Fei-fei Li, S. Seitz, G. Hager; books from R. Szeliski, R. Harley, T. Moons

Outline

- Course organization
- What is computer vision ?
- Course overview

Class organization

Course information (1)

- Introductory course
  - Audience: undergraduate and graduate students
  - 600.361/600.461
- Required background
  - Programming fundamentals, in particular data structures and either Matlab or Python experience desirable
  - Linear algebra
  - Basic probability
  - Calculus
- Teaching assistant: TBD

Course information (2)

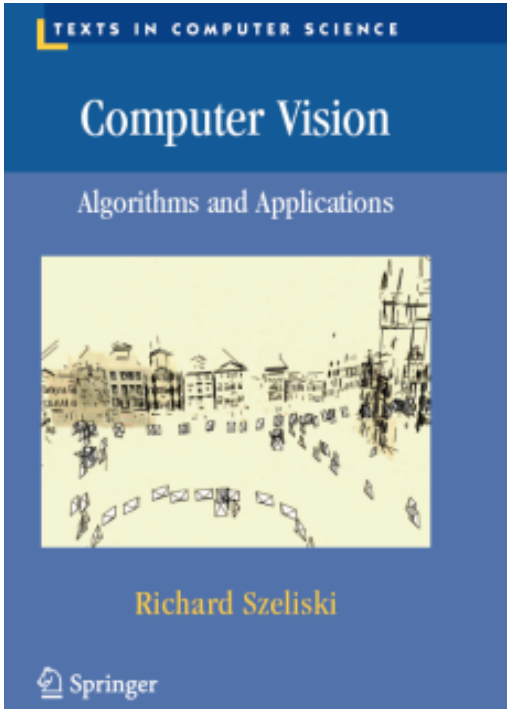
- Organization
  - 2 classes (1h15) per week
  - 4 homework, ~1 every two weeks (first half of the semester) [1/3 grade]
    - Programming assignment (Matlab)
    - Written assignment
  - TA sessions (programming, reminders, tips and assignment corrections)
  - 1 project (second half of the semester) [1/3 grade]
  - 1 midterm and final exam [1/3 grade]

- Office hours
  - By appointment
  - TA office hours = same as TA session = to be determined

Course information (3)

- Course updates:
  - Announcements/material: https://cirl.lcsr.jhu.edu/ANNOUNCEMENTS2012
  - Assignments: JHU blackboard [To confirm]

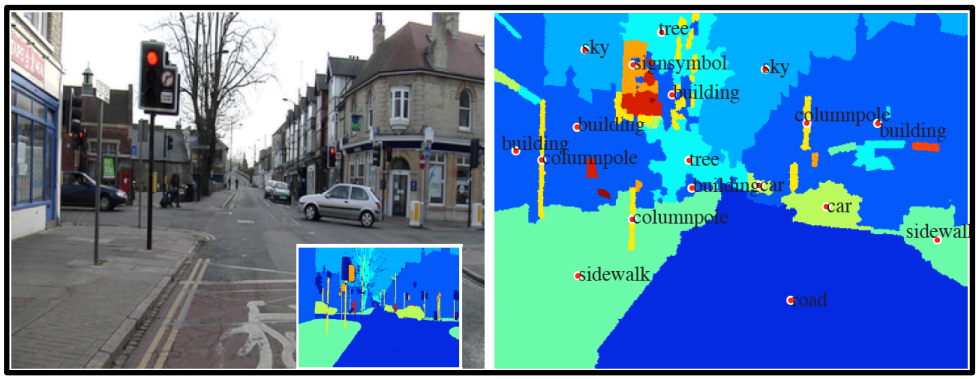
- References
  - *Computer Vision*, by R. Szeliski, Springer 2011 [online]
  - *Computer Vision: A Modern Approach*, Forsyth and Ponce, Prentice Hall, 2002
  - *Multiple View Geometry in Computer Vision*, Hartley and Zisserman, Cambridge University Press, 2010



What is Computer Vision ?

Computer vision ?

- Recovering properties from the real world using images
  - Examples:
    - detecting/recognizing objects in a scene
    - reconstructing a 3D model of the imaged buildings
- Overlaps with a few other areas from computer science, e.g.
  - Image processing
  - Machine learning / pattern recognition
  - Graphics



Micusik et al.

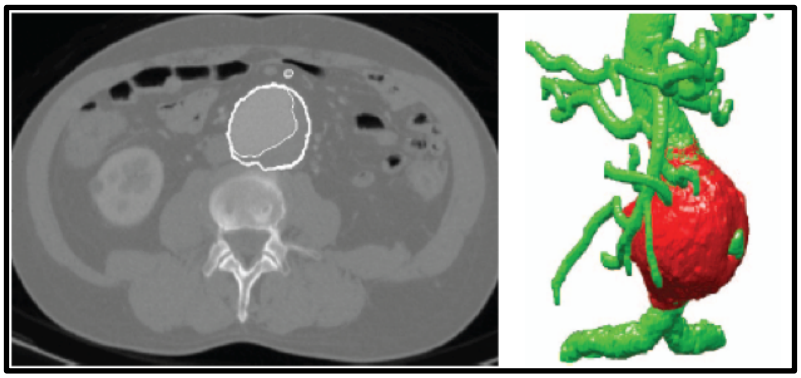
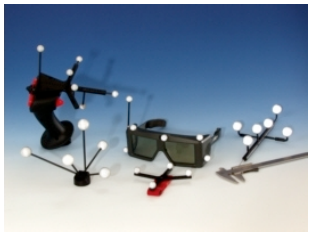

Demirci et al.

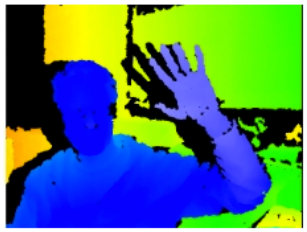
Image modalities (1)

Digital cameras



Infrared cameras

Time-of-flight cameras

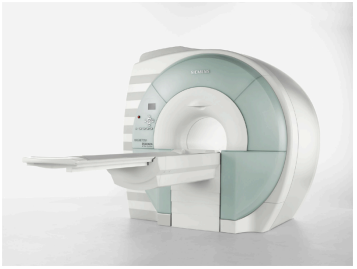


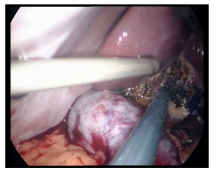
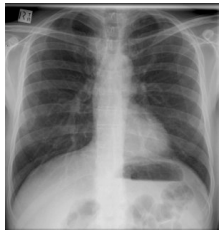
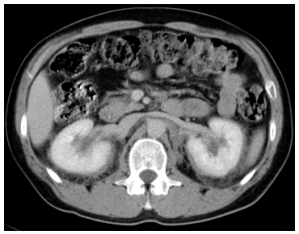
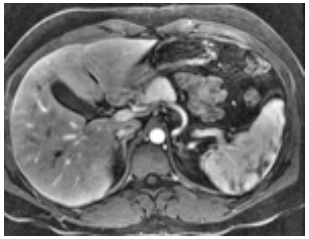
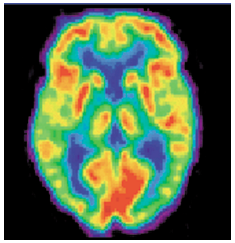
Kinect

Image modalities (2)







Ultrasound

Endoscopy

X-rays

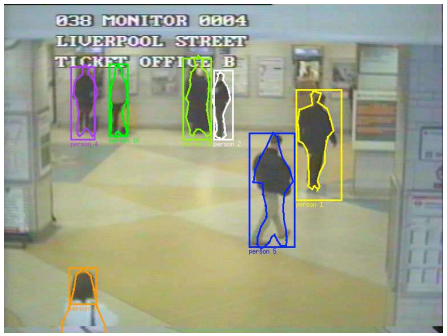
Computed tomography

Magnetic resonance

Positron emission tomography

Capsule endoscopy

Application examples



Visual surveillance (Siebel et al.)

Character recognition

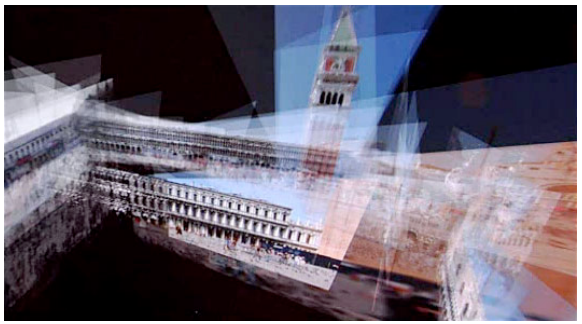
Image search

Sportvision

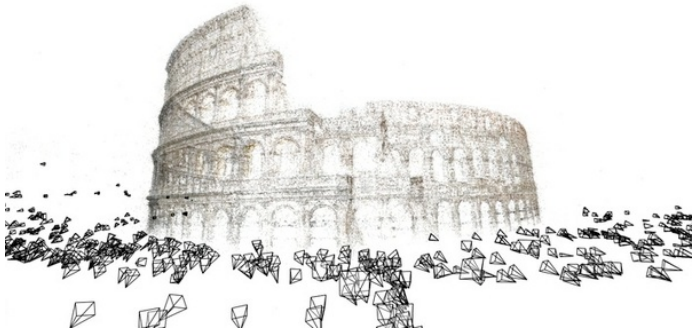
Fingerprint scanners

Robot control (NASA)

Application examples

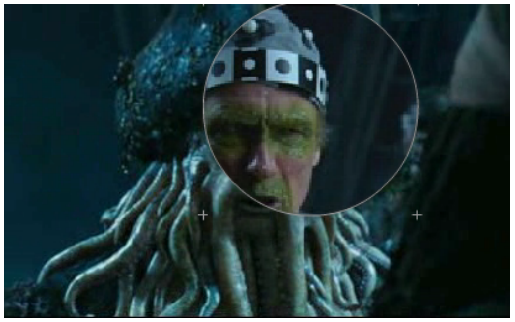


Photosynth



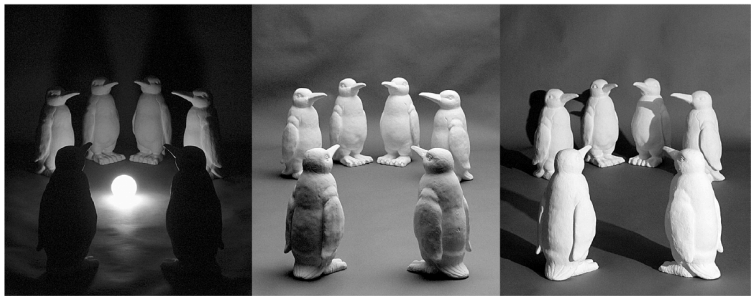
Multiple view reconstruction (Univ. Washington)

Medical augmented reality (MIT)



Special effects

Why is vision challenging…?



Illumination and viewpoint

Intra-class variation

Scale and depth

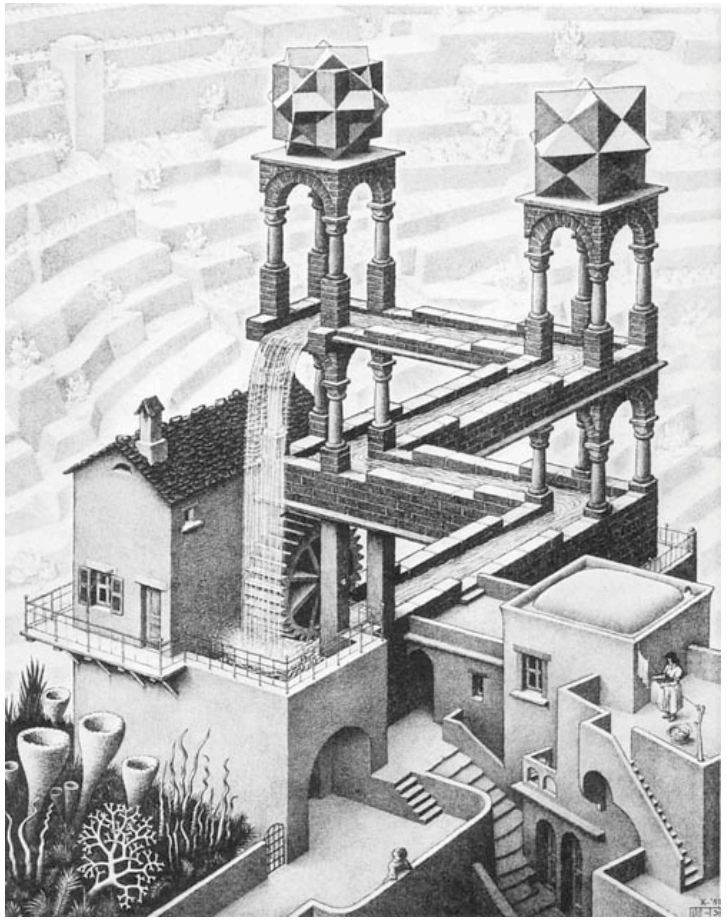
Why is vision challenging…?

Shadows

Clutter

Why is vision challenging…?

Occlusions



Illusions (Escher)

Why is vision challenging…?

Why is vision challenging…?

In short…

…additional knowledge/constraints required to remove ambiguities:

- A-priori information about the scene, context and/or the images
- Additional images from different viewpoints

A few videos…

Video: multi-view reconstruction

R. Szeliski et al. [Dubrovnik, Croatia]

Video: medical augmented reality

Traub et al.

Video: deformable surface tracking

Richa et al.

Video: real-time object detection

Hinterstoisser et al., CVPR 2010

Video: kinect object detection

Hinterstoisser et al., ICCV 2011

Video: kinect fusion

Microsoft, Siggraph 2011

Class Overview

Image processing

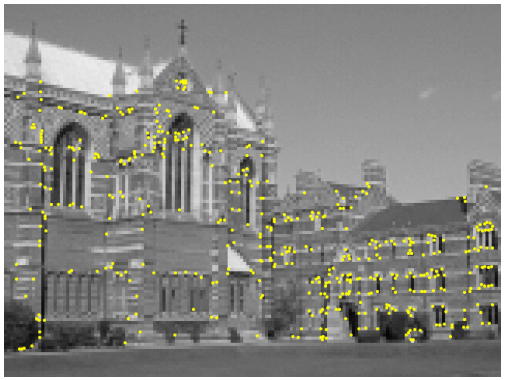
- Basic operations, linear filters, non-linear filters

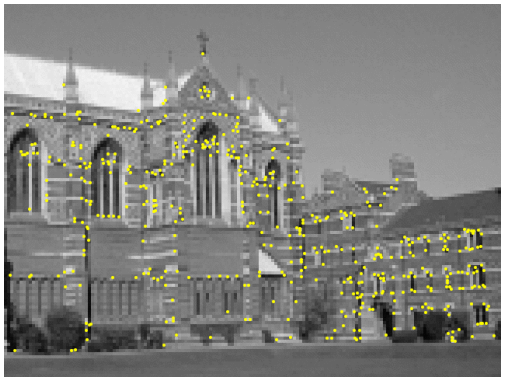


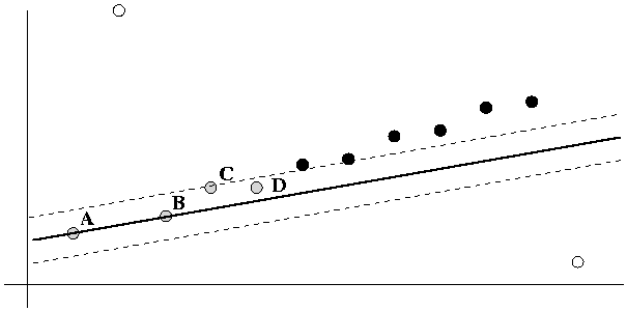



Image processing

- Edge detection, corner detection, line/circle detection/fitting

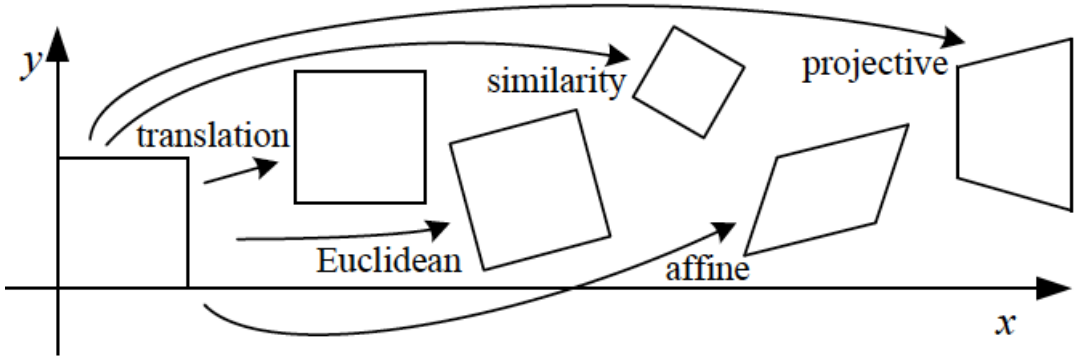






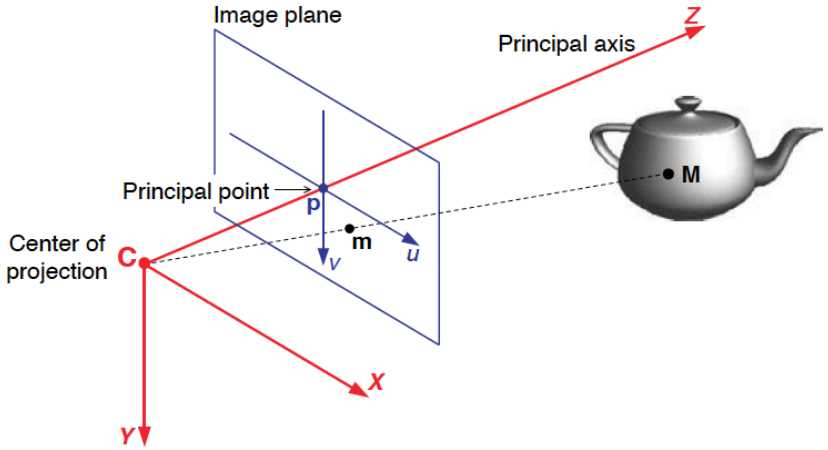
Single view geometry

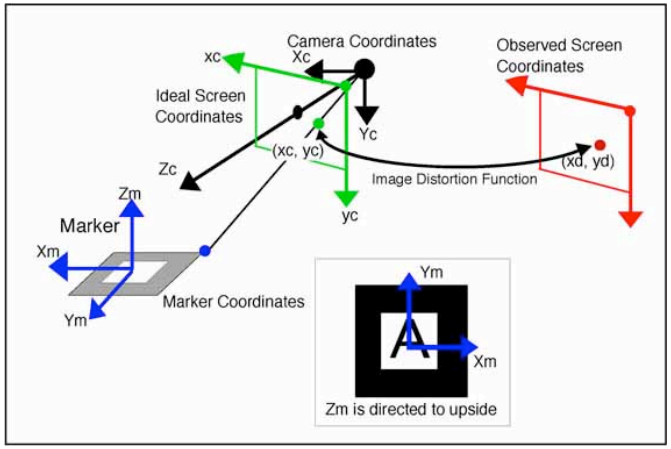
- 2D transformations, homographies, DLT, planar rectification

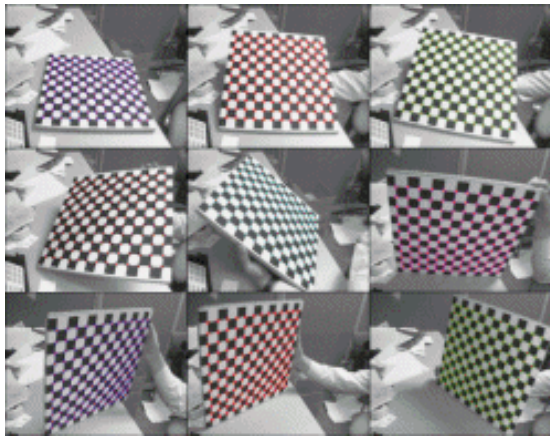


Single view geometry

- Camera models, calibration, pose estimation, mosaicing

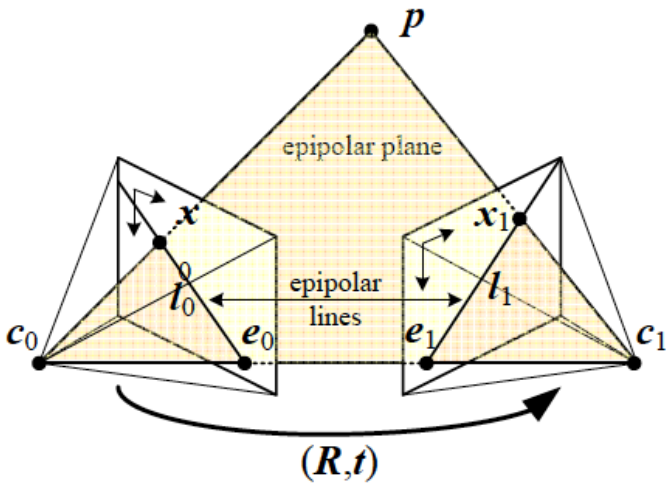


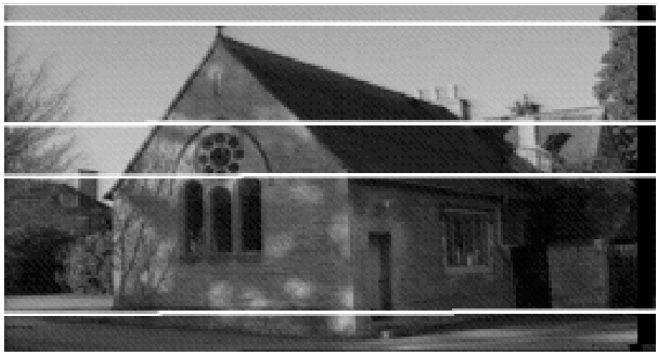


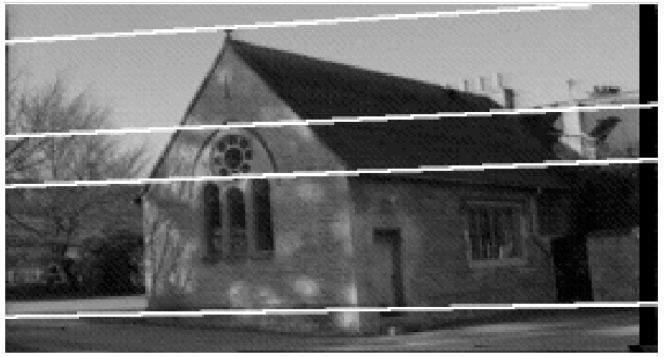


Two-view geometry

- Triangulation, stereo calibration, epipolar geometry

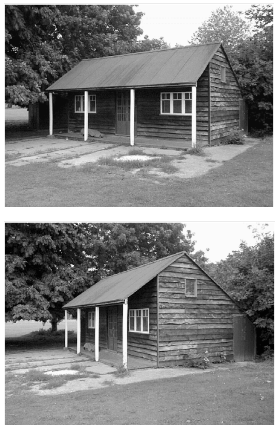
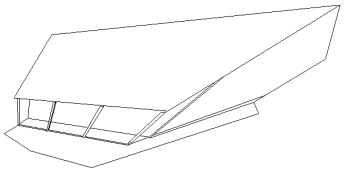
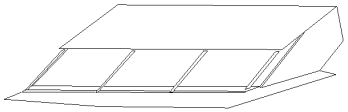
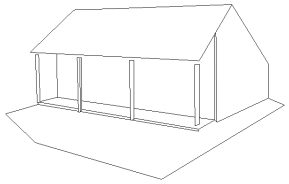
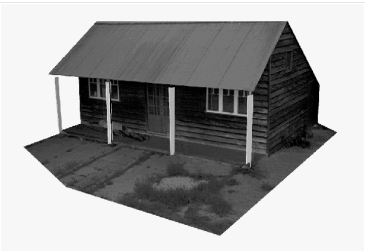




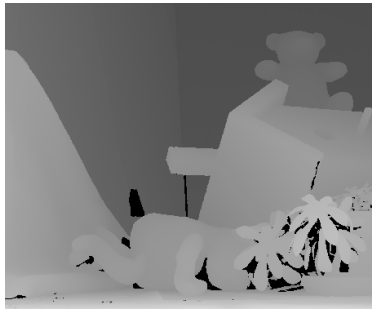
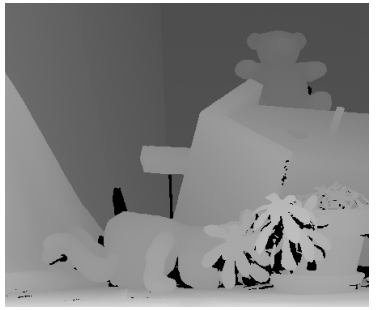


Two-view geometry

- Stratified (sparse) reconstruction, image rectification, dense reconstruction

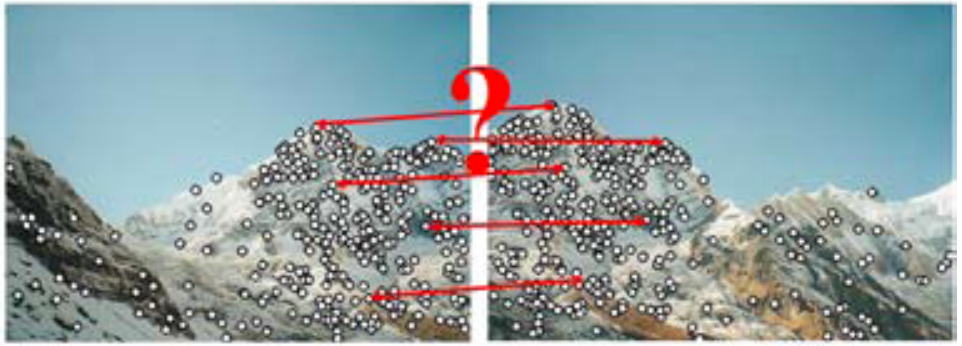






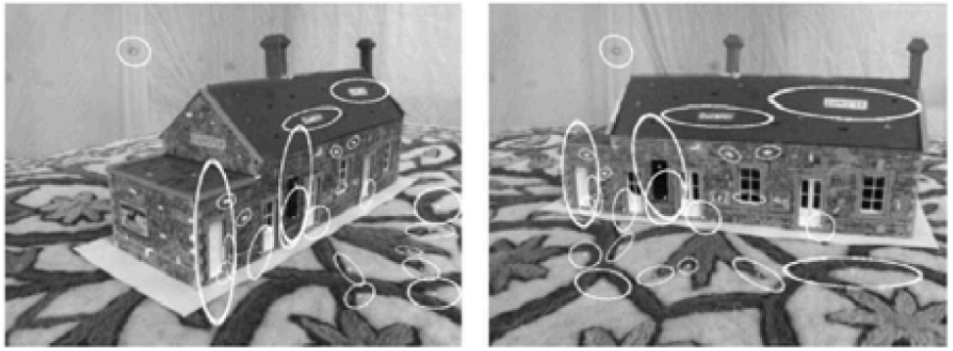


Feature descriptors

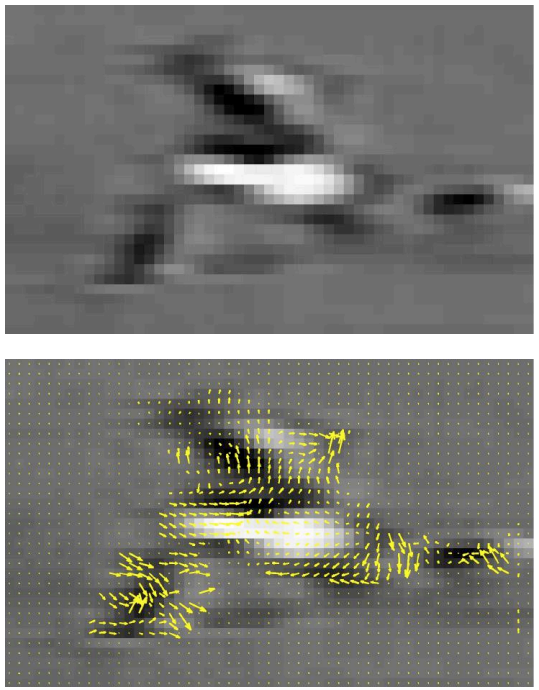
- Feature detection and matching, SIFT, SURF

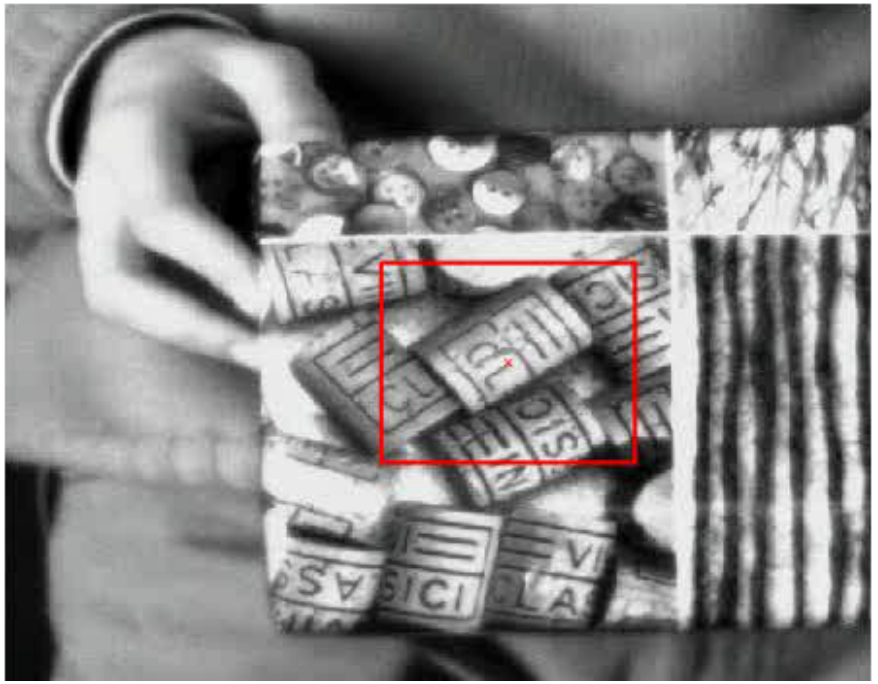




Motion & Tracking

- Optical flow, feature tracking, applications, ssd template tracking

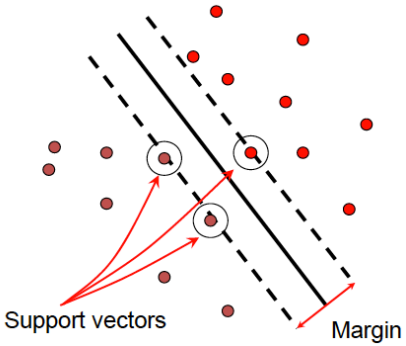


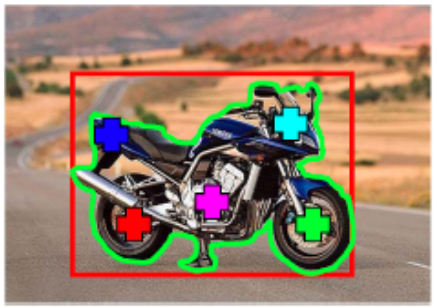


INRIA

Classification & Recognition

- Face and image classification, image features, classifiers, k-means, bag-of-words, object detection

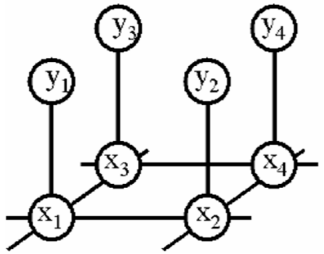


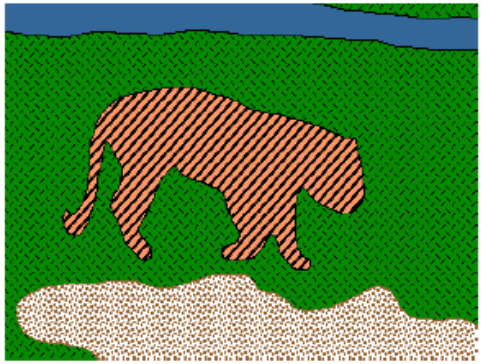
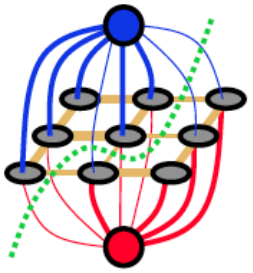


Segmentation & Clustering

- Split&merge, mean-shift clustering, MRF formulation, graph-cuts, grab-cut

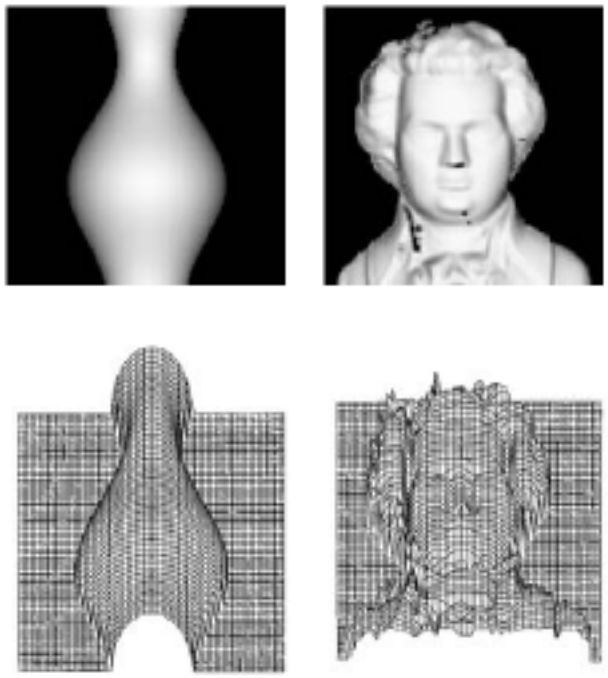


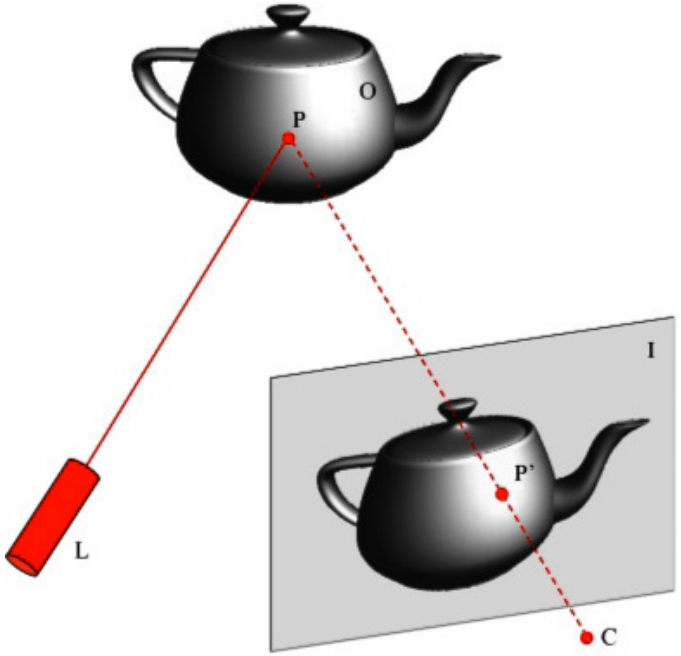


Vision III

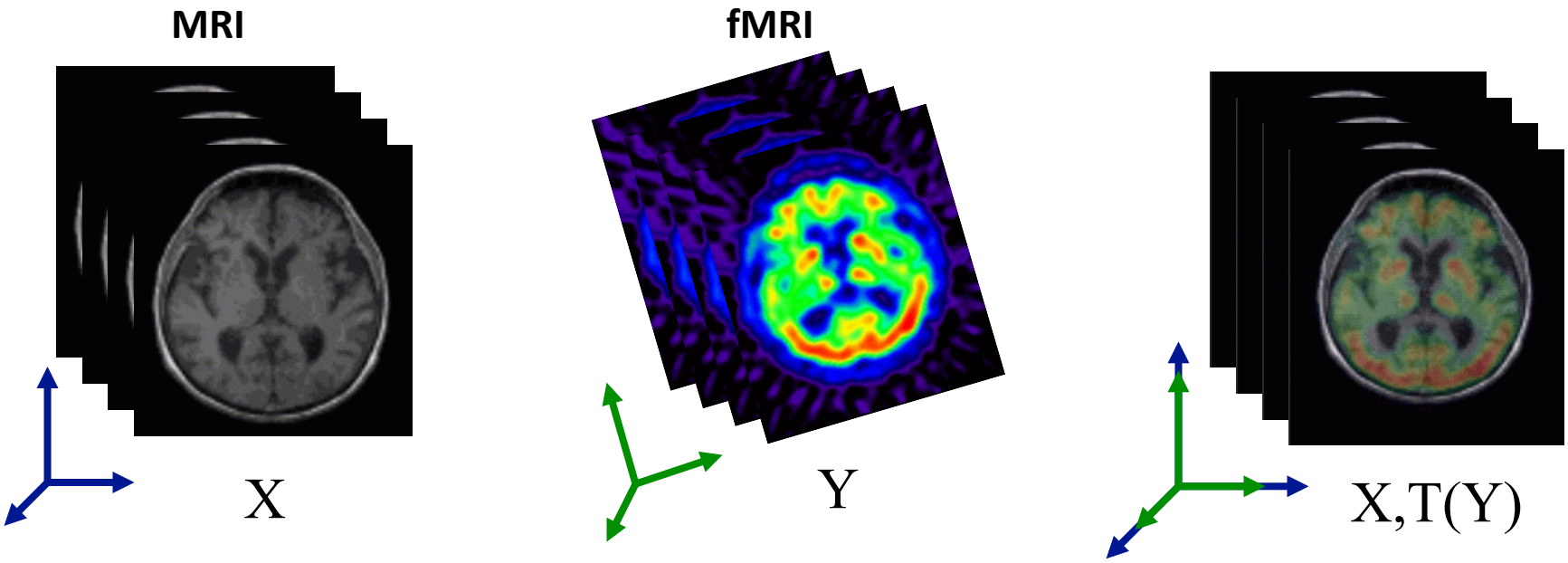
- Shape from shading, structured light, photometric stereo





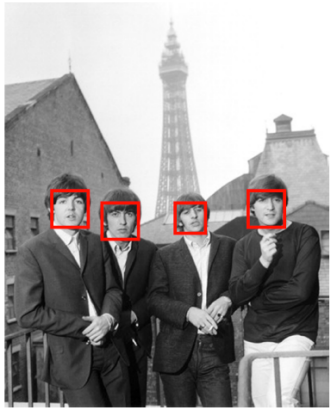
Applications

- Introduction to medical image registration



Applications

- Panoramas, face detection, 3D real-time detection of texture-less objects, deformable tracking



Does computer vision work ?
Live demos…

Programming resources

- Matlab (student version available)
- ARToolkit (University of Washington)
- Python tools (still getting organized)
- OpenCV library (c/c++)
  - Book:

    Learning OpenCV: Computer Vision with the OpenCV Library, G. Bradsky, A. Kaehler, O'Reilly 2008

What is next ?

- Images
- Basic image operations
- Linear and non-linear filtering

- Please check:
  - access to webpage
  - access to blackboard
  - access to Matlab
  - doodle poll for TA sessions