Computer Vision
CMSC 426

Fall 2019

Logistics

- Four projects, and three homework assignments (programming + discussion): in groups of three
- Midterm exam: in class
- All information available from the Website
- Grading
  - Projects 50 %
  - Homework 25%
  - MidTerm 25%

Programming

- Python
- Primer on Python?

Administration

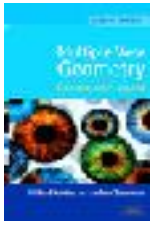
- Webpage
  - Get homework and projects
  - Syllabus
  - Other documents
- Piazza
  - Ask questions
    - do not post solutions
    - do not ask if your answer or approach is correct
  - Discuss issues
  - Public versus private
- ELMS
  - Hand in homework and projects
  - See grades

Recommended Texts

Computer Vision: Algorithms and Application, Richard Szeliski
Online: http://szeliski.org/Book/

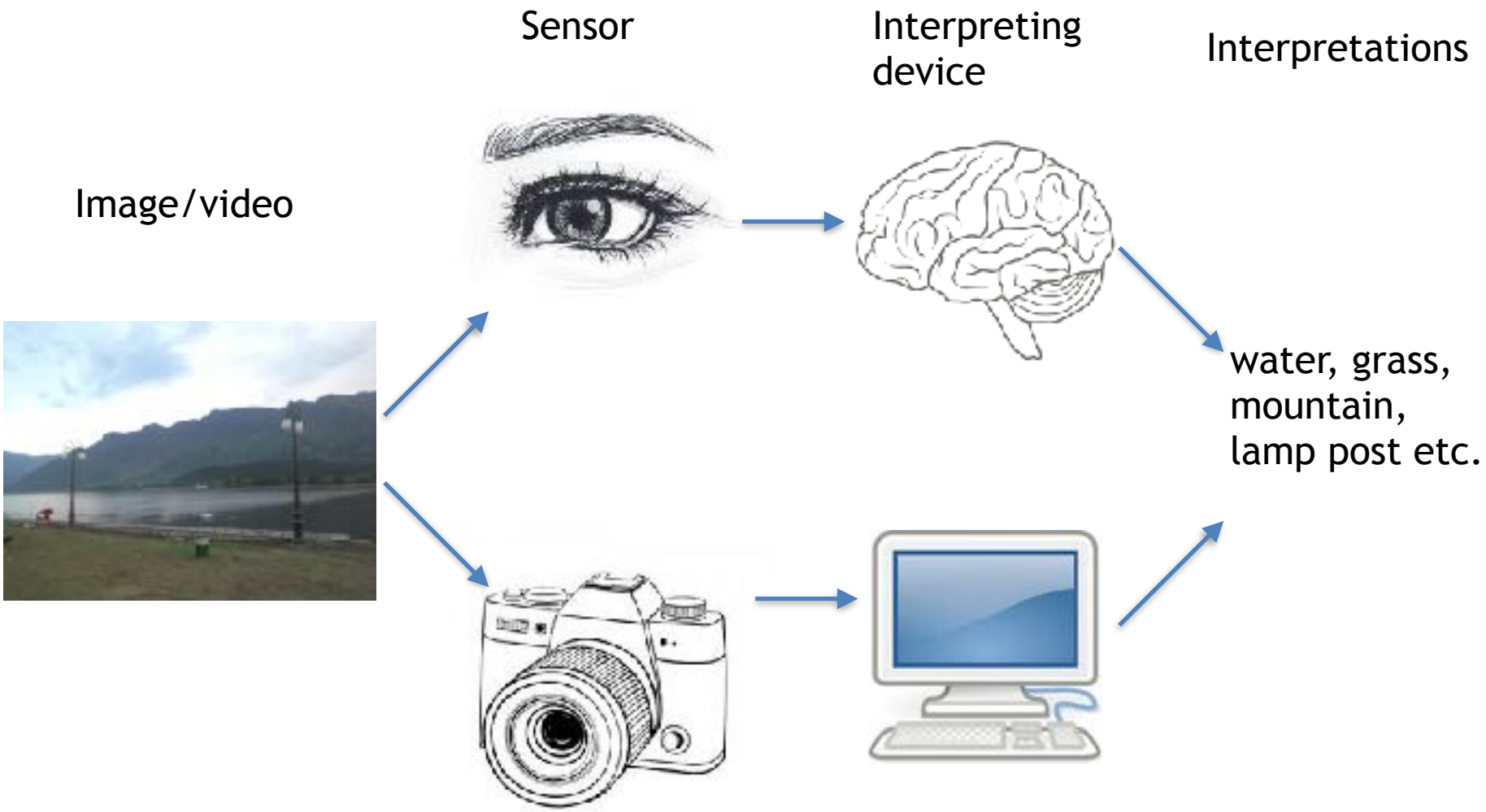
Computer Vision: A Modern Approach
David Forsyth and Jean Ponce
Online: http://www.csd.uwo.ca/~olga/Courses/Winter2010/CS4442_9542b/CVbook.pdf

Digital Image Processing, Prentice Hall, Rafael Gonzalez, and Richard Woods, 2008.
Online: http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/Digital_Image_Processing_2ndEd.pdf

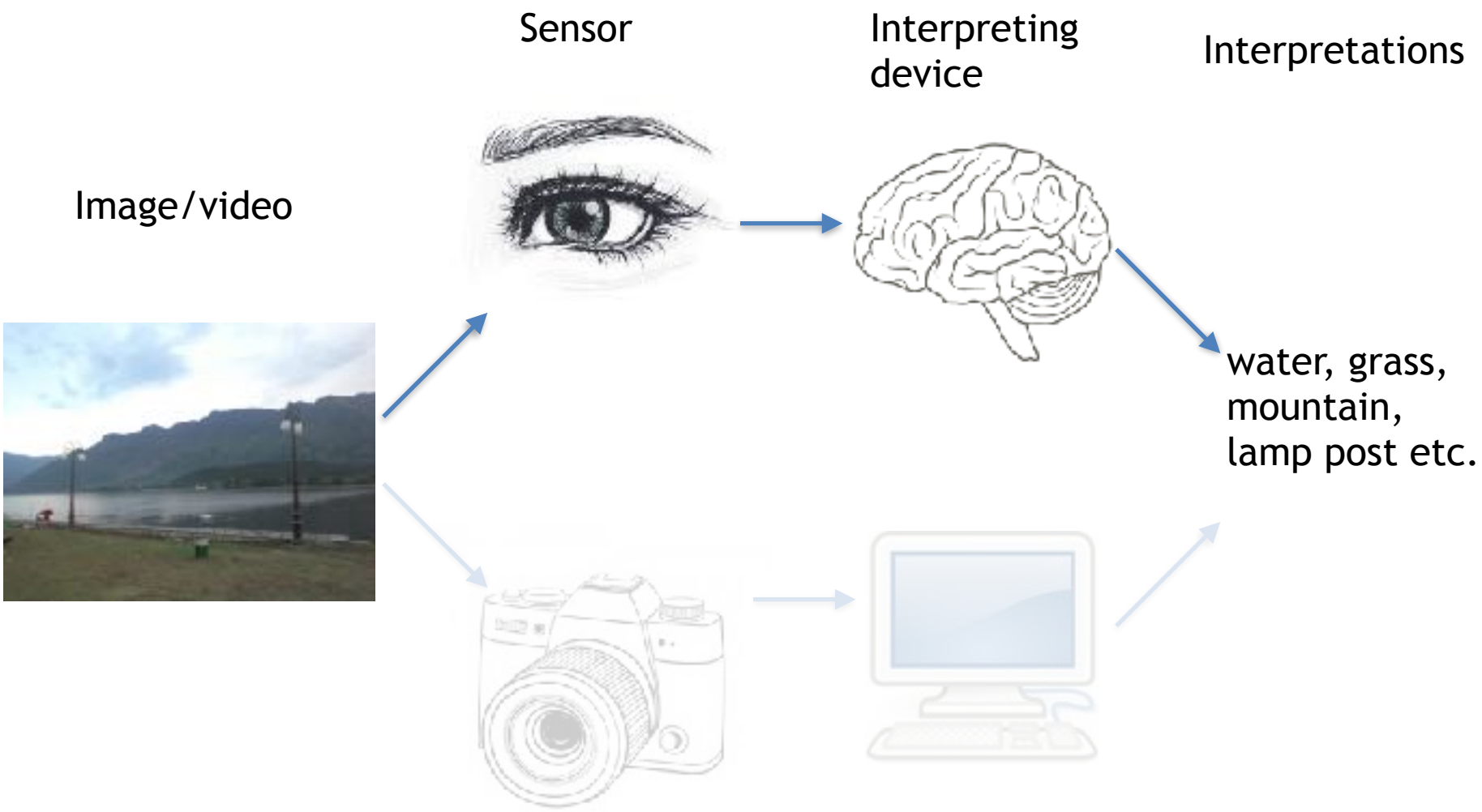


Multiple View Geometry in Computer Vision
Richard Hartley and Andrew Zisserman University Press, 2004,
selected chapters available online: http://www.robots.ox.ac.uk/~vgg/hzbook/

What is Computer Vision



What is Computer Vision



What is Computer Vision

Sensor

Interpreting device

Interpretations

Image/video

water, grass, mountain, lamp post etc.

Goals of Computer Vision

- Build machines and develop algorithms which can automatically replicate some functionalities of biological visual system

- Systems which navigate in cluttered environments
- Systems which can recognize objects, activities
- Systems which can interact with humans/world

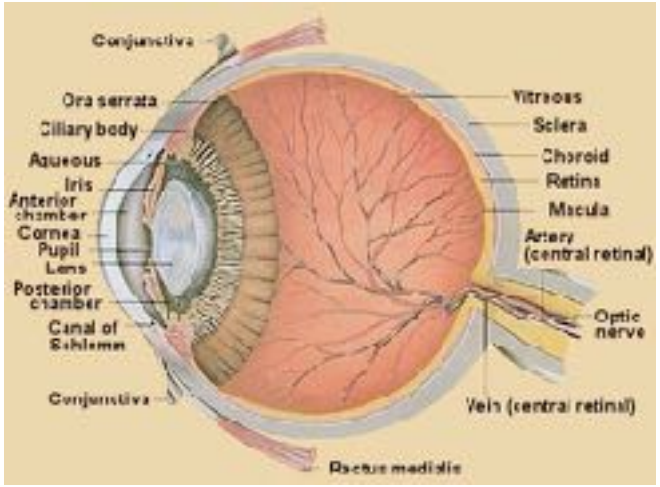
- Synergies with other disciplines and various applications Artificial Intelligence - robotics, natural language understanding

- Vision as a sensor - medical imaging, Geospatial Imaging, robotics, visual surveillance, inspection

Computer Vision

Visual Sensing

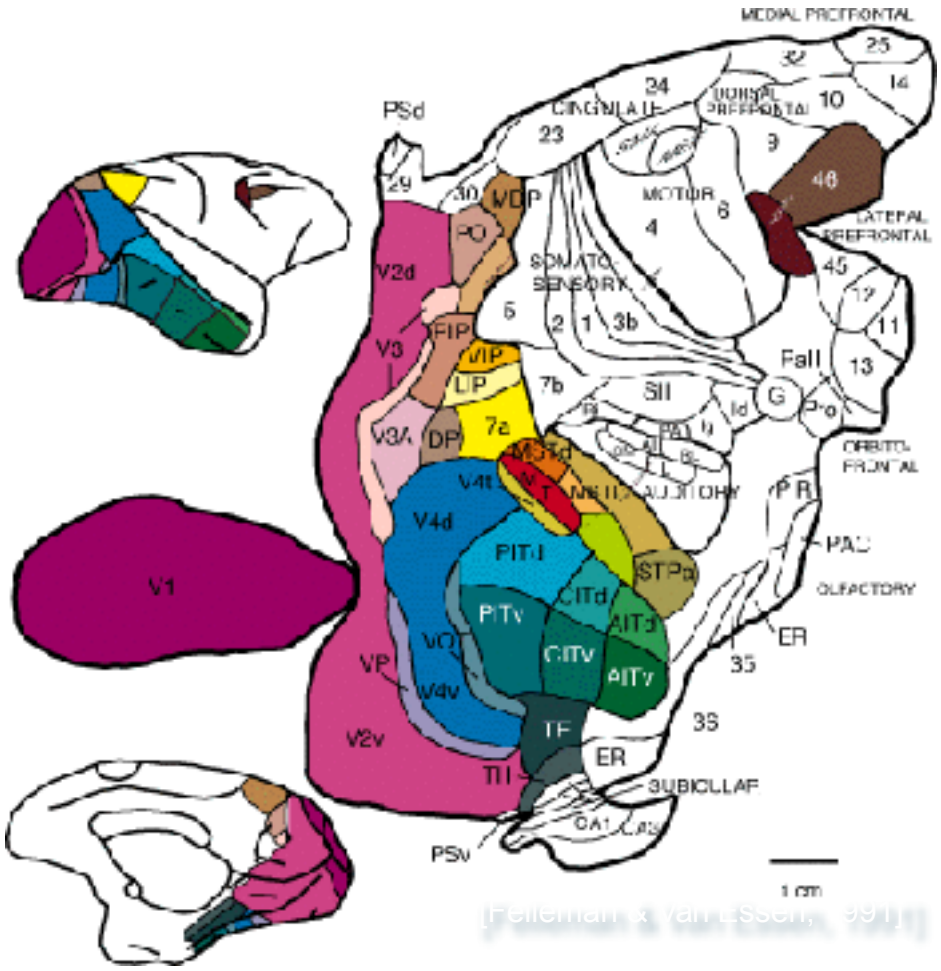
Images I(x,y) – brightness patterns



- image appearance depends on structure of the scene
- material and reflectance properties of the objects
- position and strength of light sources

Visual Information Processing

This is the part of your brain that processes visual information

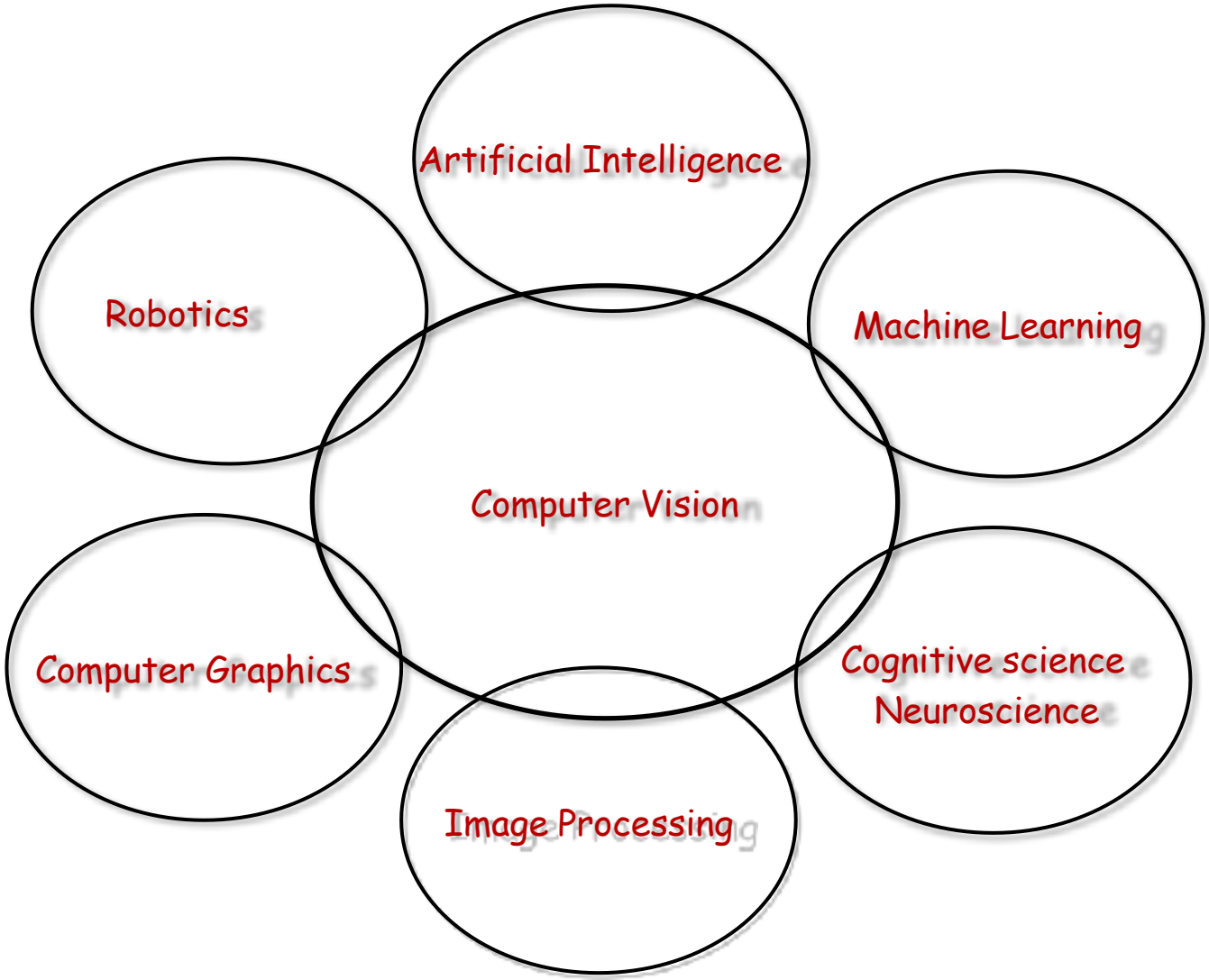


Challenges/Issues

- About 40% of our brain is devoted to vision
- We see immediately and can form and understand images instantly

- Applications and examples

Connections to other disciplines



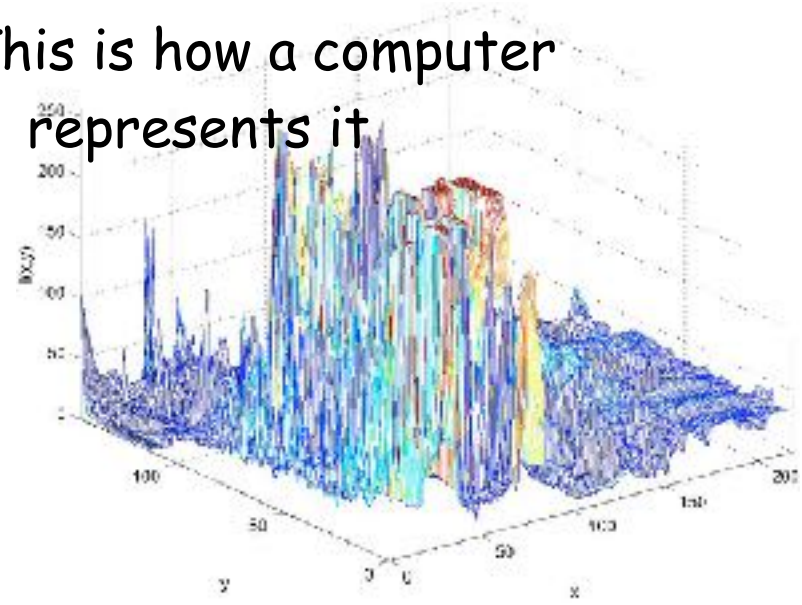
Goal of Computer Vision

what we see

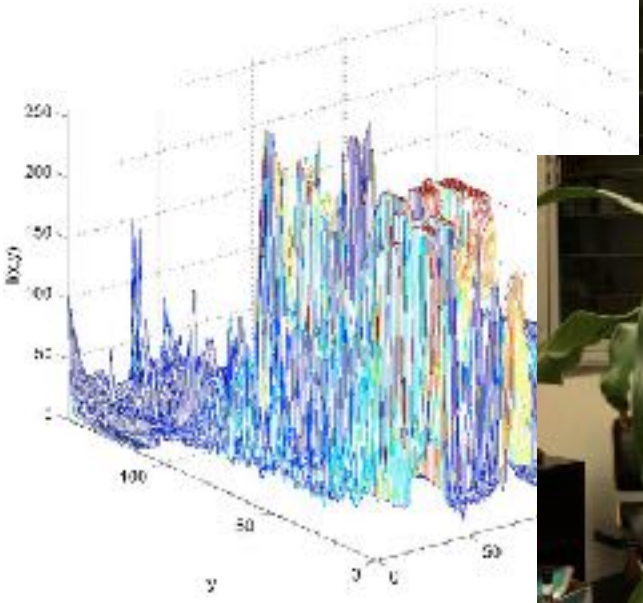
what computers see

230	25	34	123
45	0	10	52
65	11	210	42
78	87	56	90
23	18	29	61

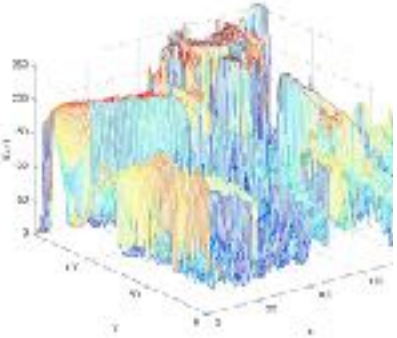
This is how a computer represents it

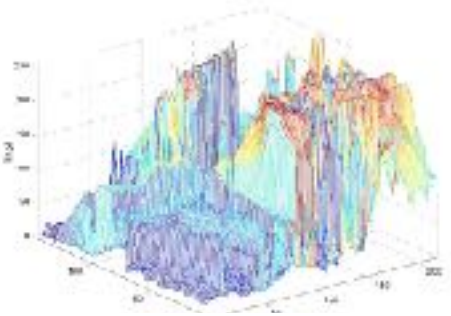


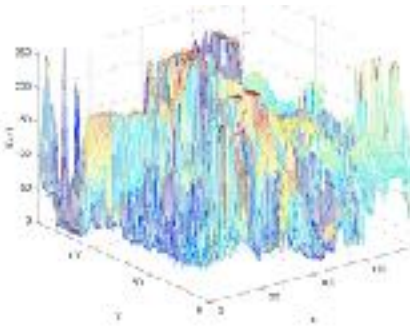
And so is this …

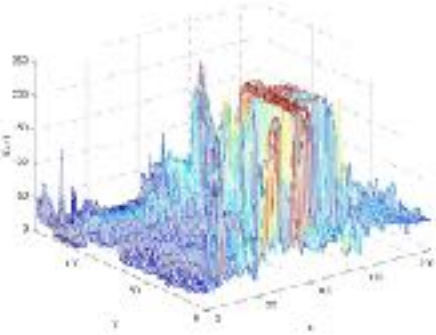


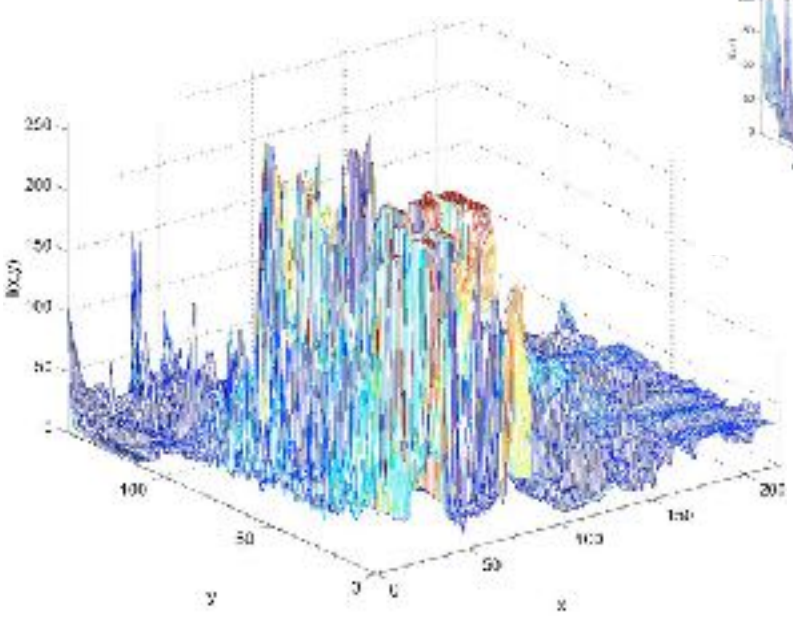
And so are these!

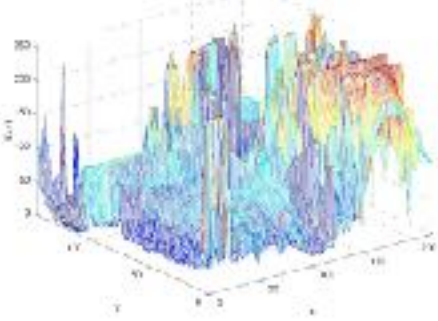






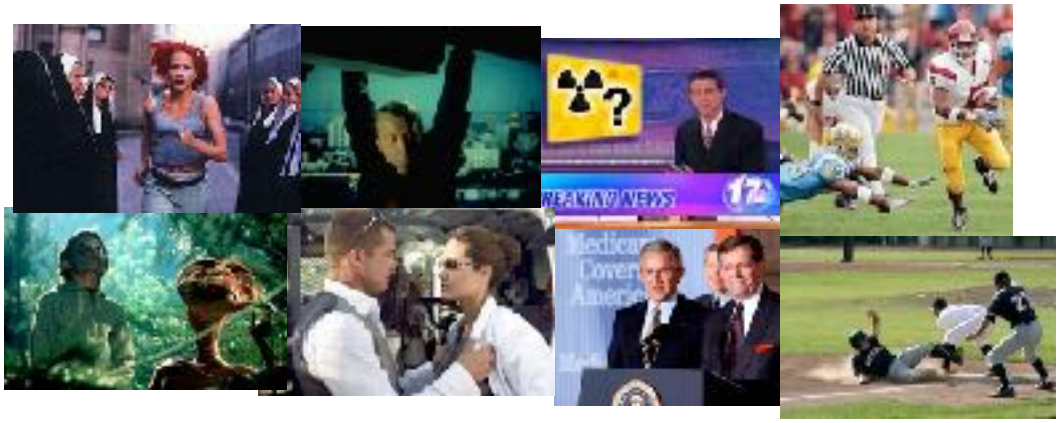




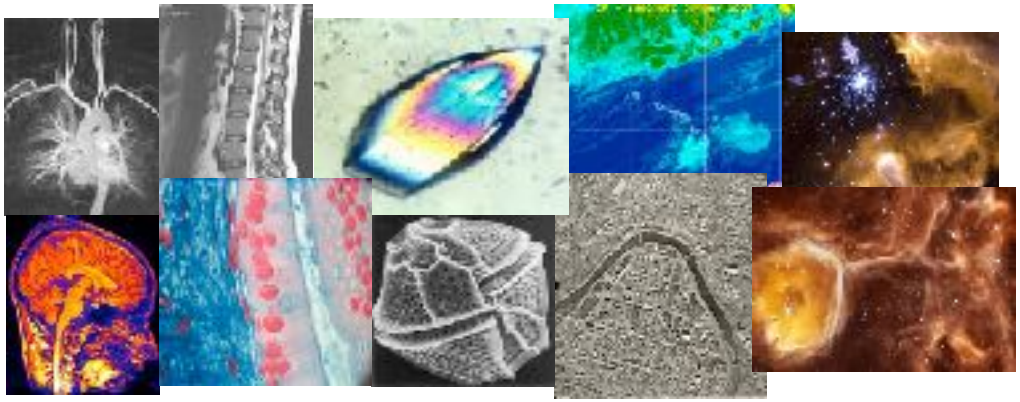


Why study computer vision?

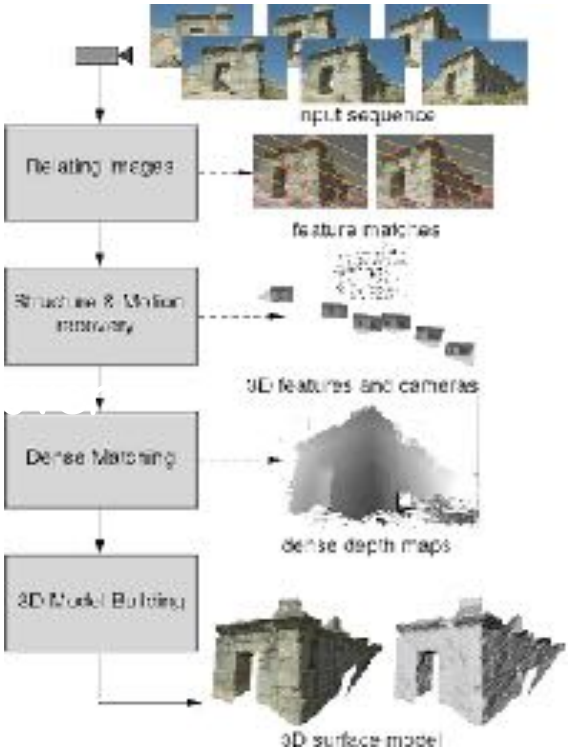
- Vision is useful: Images and video are everywhere!

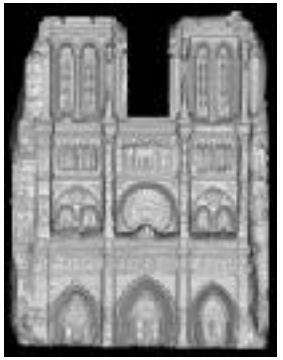


Google Image Search Picasa flickr websh ots picsearch YouTube



Vision as measurement device

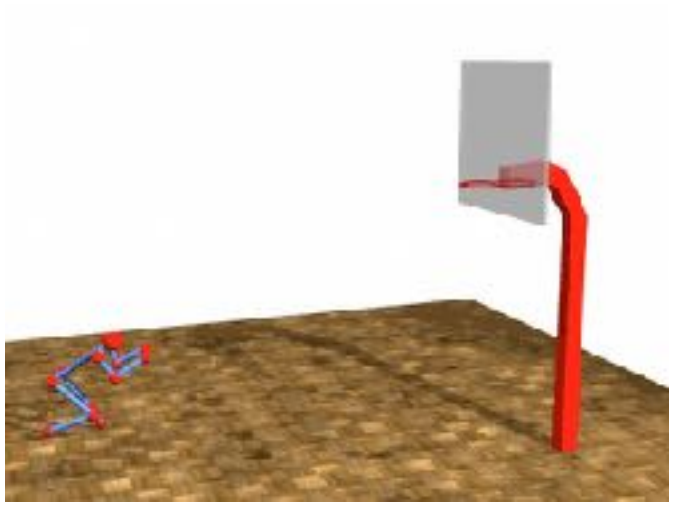




Special effects: shape and motion capture

Source: S. Seitz

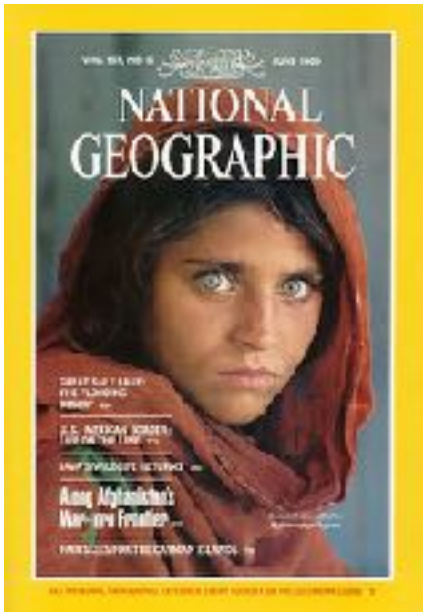
3D Modeling



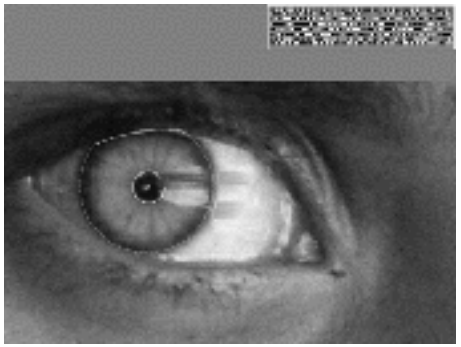
http://www.photogrammetry.ethz.ch/research/cause/3dreconstruction3.html

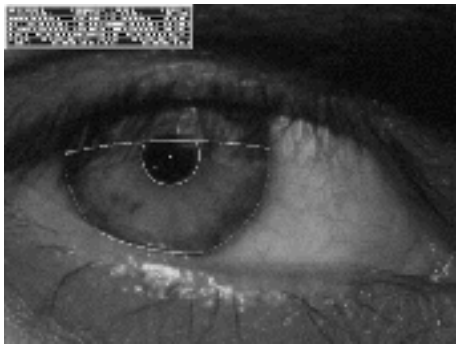
Face recognition: Apple iPhoto software

http://www.apple.com/ilife/iphoto/



How the Afghan Girl was Identified by Her Iris Patterns





Source: S. Seitz

Biometrics

Fingerprint scanners on many new laptops, other devices

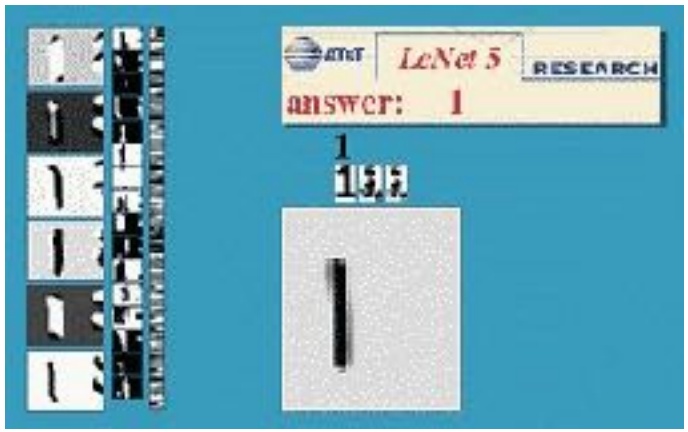
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Source: S. Seitz

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs

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

Source: S. Seitz

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

Google goggles labs


Landmark

Book


Contact Info.

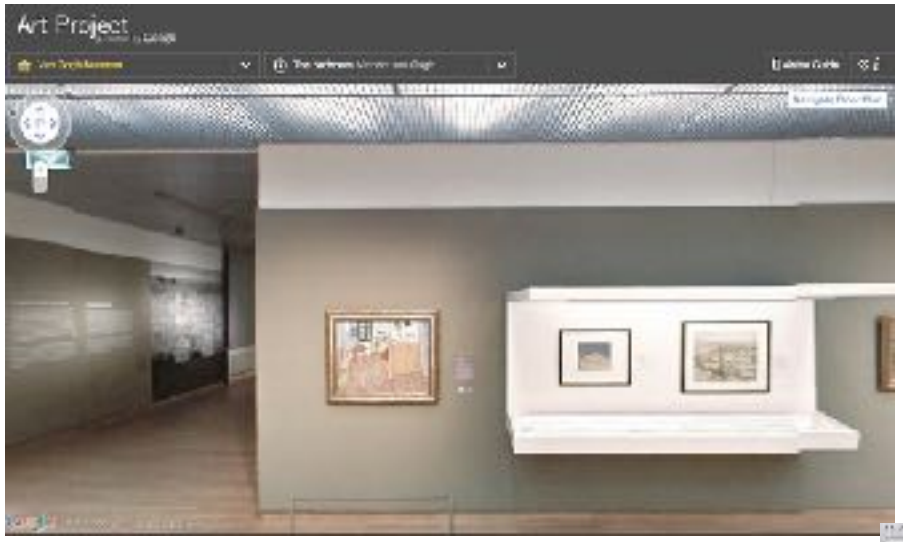

Artwork


Places

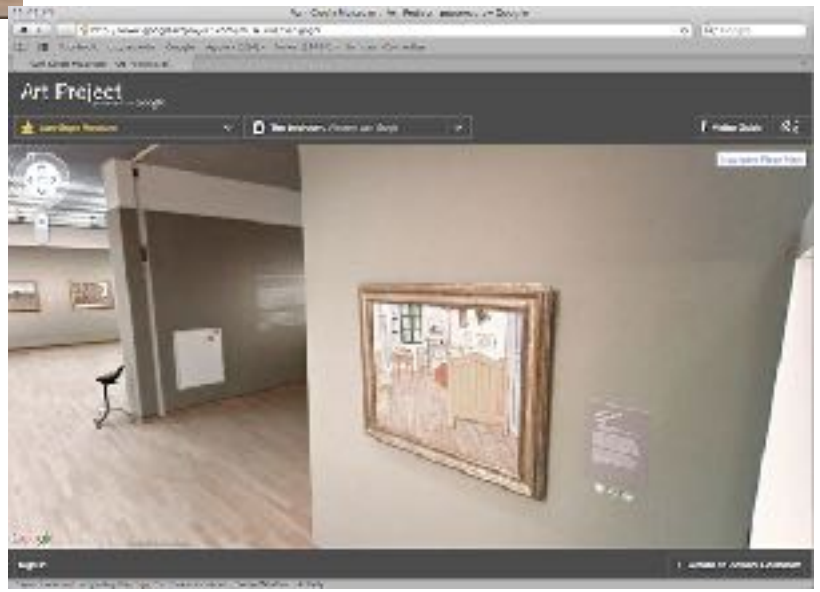

Wine

Logo

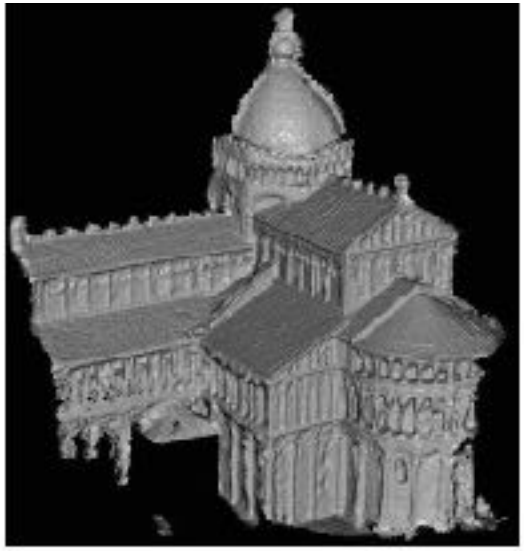
Google Art Museum Project



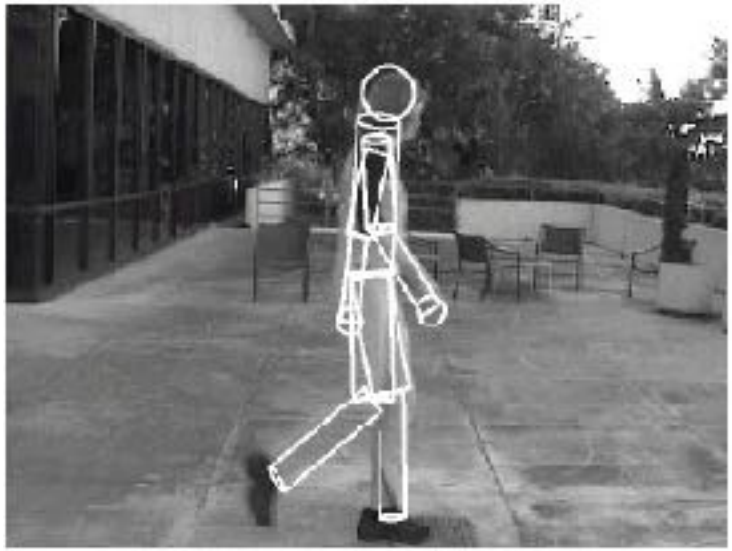
Navigate museums of the world



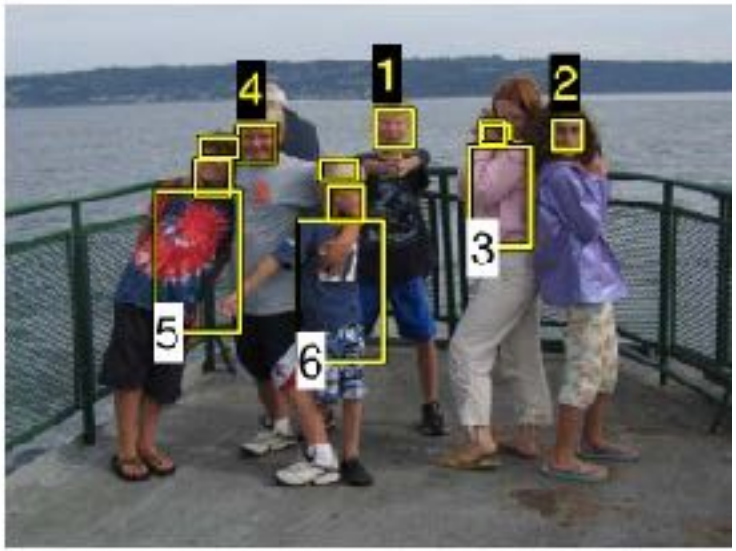
(a)



(b)



(c)



(d)

Automotive safety

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - "In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians."

Source: A. Shashua, S. Seitz

Vision in supermarkets

LaneHawk by EvolutionRobotics

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… "

Source: S. Seitz

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

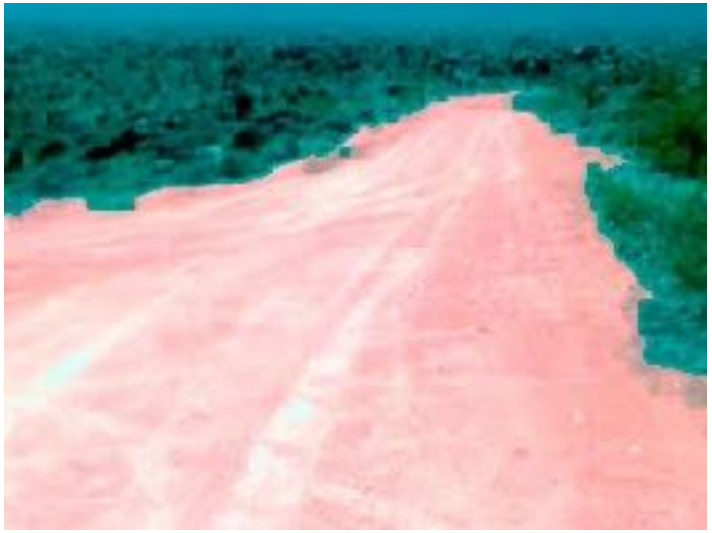
Sony EyeToy

Assistive technologies

Xbox and Kinect sensor

Source: S. Seitz

Classification



Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

Object categorization

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

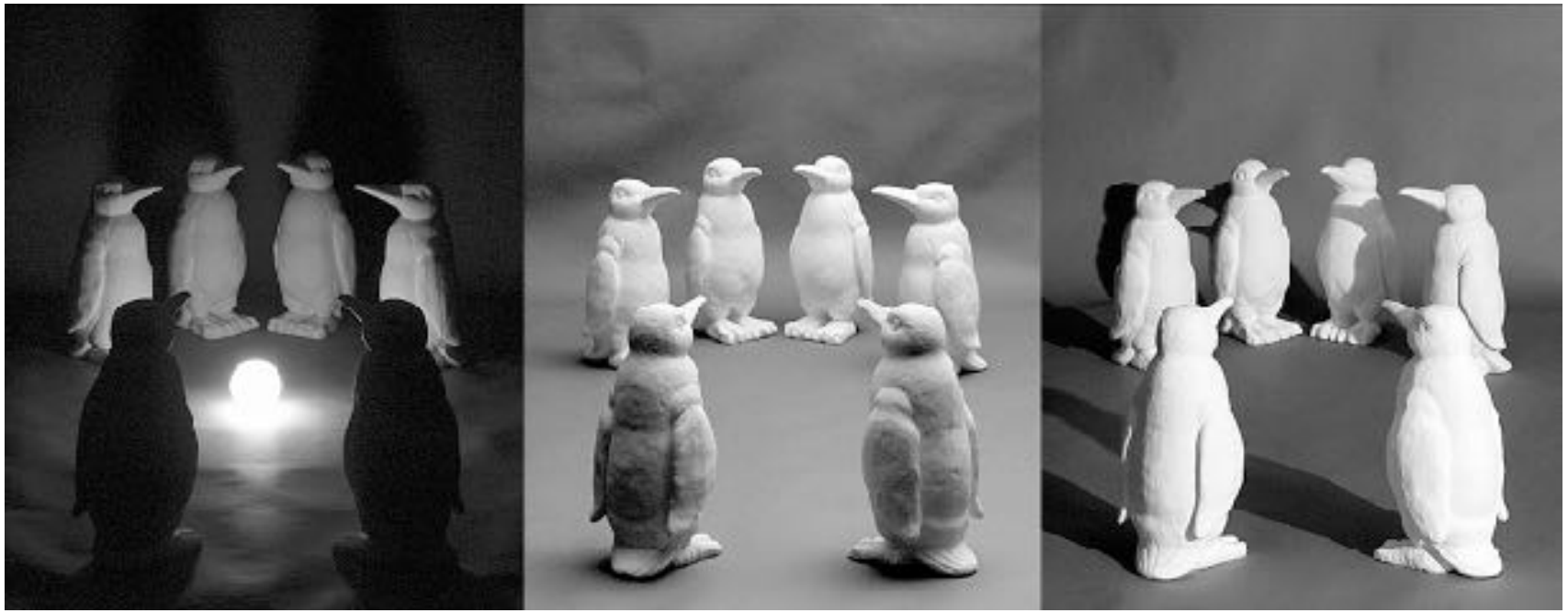
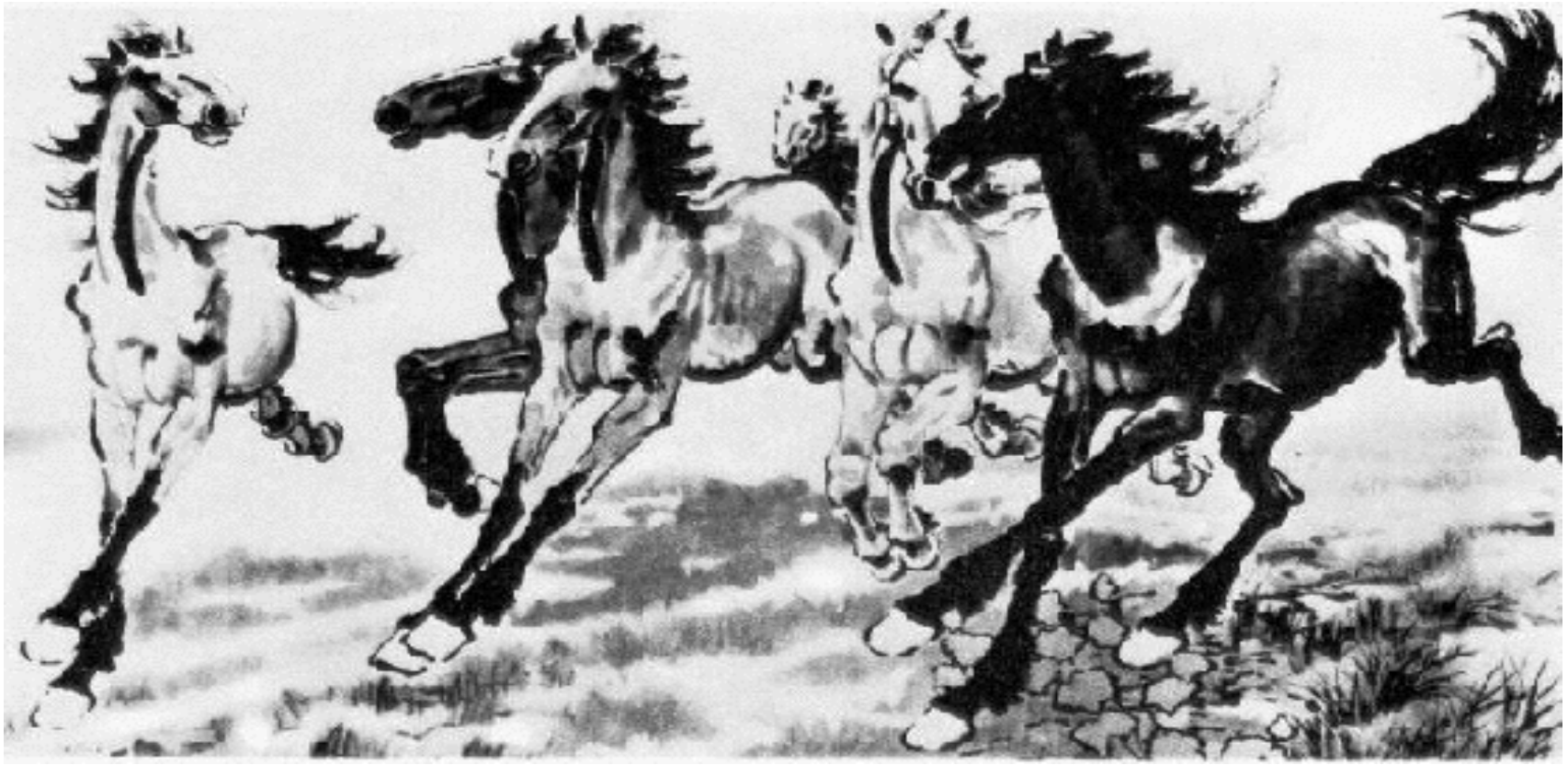


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

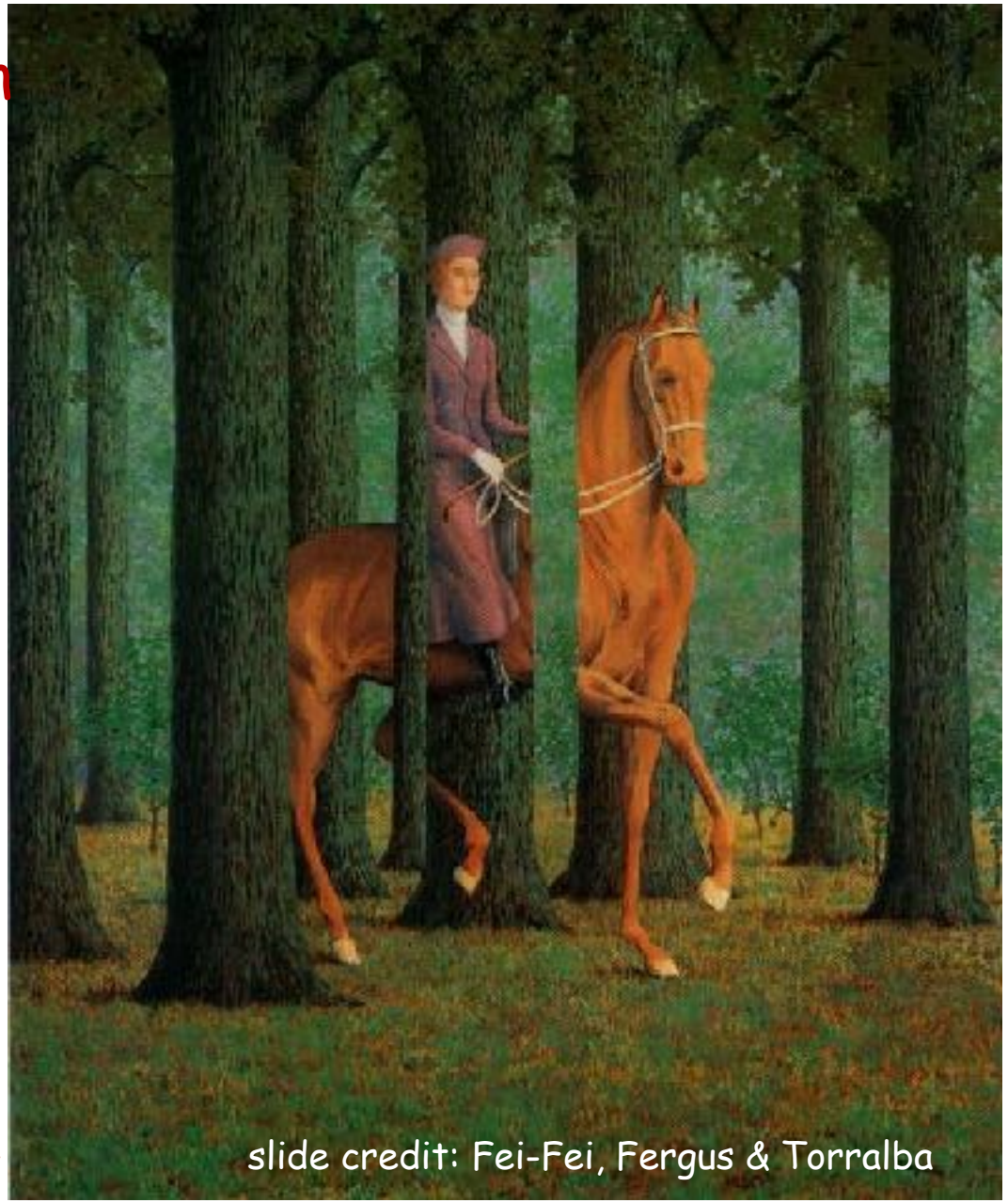
Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

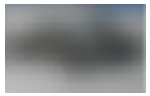
Challenges: background clutter

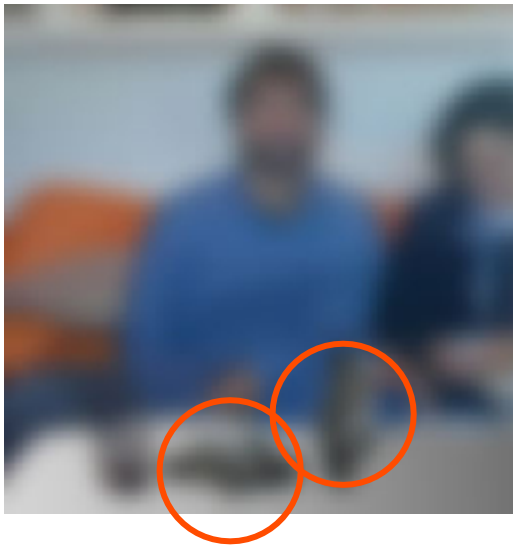
Challenges: Motion

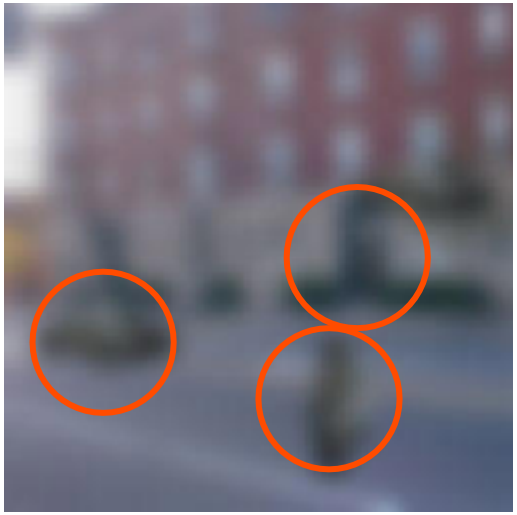
Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity







slide credit: Fei-Fei, Fergus & Torralba

Levels of complexity

- Early vision – local operations, compute maps, or statistics of individual pixels (edges, motion fields, depth maps)
- Midlevel vision – assembly of local information (segmentation, contour completions, grouping)
- Scene analysis – recognition of objects, scenes
- Active vision – how to control and use the resources to adjust the sensor to gather additional information
- Goal directed vision – control behaviors based on visual information

Contents of the Class

Image Processing, Low-level and Mid-level Vision :

- Image sensing, lenses
- Non-traditional sensors & perceptual coordinate systems
- Photometry and Color
- Filtering, correlation, convolution, noise
- Fourier transform
- Edge detection, Boundary detection
- Hough transforms
- Features, Corners, SIFT features
- Image and Motion
- Segmentation
- Texture Analysis

Multiple view Geometry for Robotics:

- Geometric transforms
- Projective geometry
- Camera Calibration
- Epipolar geometry
- Stereopsis
- Optical flow
- Tracking

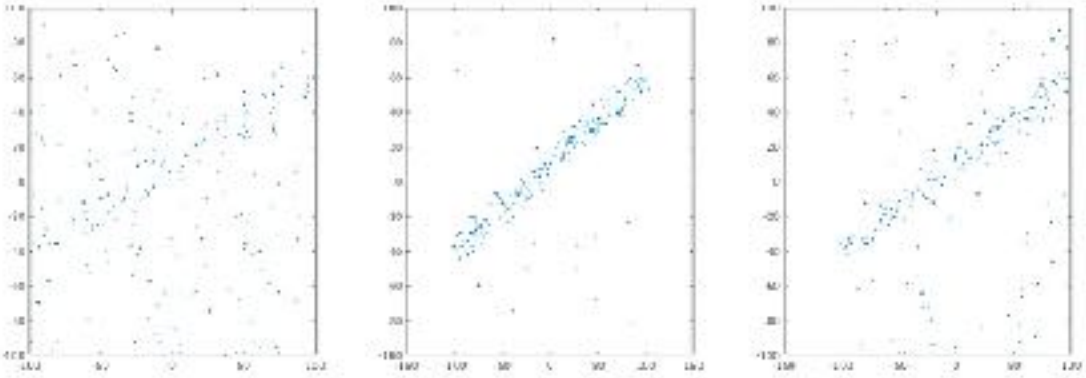
Image Recognition

- Recognition of specific objects
- Recognition using Machine Learning, SVM, HOG features
- Recognition using Neural Networks
- Applications of Recognition

Short description of Projects

Homework 1

- Review of estimation
- LS estimation TLS estimation , LS with Regularization and RANSAC, applied to the problem of line fitting

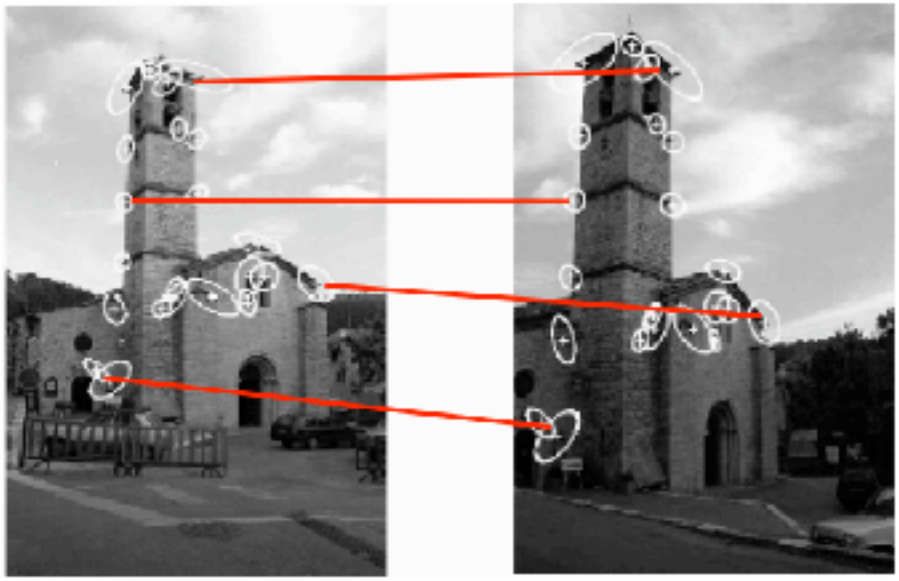


Project 1: Color Segmentation with GMM

Detect the ball in images “seen” by Nao.

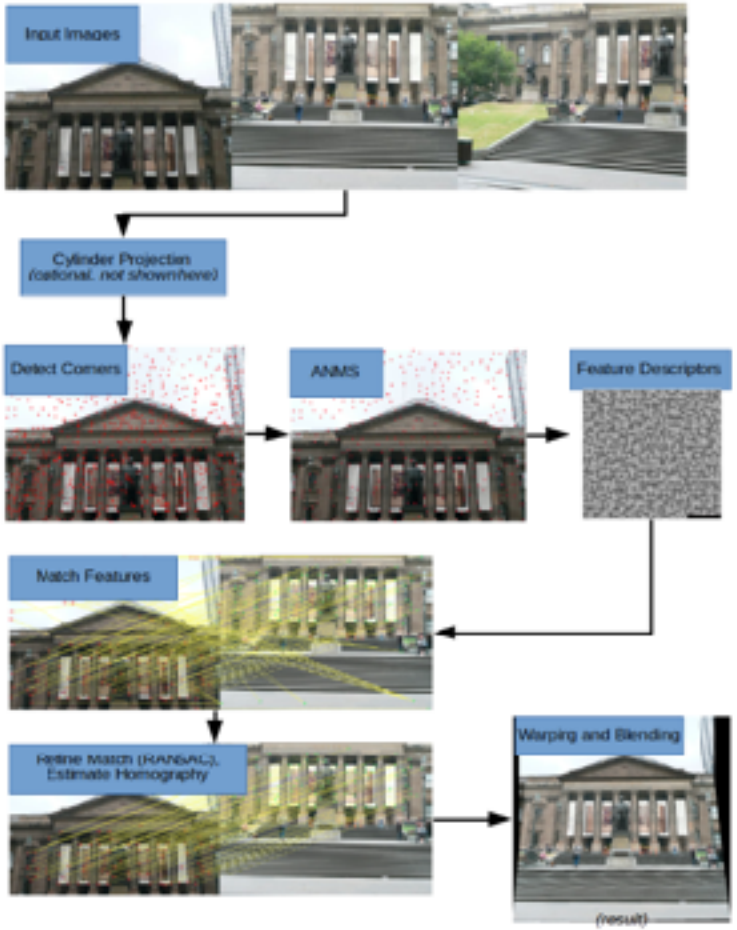
You will learn about Color imaging
and about Clustering approaches (K-mean and GMMs)

Homework 2: Image Features and Warping



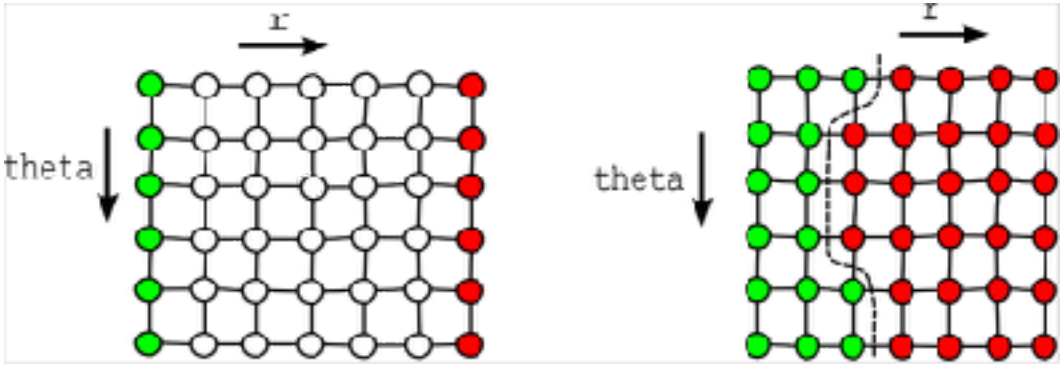
The project involves: corner detection and
geometric transformations between image planes

Project 2: Panorama Stitching



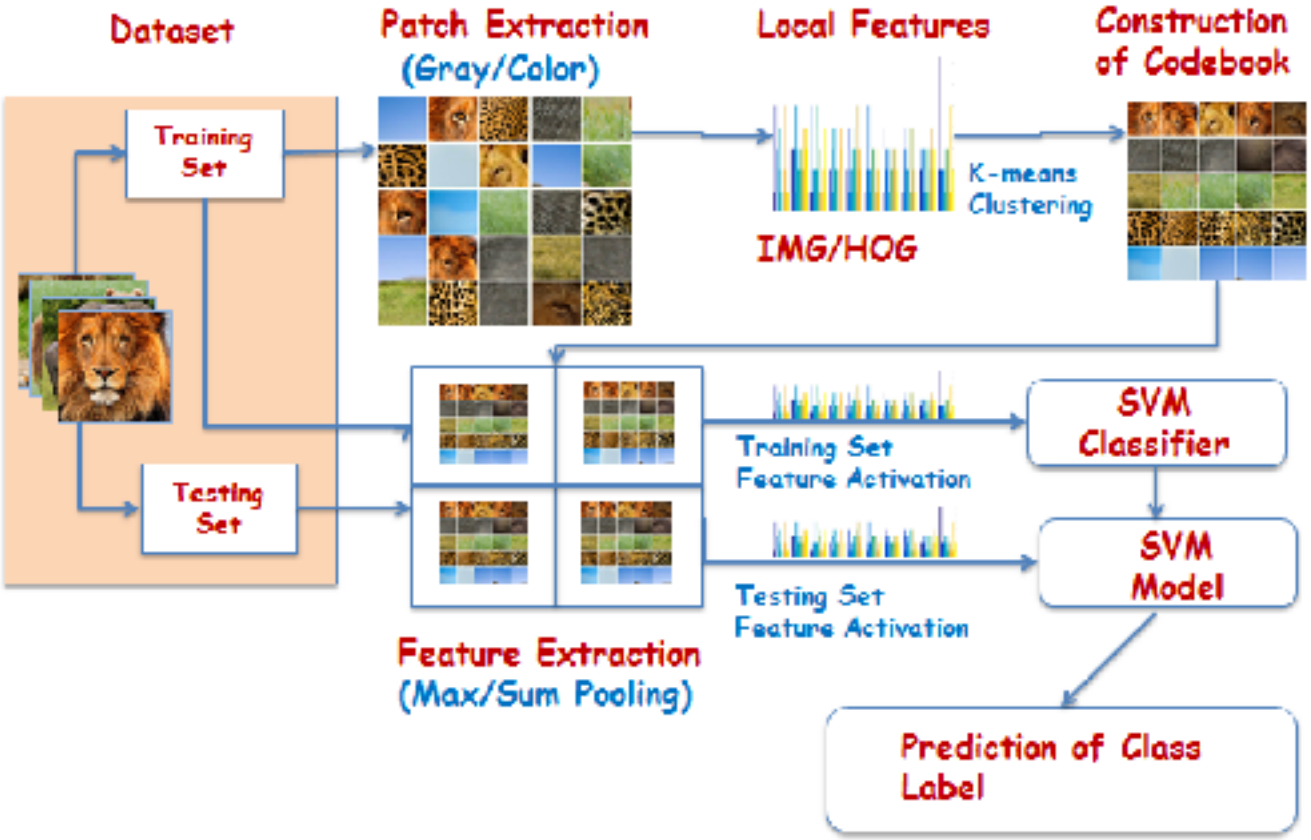
Use the image features to derive the transformation between images and blend

Project 3: Segmentation with Graphcuts

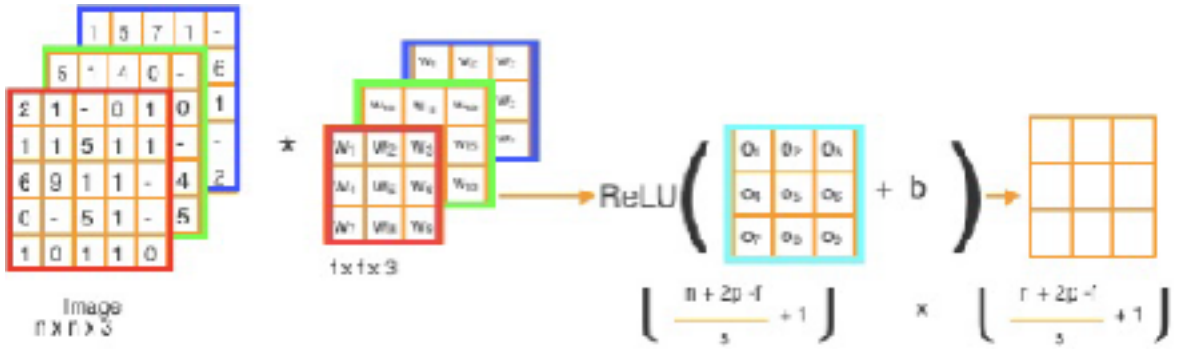


Segment foreground from background in log-polar coordinates using Edge, Color, Texture, and Motion information.

Homework 3:Image classification using HOGs and Bag of Words



Project 4:Image classification using CNNs



Write a CNN, then train it as a classifier