BIL 722
Advanced Topics in Computer Vision

Oct. 4, 2012

Erkut Erdem
Dept. of Computer Engineering
Hacettepe University

Introduction

Instructor and Course Schedule

- Dr. Erkut ERDEM
- erkut@cs.hacettepe.edu.tr
- Office: 114
- Tel: 297 7500 / 149

- Lectures: Thursday, 13:30-16:15@D5
- Office Hour: *To be announced!*

Communication

- The course webpage will be updated regularly throughout the semester with lecture notes, programming and reading assignments and important deadlines.
  http://web.cs.hacettepe.edu.tr/~erkut/bil722.f12

- All other communications will be carried out through Piazza. Please enroll it by following the link
  https://piazza.com/hacettepe.edu.tr/fall2012/bil722

Today

- Introduction
- Administrative stuff
- Overview of the course
- Topics covered in this semester

A bit about Computer Vision at our department


Ahmet Burak Can


Aykut Erdem


Erkut Erdem


Nazli Ikizler-Cinbis

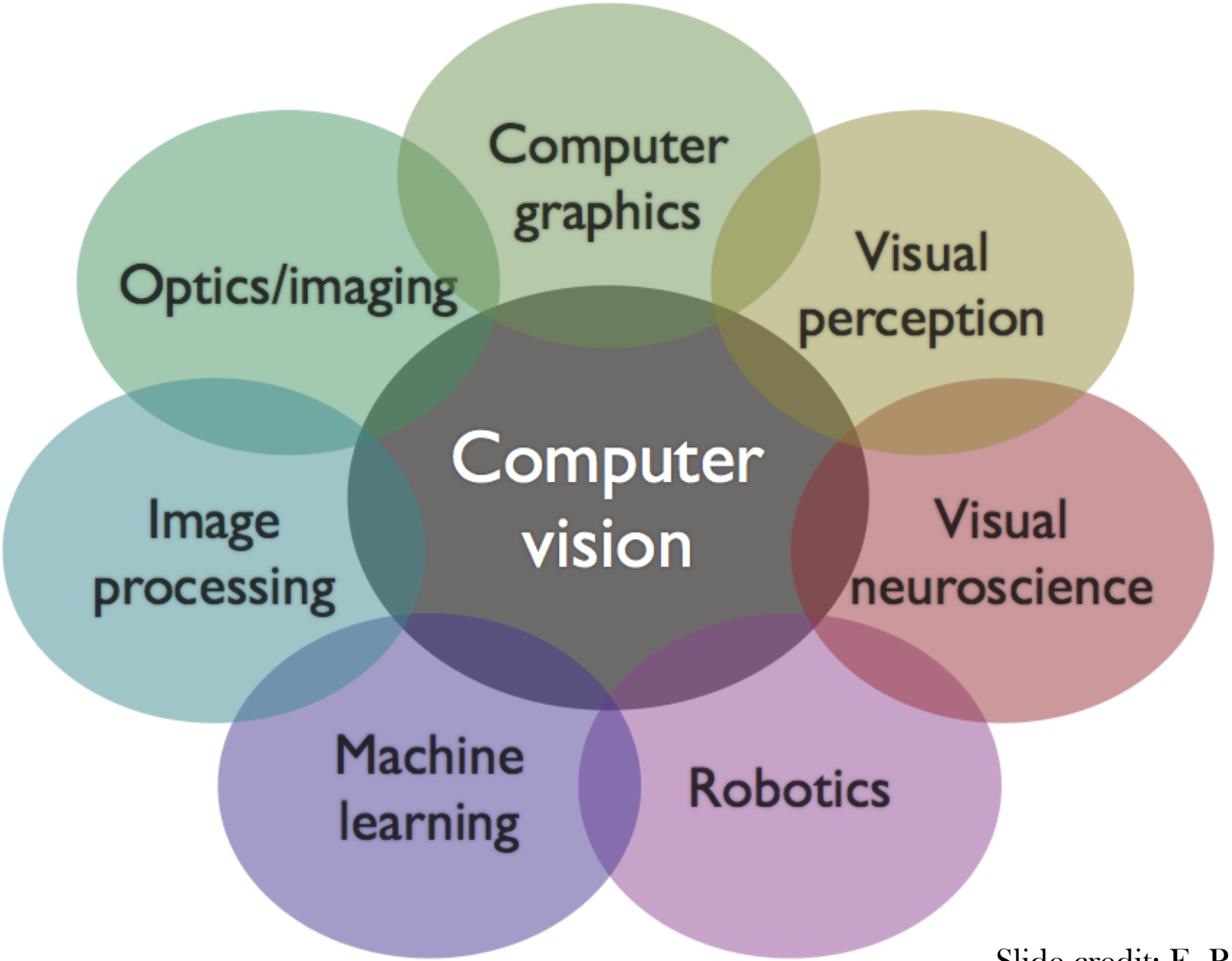
huvision

http://vision.cs.hacettepe.edu.tr

About BIL722

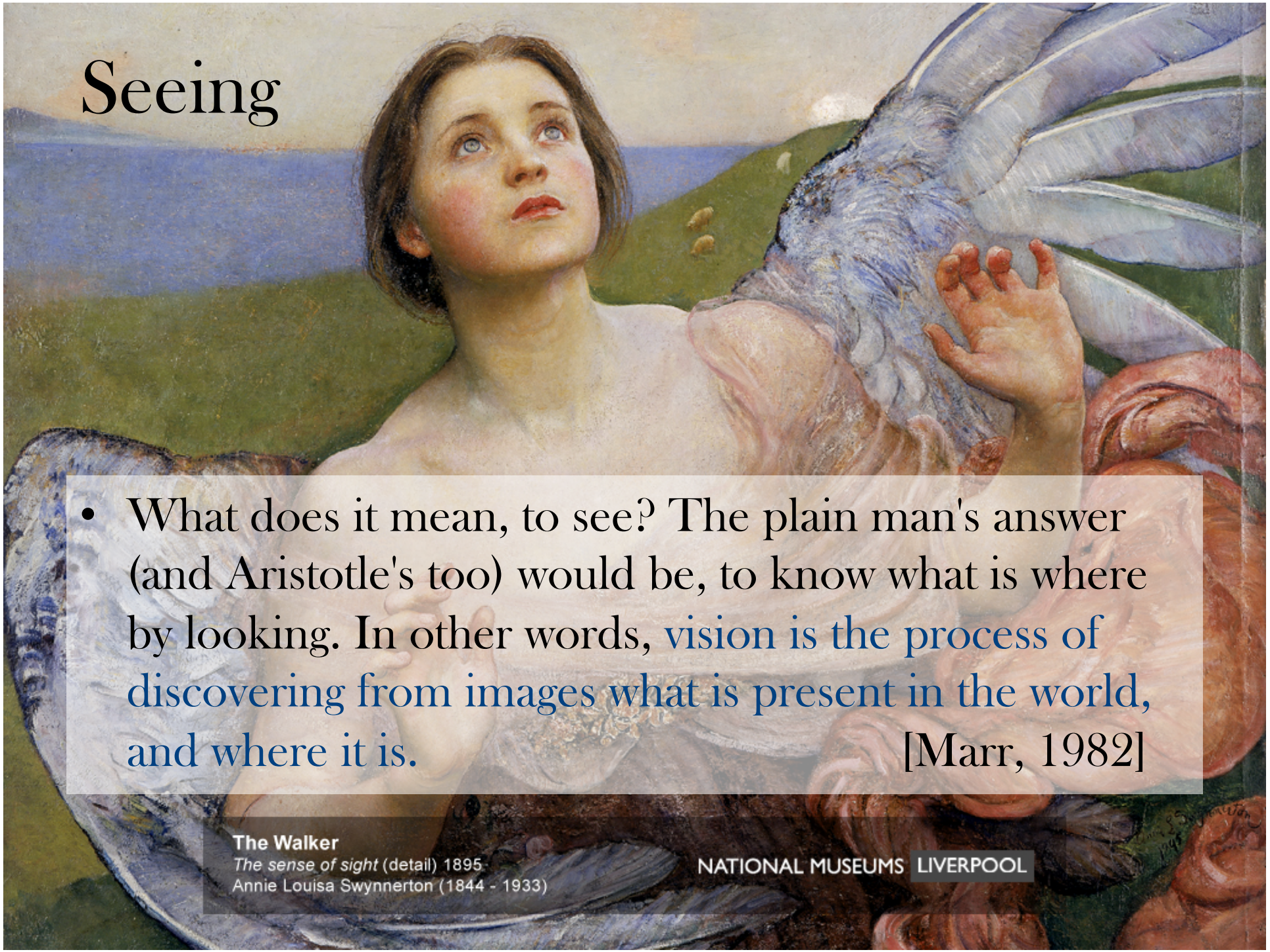
- An advanced-level graduate seminar course which takes an in-depth look at a wide selection of important topics in computer vision.

Related Disciplines



Slide credit: E. P. Simoncelli

Seeing



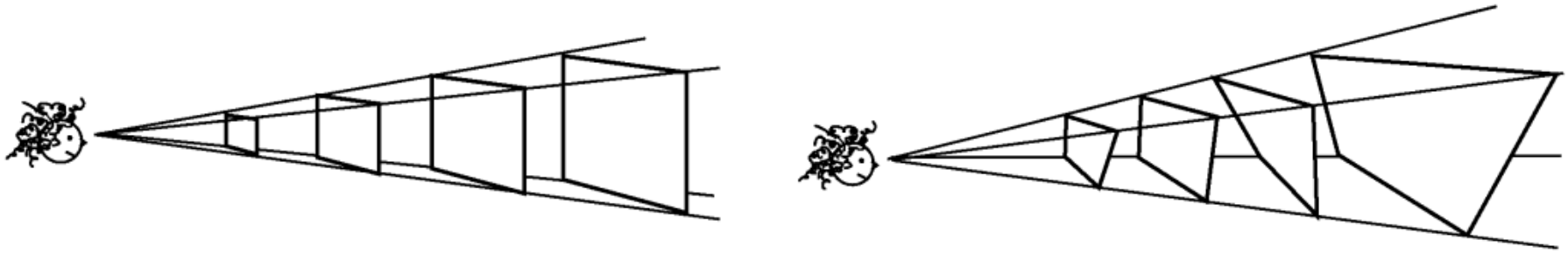
- What does it mean, to see? The plain man's answer (and Aristotle's too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is. [Marr, 1982]

The Walker
The sense of sight (detail) 1895
Annie Louisa Swynnerton (1844 - 1933)

NATIONAL MUSEUMS LIVERPOOL

Why is vision hard?

- A typical image includes many objects organized in many different configurations.
- Vision requires solving ill-posed problems.
- Images are both complicated and highly ambiguous.
- Same object can generate very different images.
- Different objects can generate similar images.



Figures: Steven Pinker, How the Mind Works, 1997

Challenges: Illumination

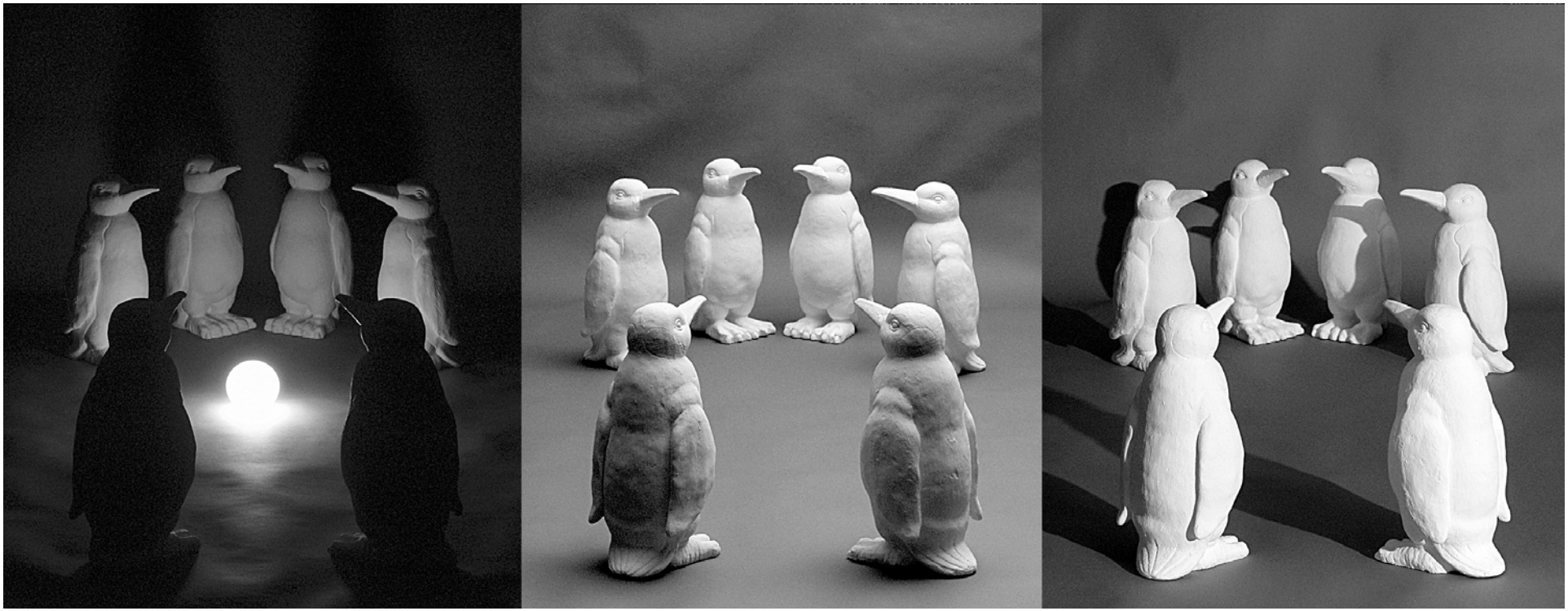


Figure: J. Koenderink

Slide credit: L. Fei-Fei

Challenges: viewpoint variation

Michelangelo 1475-1564

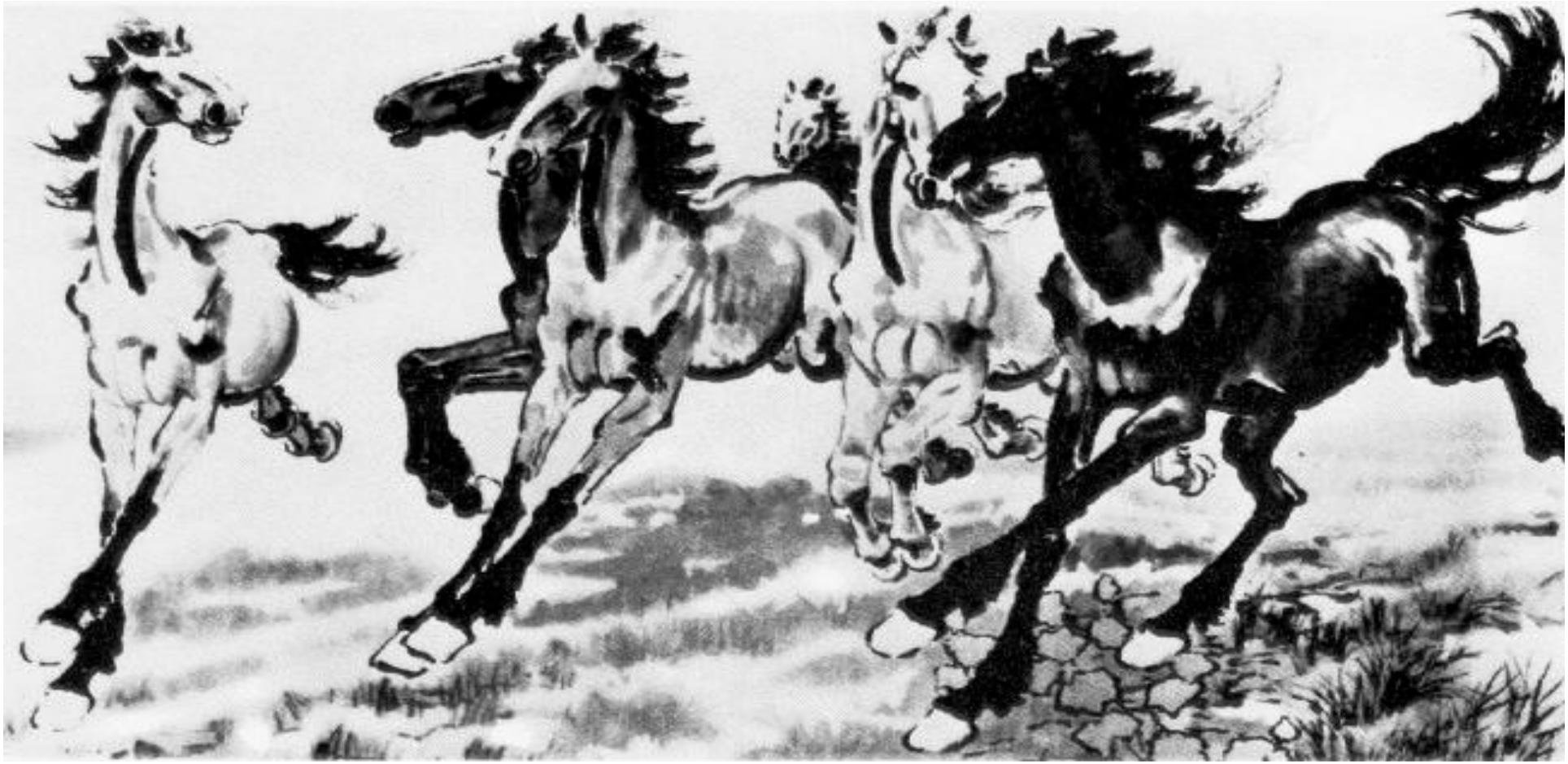
slide credit: Fei-Fei, Fergus & Torralba

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba

Challenges: Scale

Slide credit: L. Fei-Fei,
R. Fergus and A. Torralba

Challenges: Deformations



Xu, Beihong 1943

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba

Challenges: Occlusion

The Blank Check,
by René Magritte

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba

Challenges: background clutter

Slide credit: S. Lazebnik

Challenges: Motion

Slide credit: S. Lazebnik

Challenges: Some things have strong variations in appearance

Slide credit: B. Freeman and A. Torralba

Challenges or opportunities?

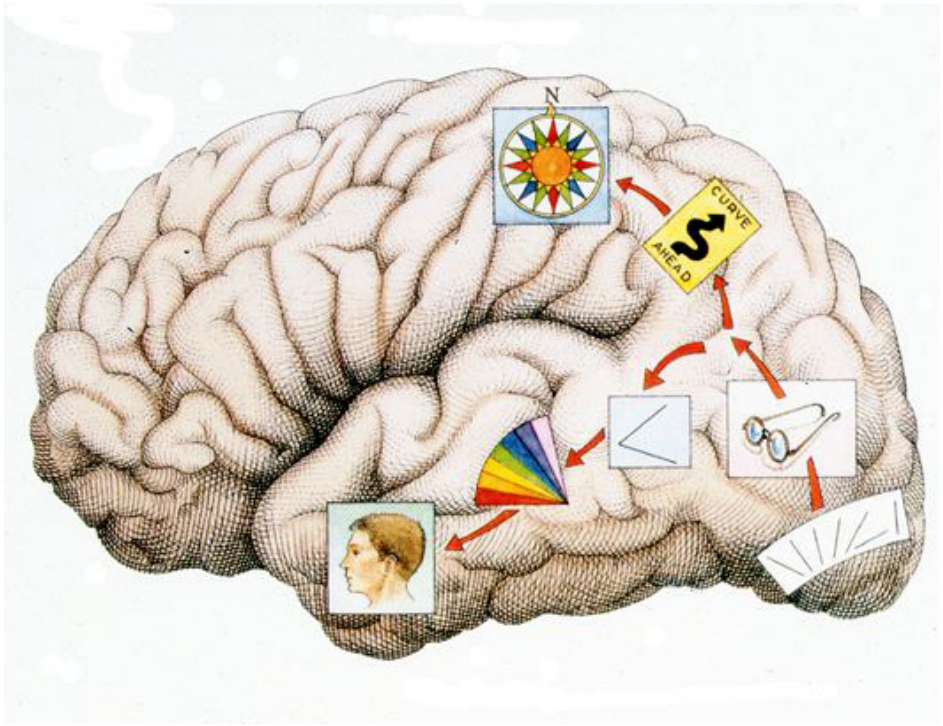
- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

Image: J. Koenderink

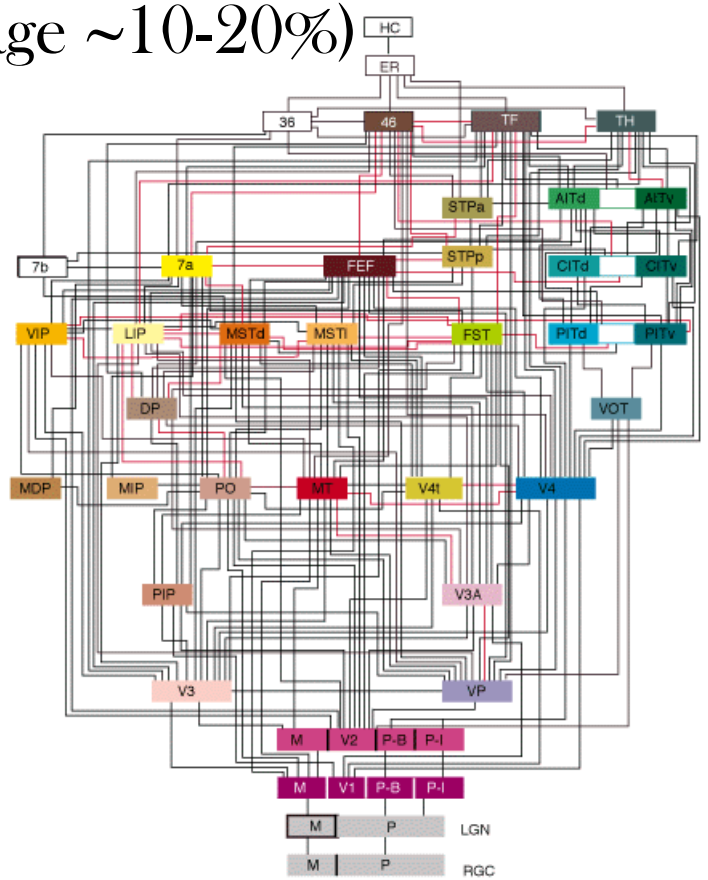
Slide credit: S. Lazebnik

Why does vision appear easy to humans?

- Our brains are specialized to do vision.
- ~50% of the cortex in a human brain is devoted for visual processing
  (cf. motor control ~20-30%, language ~10-20%)

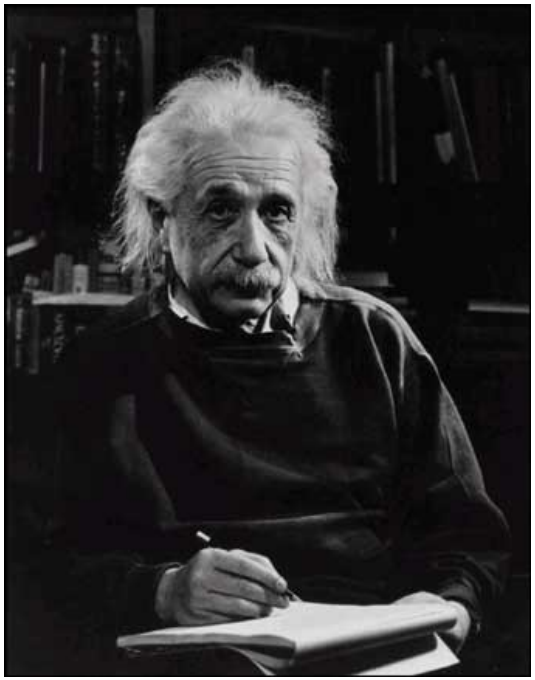


[David Heeger, 2006]



[Felleman and van Essen, 1991]

Computer Vision



What we see

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 a computer sees

Slide credit: S. Narasimhan

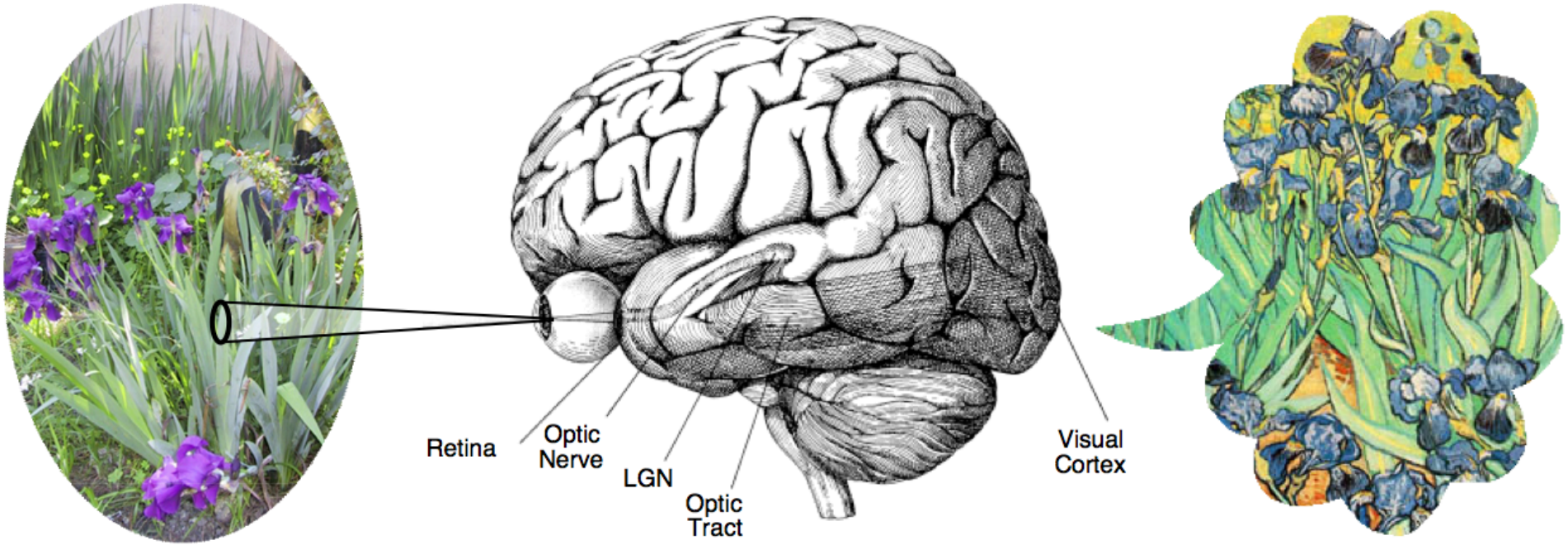
Marr's observation: Studying vision at 3 levels

"Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information"

Vision as an information processing task

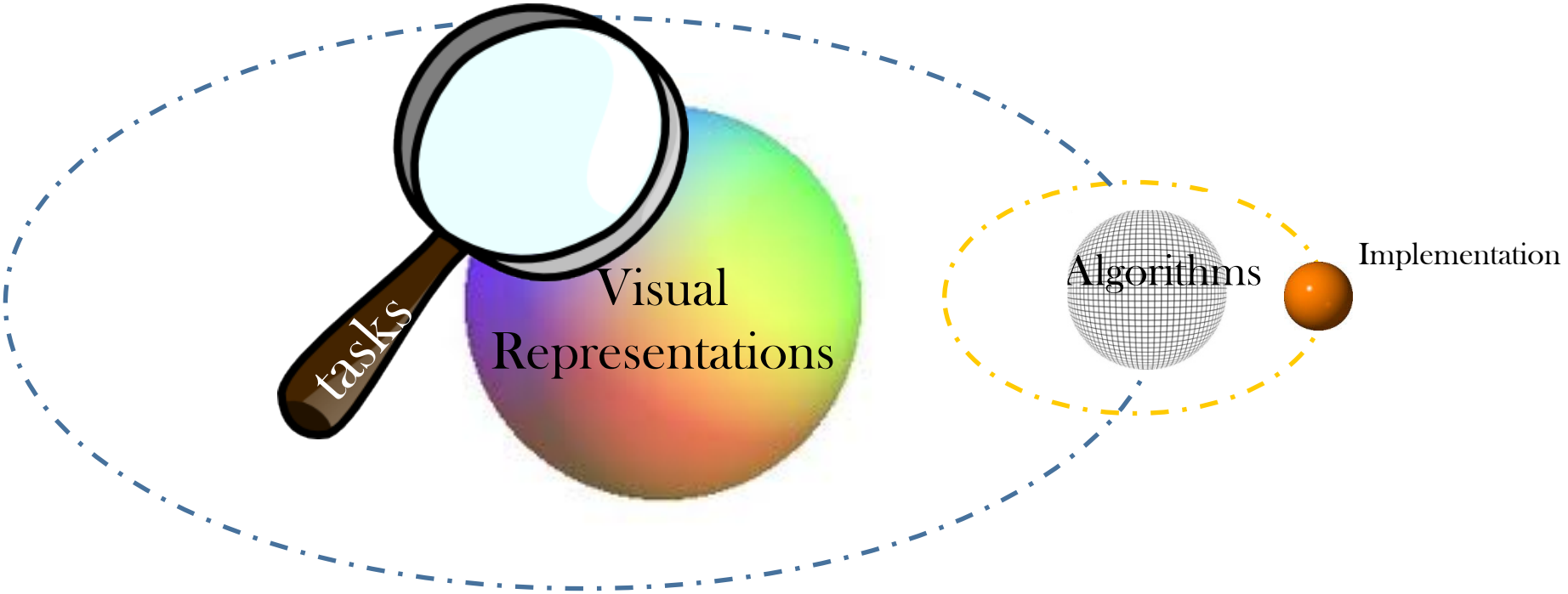
- 3 levels of understanding:
  1. Computational theory
     - What is computed? Why it is computed?
  2. Representation and Algorithm
     - How it is computed?
     - Input, Output, Transformation
  3. Physical Realization
     - Hardware

Marr's observation: Studying vision at 3 levels



Slide Credit: E. Simoncelli

Marr's observation: Studying vision at 3 levels



Slide Credit: S. C. Zhu

In this course

- We will survey and discuss recent conference papers and research articles related to several computer vision topics.

Goals

- Give a deeper understanding of the state- of-the-art methods in vision literature
- Provide students the ability to identify open issues and possible directions for future research

- Read and discuss some interesting recent papers together
  - Learn how to speak,
  - Learn how think critically

- Complete a vision project throughout the semester

Course Organization

- 20% Paper presentations
- 15% Class participation
- 10% Response papers
- 20% Project (progress report, presentation)
- 35% Final exam (final project report)

Paper Presentations

- Each week we will be discussing 2-3 papers on a specific topic.
- Start picking your papers from the list.
- Prepare a good, conference-quality presentation (30+15 mins)
  - Clearly state the problem
  - Discuss key technical ideas, main contributions, assumptions, strengths and weaknesses
- Send your slides the day before the class
- Not just present but also defend the paper in front of the class

Paper Presentations

- For some of the papers, its code may be available on the web
- If so,
  - Run it on new data,
  - Play with parameters,
  - Discuss the paper accordingly.

Class Participation

- Come to class as prepared as the presenter
- Actively participate in class discussions
- Voice your ideas, comments and opinions

Paper Responses

- For half of the papers in the list,
  - Hand in a brief summary (1-2 pages) of each paper
    - Summary of the paper in your own words,
    - Main contributions of the paper,
    - Strengths and weaknesses of the paper,
    - Details of the experimental evaluation,
    - Areas of further improvements
  - Send your responses via e-mail before the beginning of the class

Project

- Each student will work individually.
- Design of a novel approach with strong experimental evaluation.

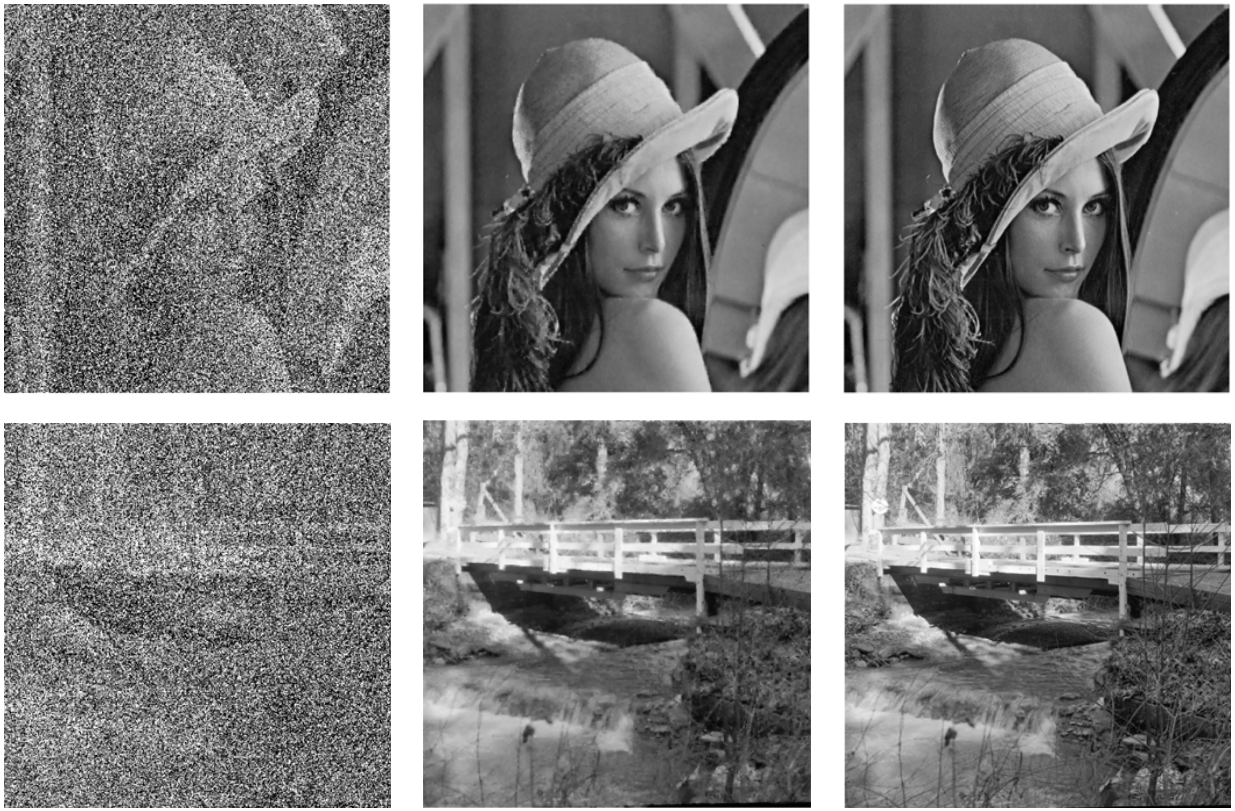
- Project proposals in a few weeks.
- Progress reports in the middle of semester.
- Project presentations at the end of semester.

Course Schedule

- (1 week) Introduction to the course
- (1 week) Image smoothing, restoration and enhancement
- (2 weeks) Role of context
- (2 weeks) Action
- (1 week) Scene Understanding
- (2 weeks) Visual saliency
- (1 week) Video processing
- (1 week) Image retrieval
- (1 week) Miscellaneous
- (1 week) Project Presentations

Image Smoothing

- Images are corrupted with 70% salt-and-pepper noise

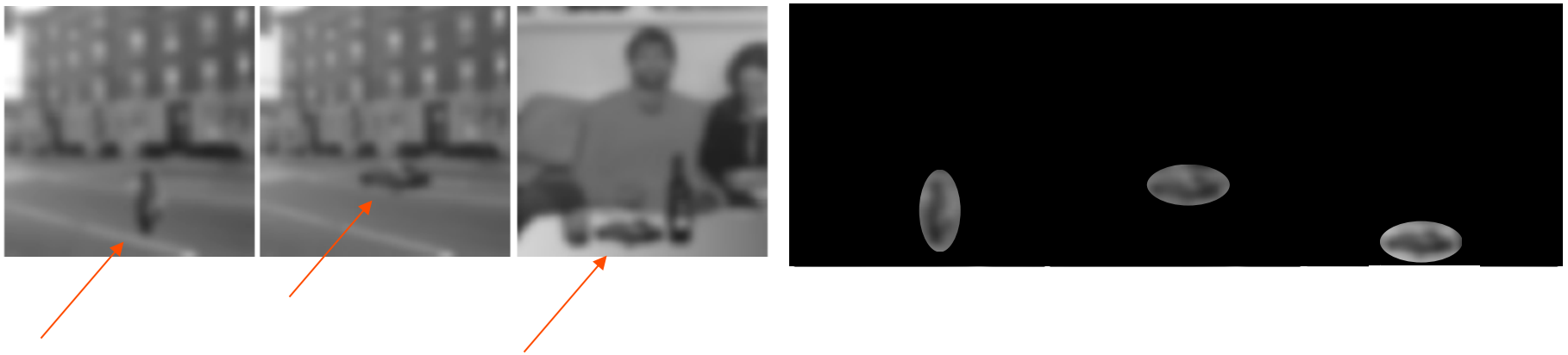


Noisy input Recovered image Original image

What do these examples demonstrate?

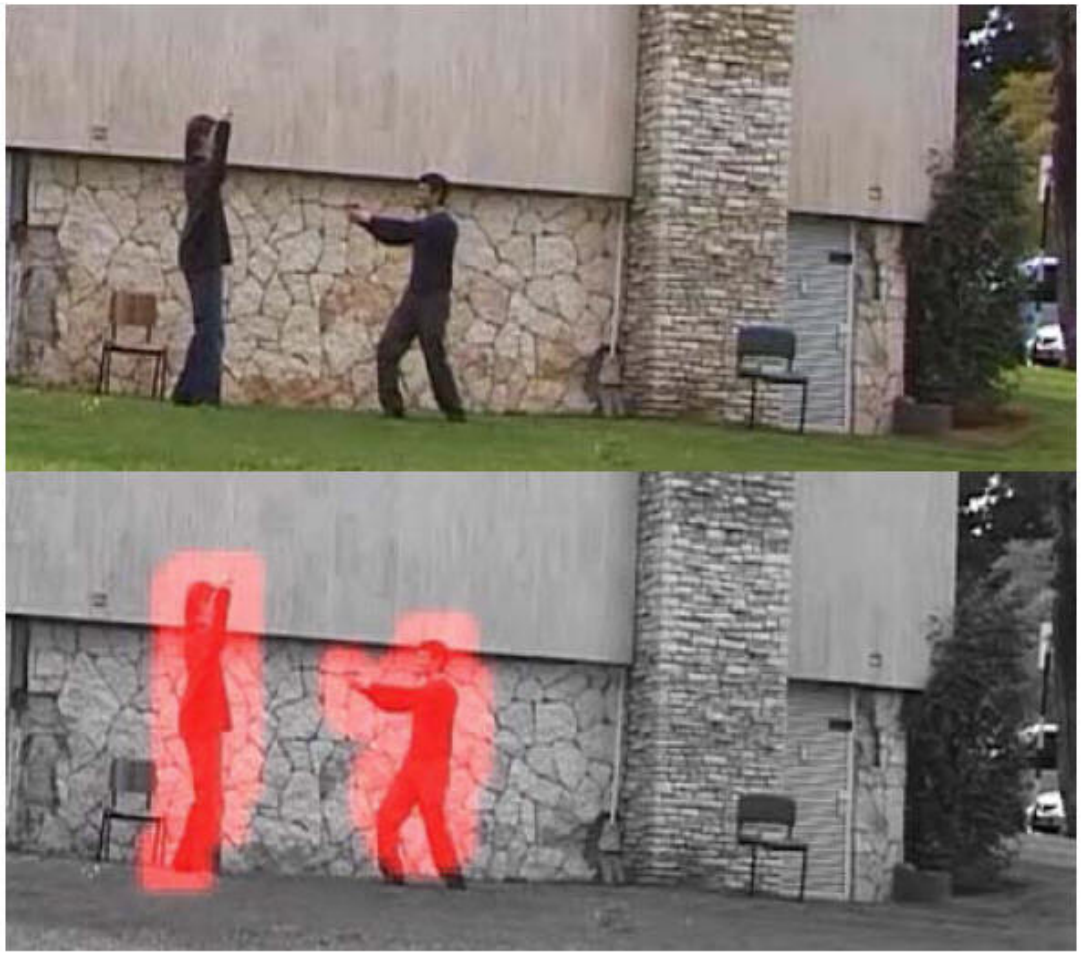
R. H. Chan, C.-W. Ho, and M. Nikolova, Salt-and-Pepper Noise Removal by Median-Type Noise Detectors and Detail-Preserving Regularization. IEEE TIP 2005

Role of context



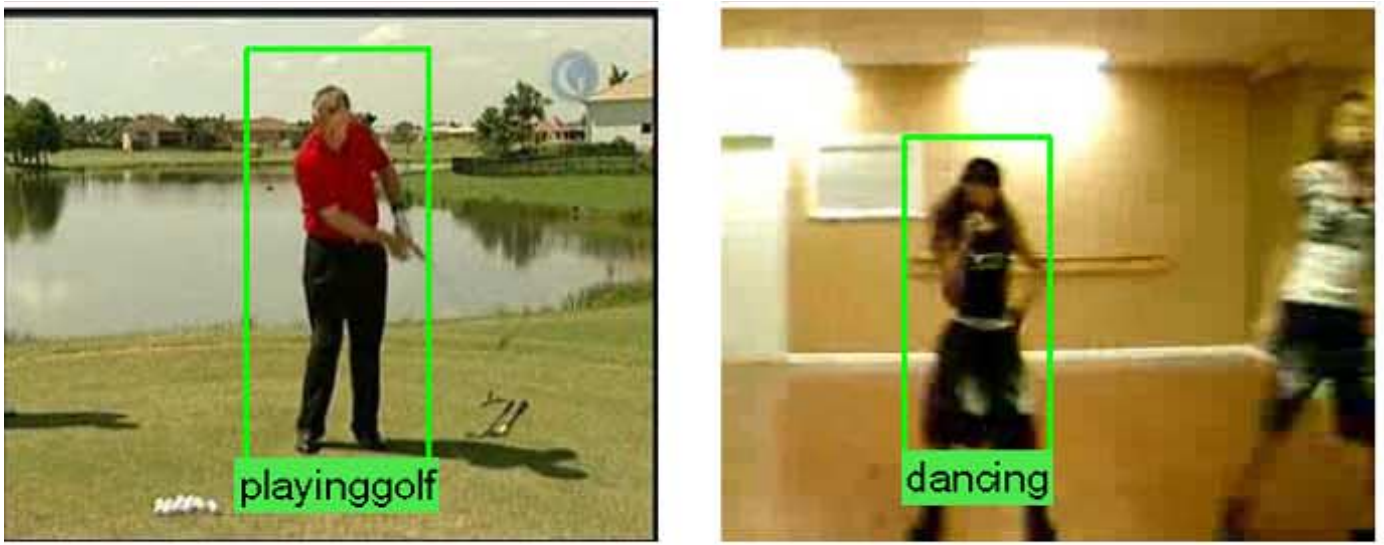
Slide credit: A. Torralba, D. Hoiem

Action



O. Boiman and M. Irani, Detecting Irregularities in Images and in Video, ICCV2005

Action



N.Ikizler-Cinbis, R. G.Cinbis and S. Sclaroff, Learning Actions From The Web, ICCV2009

Scene Understanding

slide by Fei Fei, Fergus & Torralba

Scene Understanding

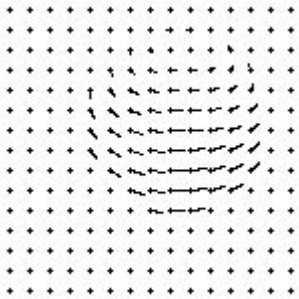
Visual Saliency

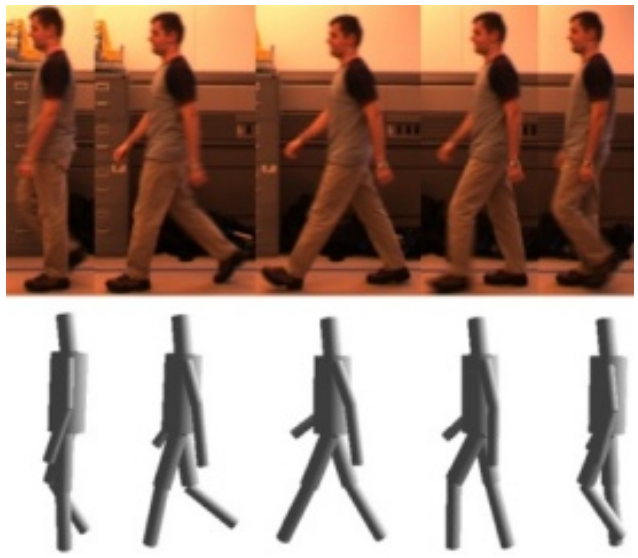
Visual Saliency

Visual Saliency

Video processing

Tracking objects, video analysis, low level motion

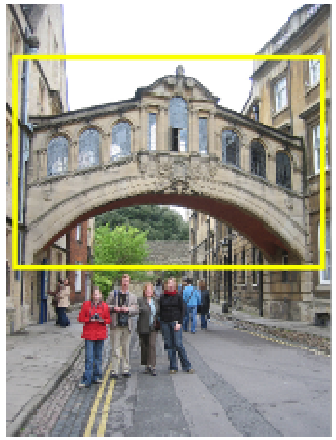


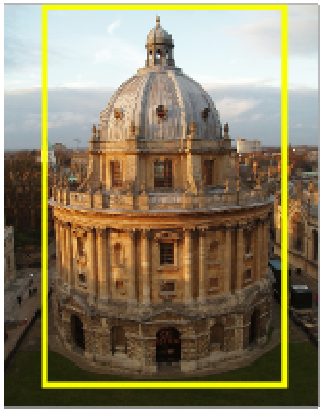


Tomas Izo

Slide credit: K. Grauman

Image retrieval





Find these landmarks

...in these images and 1M more

Slide credit: J. Sivic