Chapter 0

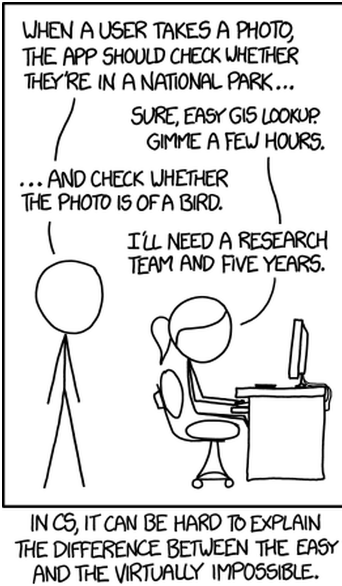
Introduction and Formalities

University of Konstanz

Lecture “Image Analysis and Computer Vision”
Winter semester 2014/15

Bastian Goldlücke

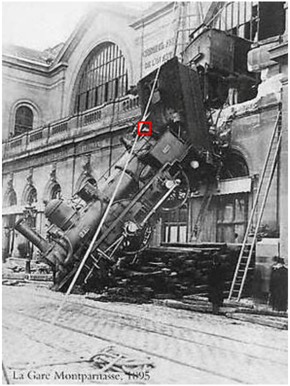
Computer Vision and Image Analysis



The goals of computer vision

- "Perceive the story behind the picture"
- Compute properties of the world
  - 3D shape and appearance
  - Names of people and objects
  - Track things moving through a scene

The goals of computer vision



180	199	219	246	86	10	90	55	52	50
221	160	255	137	15	35	69	74	76	41
184	173	241	52	7	81	22	66	67	60
113	254	181	20	55	52	24	53	51	31
157	255	105	21	5	74	155	64	50	30
252	210	57	1	52	233	192	106	90	52
248	147	9	41	218	232	204	155	166	158
232	76	39	235	253	136	181	237	255	244
225	14	27	54	28	26	17	74	84	96
77	47	31	20	50	20	40	9	34	29

What the computer “sees” is numbers:

180	199	219	246	86	10	90	55	52	50
221	160	255	137	15	35	69	74	76	41
184	173	241	52	7	81	22	66	67	60
113	254	181	20	55	52	24	53	51	31
157	255	105	21	5	74	155	64	50	30
252	210	57	1	52	233	192	106	90	52
248	147	9	41	218	232	204	155	166	158
232	76	39	235	253	136	181	237	255	244
225	14	27	54	28	26	17	74	84	96
77	47	31	20	50	20	40	9	34	29

... and by analyzing this data, you want to ...

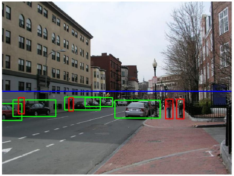
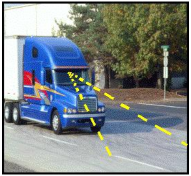
... detect faces

- Many new digital cameras now detect faces: Canon, Sony, Fuji ...

... recognize objects and people

[Nokia, Google goggles]

- Pedestrian and car detection
- Lane detection and lane departure warning
- Collision warning systems with adaptive cruise control


... compute 3D structure from photo collections

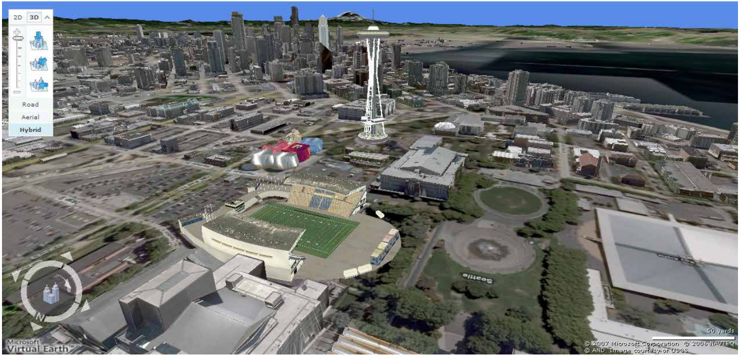


Image: Microsoft Virtual Earth, see also Google Earth, Google StreetView

... capture 3D shape for movies, computer games ...

[The Matrix movies, ESC Entertainment, XYZRGB, NRC]

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of the Macaque brain
  - More human brain devoted to vision than anything else

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of the Macaque brain
  - More human brain devoted to vision than anything else

Slide: J. Hays

The beginning of computer vision ...

Computer vision assigned as a summer project to student research assistants at MIT in 1966:

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group July 7, 1966

Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

"The final goal is OBJECT IDENTIFICATION which will actually name objects by matching them with a vocabulary of known objects."

Reality check

Rick Szeliski, 2010

[...] despite all of the 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 it so difficult?

- It is an **inverse problem**: recover some unknowns given insufficient information to fully specify the solution
- In general, to disambiguate between solutions we resort to
  - physics-based model, e.g., geometry, light
  - probabilistic models
- These are two different schools that are typically in conflict
- Current research moves towards unifying both

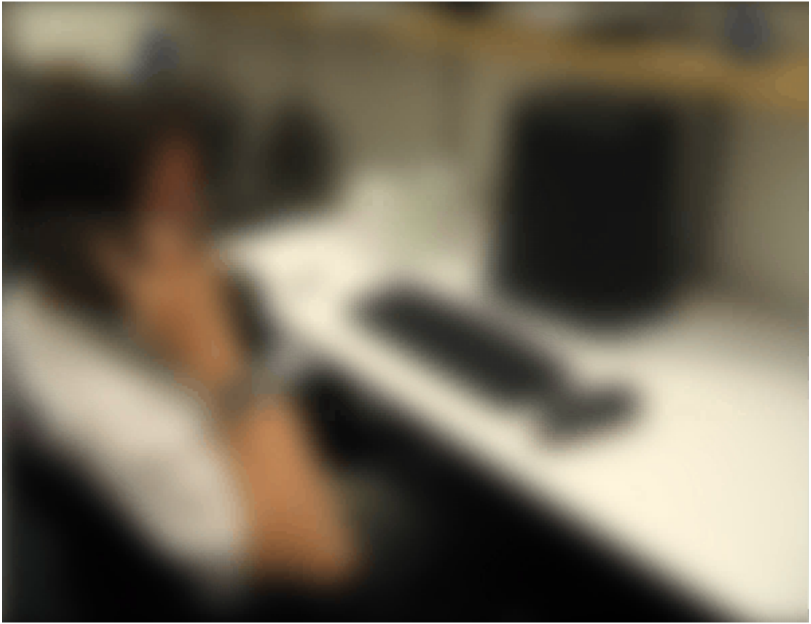
Challenges: inherent ambiguity of projection

Many possible scenes can give rise to the same image.

Challenges: local ambiguity of patches

Slide: Fei-Fei, Fergus and Torralba

Challenges: local ambiguity of patches



Images: Rob Fergus and Antonio Torralba

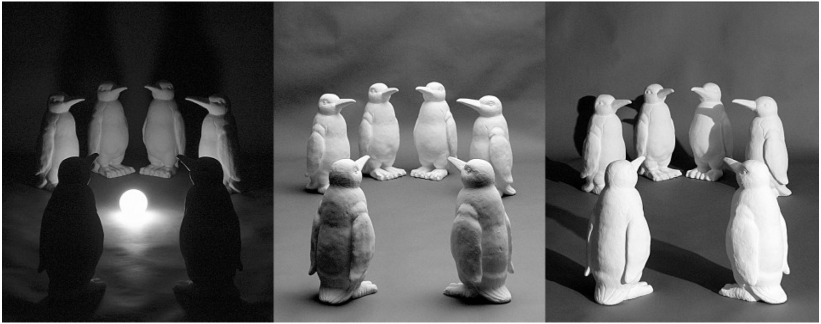
Challenges: local ambiguity of patches

Images: Rob Fergus and Antonio Torralba

Challenges: viewpoint variation

The same or different statues?

Challenges: illumination variation

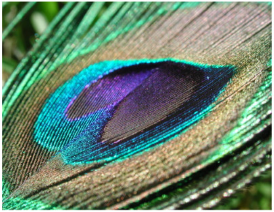


Challenges: complex materials and geometry









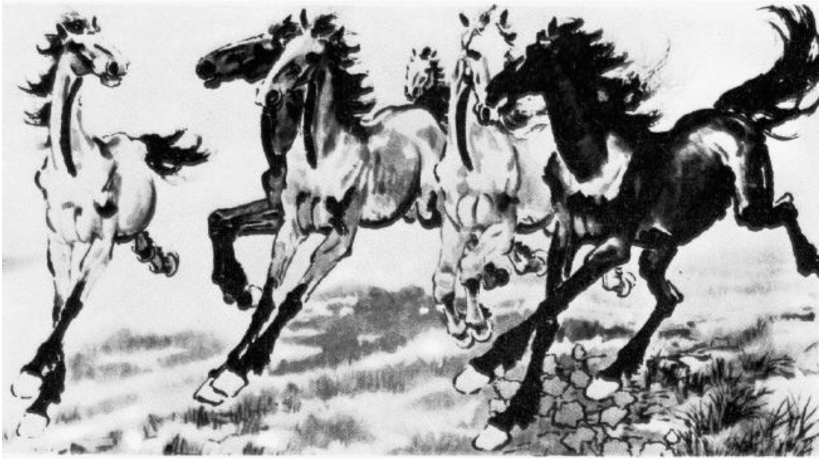

Slide: S. Lazebnik

Challenges: intra-class variation

Challenges: scale variation

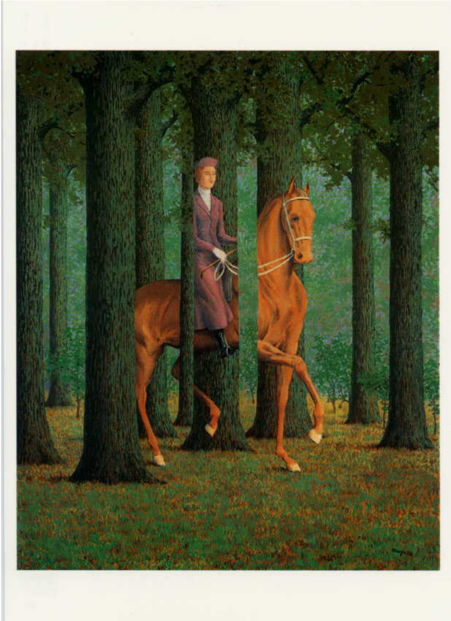
Slide: Fei-Fei, Fergus and Torralba, Image (c) Apple inc.

Challenges: deformation



Slide: Fei-Fei, Fergus and Torralba, Painting: Xu, Beihong 1943

Challenges: occlusion



Slide: Fei-Fei, Fergus and Torralba, Painting: Magritte, 1957

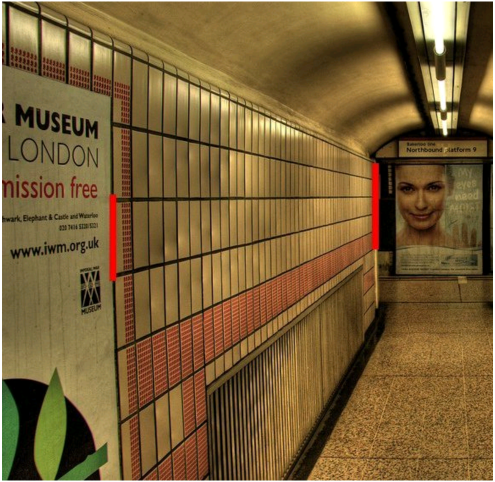
Challenges: complex background

Challenges: motion

Is human vision always “better”?

- The notion of what is hard actually differs somewhat between humans and computers.
- One reason is that a human cannot shut down certain (very useful) automatic systems.
- As a result, computers are better at certain things, in particular absolute measurements in images.

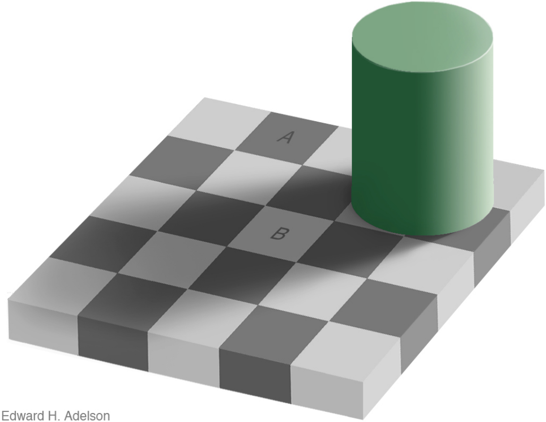
Which of the red lines is longer?



Depth perception is automatic

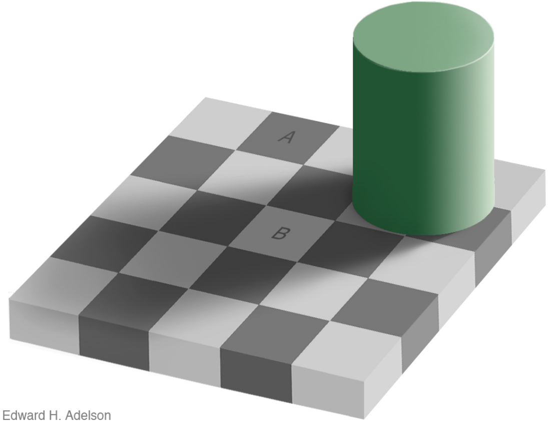
Figure: Walt Anthony 2006

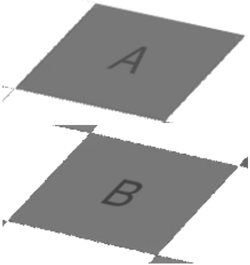
Which square is brighter, A or B?



Edward H. Adelson

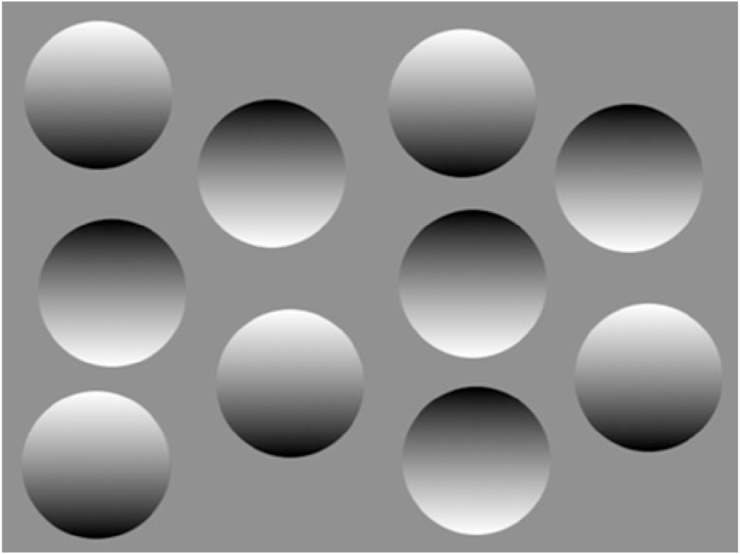
Which square is brighter, A or B?





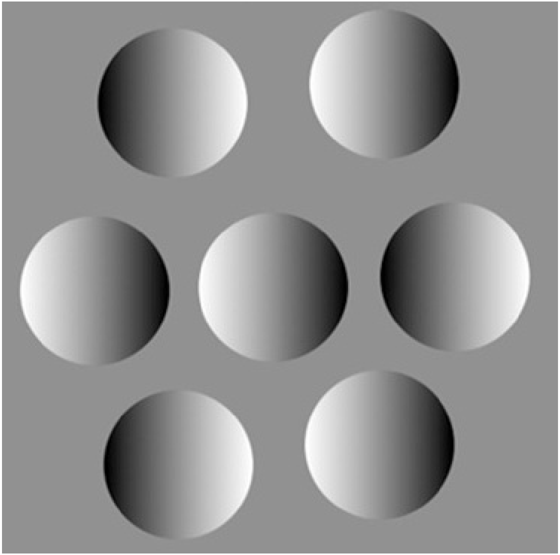
Color constancy is automatic

Shape from shading



Do the spheres pop in or pop out?

Shape from shading



What about now?

Human vs. computer vision

- Human vision has many automatic systems for low-level vision, which work effortlessly.
- In most cases, they cannot be turned off, but in case of ambiguities, one can sometimes concsiously “flip” between interpretations.
- The actual methods which are employed in the brain are usually not well understood, if at all.
- In image analysis, we study algorithmic emulations for some of these systems which are efficient and work well.

Computer vision: detective work?

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

Depth cue: linear perspective

Depth cue: aerial perspective

Depth ordering cue: occlusion

Image (c) aeon magazine

Shape cue: texture gradient

Shape and lighting cue: shading

Position and lighting cue: shadows

Grouping cue: similarity (color, texture, proximity)

So what will we we learn in this lecture?

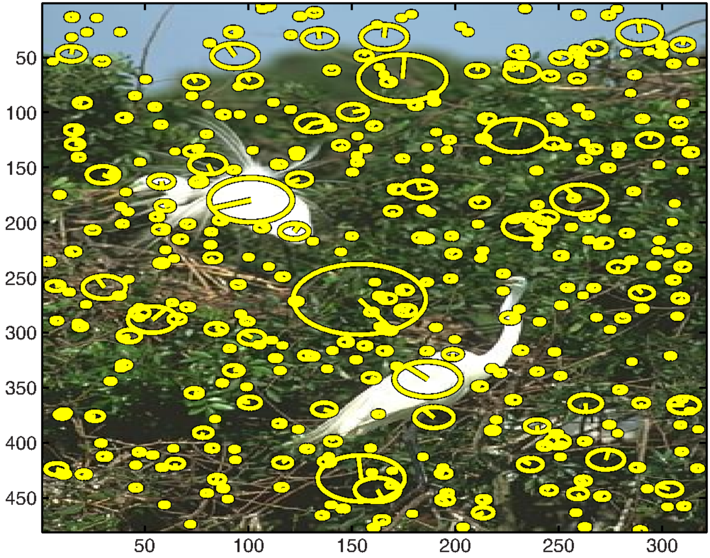
Lecture: Image Analysis and Computer Vision

Basic introduction, based on material in books
4+2 SWS, 9 ECTS, Target group: late Bachelor, Master
Focus on practical problem solving and exercises (MATLAB)

- Image processing and features
- Image formation and camera models
- Structure from motion (aka 3D point clouds from images)
- Tracking and dense correspondence
- Machine learning and recognition

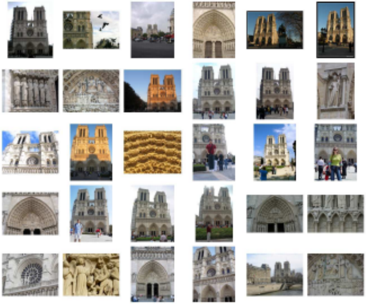
... to be continued next semester

Part I: Image processing and features



Project: SIFT and your very own panorama stitching app

Part III: Structure from Motion

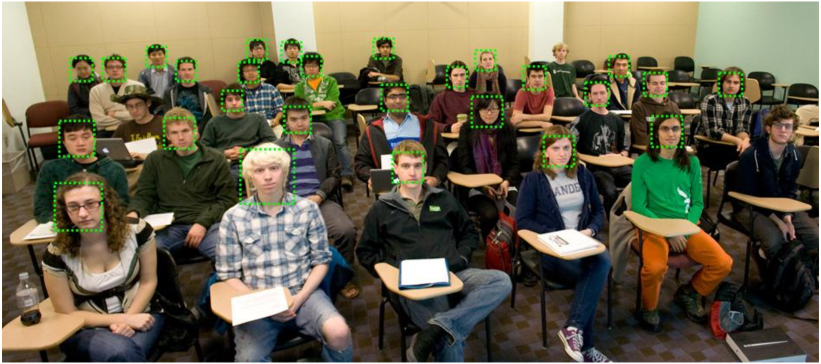


Project: generate a 3D point cloud from your own photographs

Part IV: Tracking and dense correspondence

Project: implement a method for dense optical flow

Part IV: Machine learning and recognition



Project: implement a simple face detector

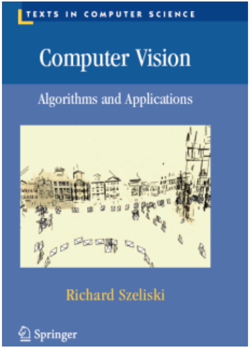
A note on terminology

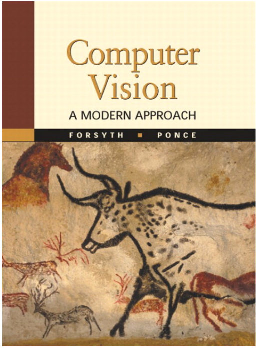
- **Image Analysis: “low to mid level”, “early” vision**
  - Images to images (image processing/enhancement)
  - Image labeling (segmentation, depth, motion, features)

- **Computer Vision: “high level”, scene interpretation**
  - Images to models (3D, texture, lighting)
  - Recognition (objects, activities ...)

Literature

- Two recommended main reference books:
  - Rick Szeliski, *Computer Vision: Algorithms and Applications*
  - David Forsyth and Jean Ponce, *Computer Vision: A Modern Approach*





- Check for more references and links on the course web page.

Getting help and more information

- We are on **ILIAS**, please sign up in order to receive organisatory e-mails
- Scripts are available on ILIAS as well
- There will also be a forum where you can ask questions and discuss exercises, lectures ...
- From time to time, myself or Ole will check it out and see if we have time to write some quick answers

Exercises

Teaching assistant: Ole Johannsen

- Usually one exercise sheet per week (only a quick one or maybe none when projects are due)
- Turnin on Tuesdays (late policy: -30% points each day)
- Solutions will be discussed Friday in the exercise group

Grading

- Programming stuff yourself is a vital part of learning computer vision (or indeed, anything else in CS)
- Therefore, programming projects will count towards final grade - in particular, they are necessary to get the best marks.
- Scoring scale example (**not final numbers, just so you get the idea**):
  - 80 total points from exam,
  - 50 total points from projects,
  - a total of 100 of points is a 1.0,
  - 50 is minimum for success.

Summary: Computer Vision and Image Analysis

- Computer Vision deals with difficult inverse problems
- Many different cues about a scene can be found in images, and can potentially be exploited in algorithms
- A good idea is to look at how humans gain information, and try to emulate these systems
- Lots can be learned about human perception from visual illusions
- In image analysis, one learns how to investigate images on a low level in order find information to feed into higher level algorithms
- We will start next lecture with basic image processing