CSE 6239
Computer Vision

Introduction

Dr. Sk. Md. Masudul Ahsan

Professor, Dept of CSE, KUET

Computer Vision

Make computers understand images and video.

- Computing properties of the 3D world from visual data *(measurement)*
- Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*

What kind of scene?

Where are the cars?

How far is the building?

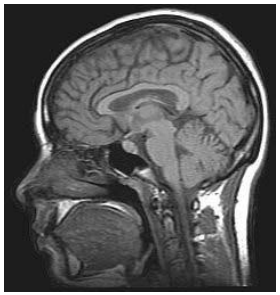
…

What is Computer Vision?

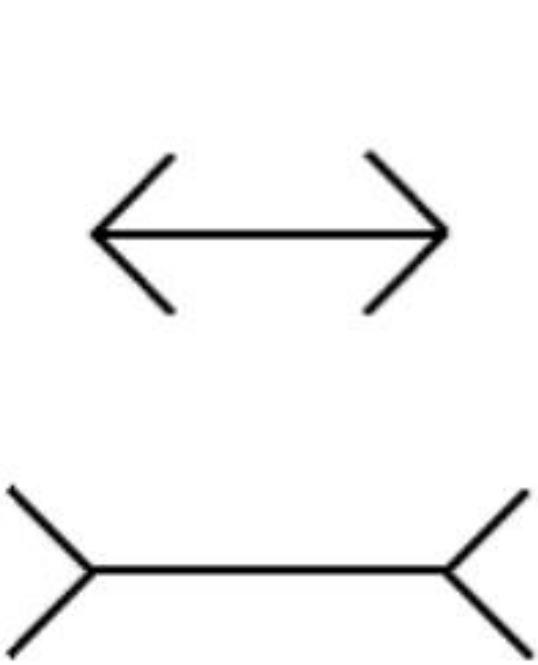
- **Computer vision** is the science and technology of machines that see.

- Concerned with the theory for building artificial systems that obtain information from images.

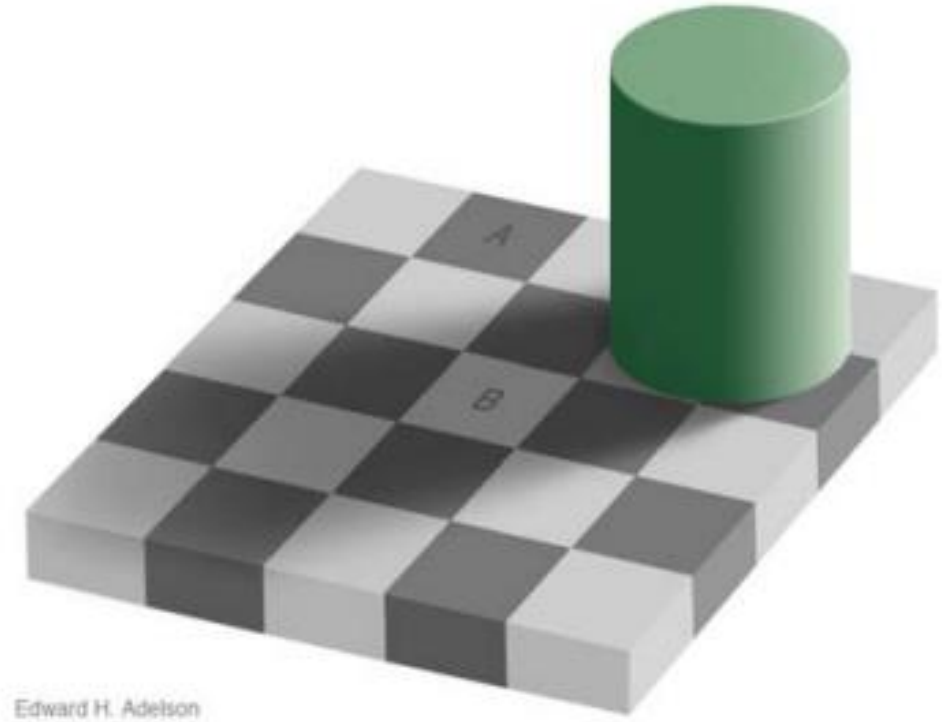
- The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner



Vision is Challenging



(a)



(b)

Important note:

In general, computer vision does not work

(except in certain situations/conditions)

What is an image?

We can think of an **image** as a function, *f*, from R^2 to R:

- $f(x, y)$ gives the **intensity** at position (x, y)
- Realistically, we expect the image only to be defined over a rectangle, with a finite range:
  - $f: [a,b]\mathrm{x}[c,d] \rightarrow [0,1]$

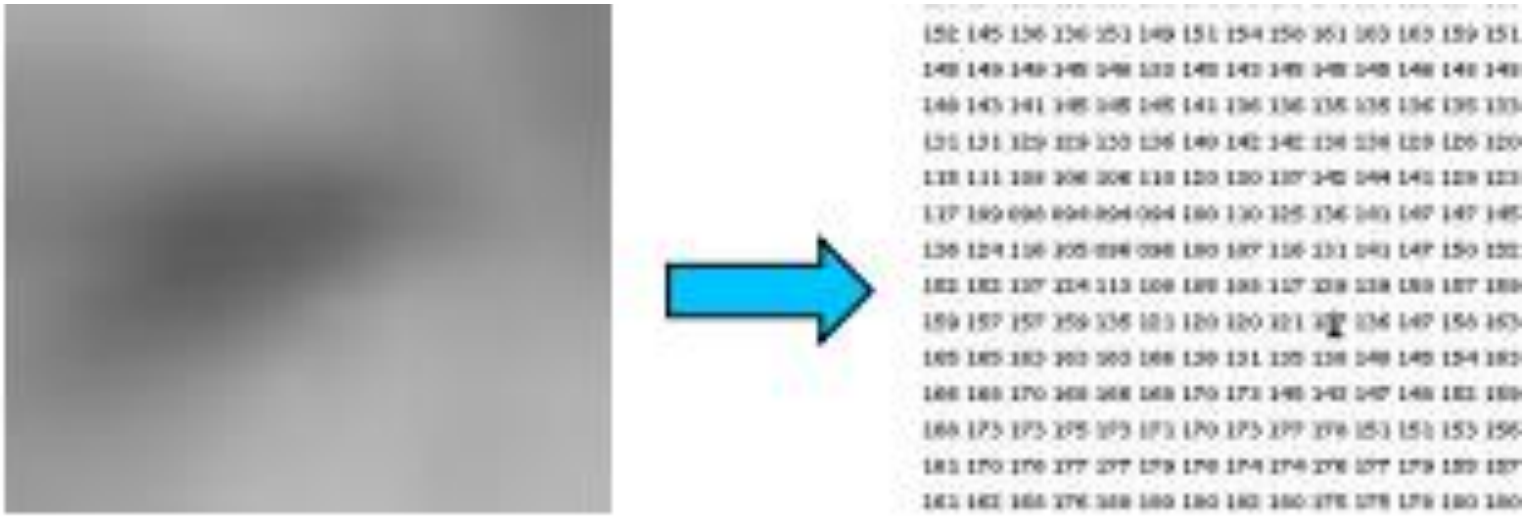
A color image is just three functions pasted together. We can write this as a "vector-valued" function:

$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

Digital Image

A digitized image is one in which:

- Spatial and grayscale values have been made discrete.
- Intensities measured across a regularly spaced grid in x and y directions are sampled to
  - 8 bits (256 values) per point for black and white,
  - 3x8 bits per point for color images.
- Stored as a 2D arrays of gray-level values. The array elements are called pixels and identified by their x, y coordinates.



Every Picture Tells a Story

Black and white photo by H Roger-Viollet of a train accident at La Gare Montparnasse station in Paris on 22 October, 1895 when engine 120-721 failed to stop at the platform, went through a first-floor window and crashed down onto the street.

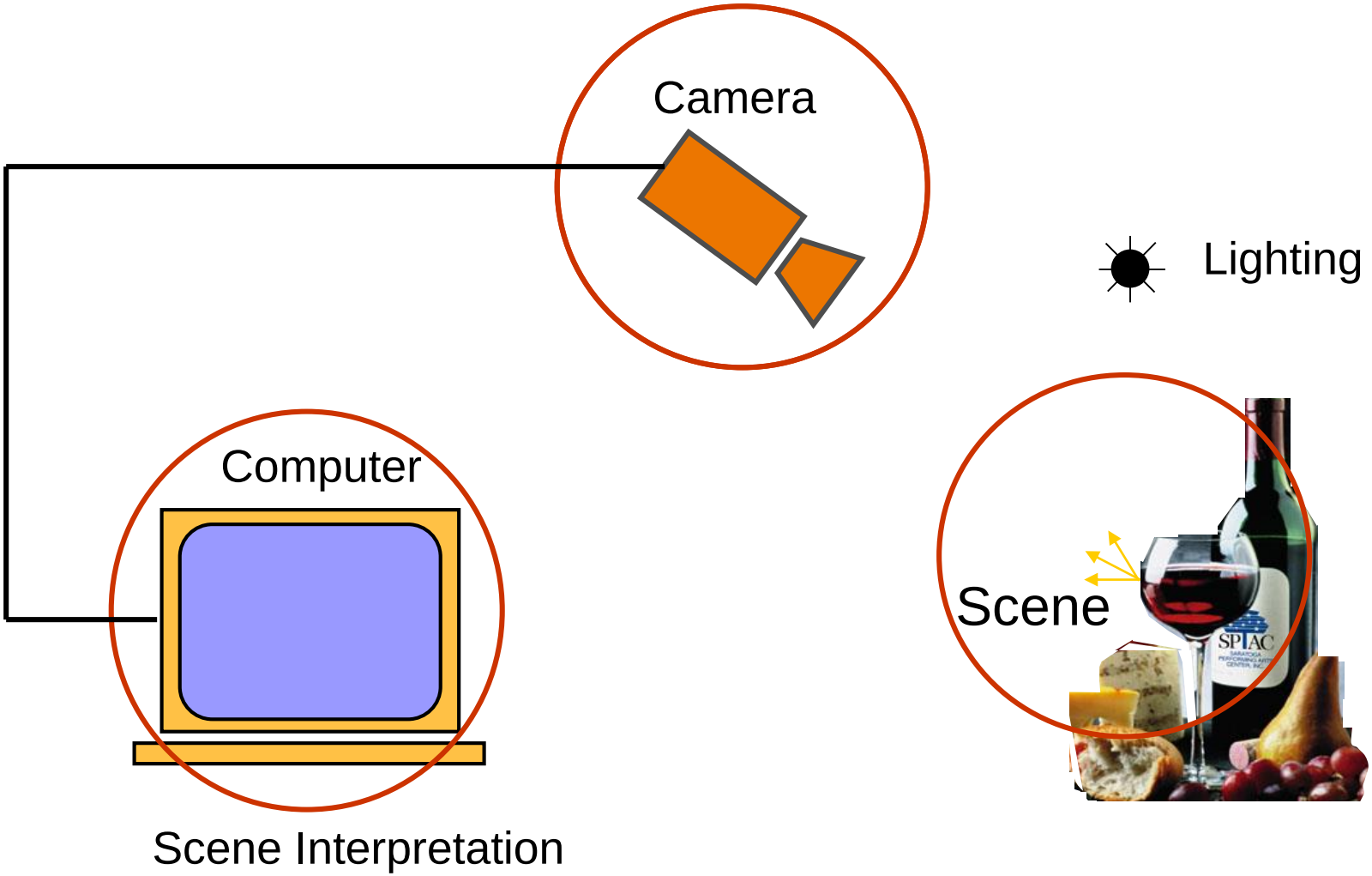
Goal of computer vision is to write computer programs that can interpret images.

Can computers match (or beat) human vision?

Yes and no (but mostly no!)

- humans are much better at “hard” things
- computers can be better at “easy” things

Components of a computer vision system



Srinivasa Narasimhan's slide

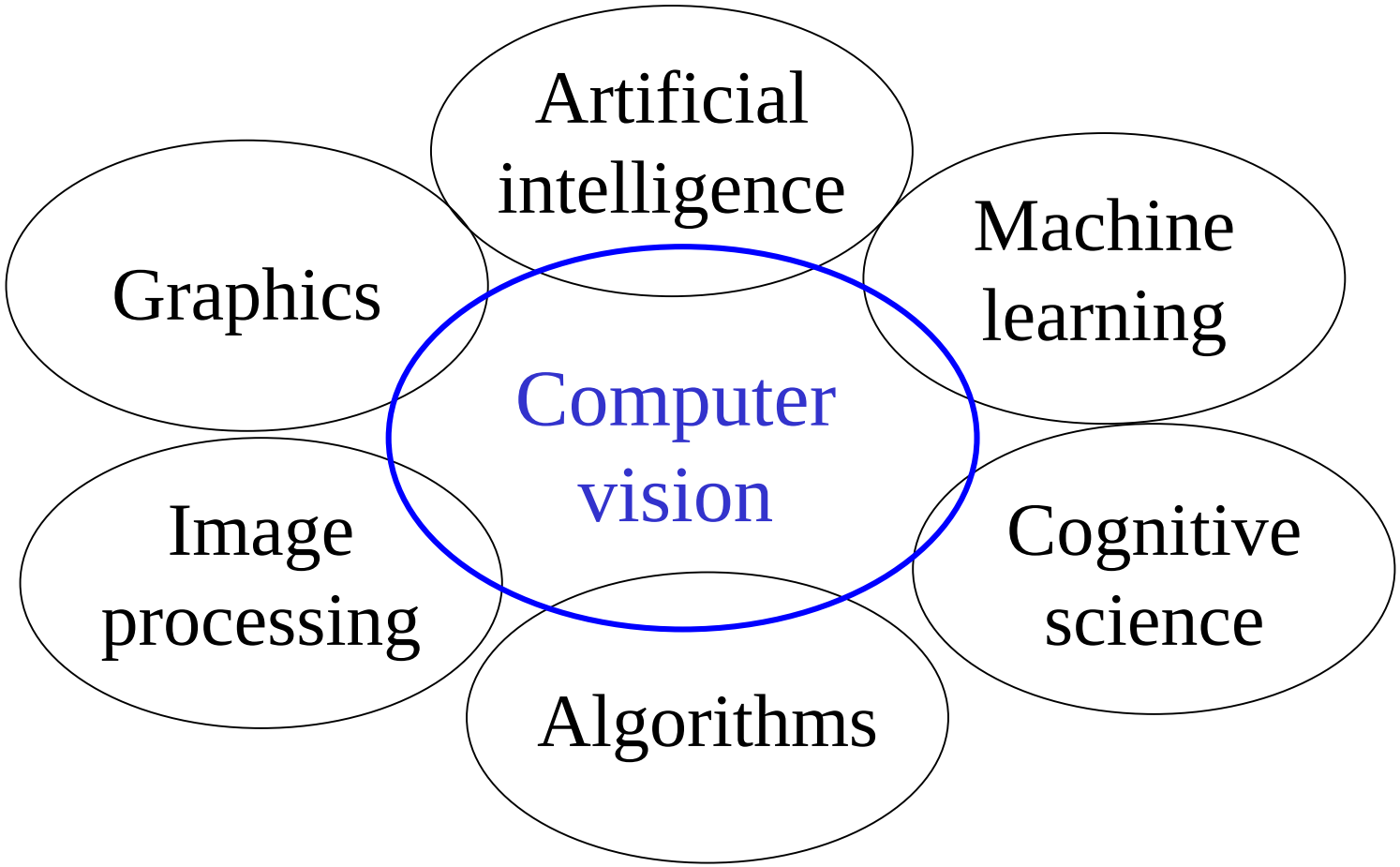
Computer vision vs human vision

What we see

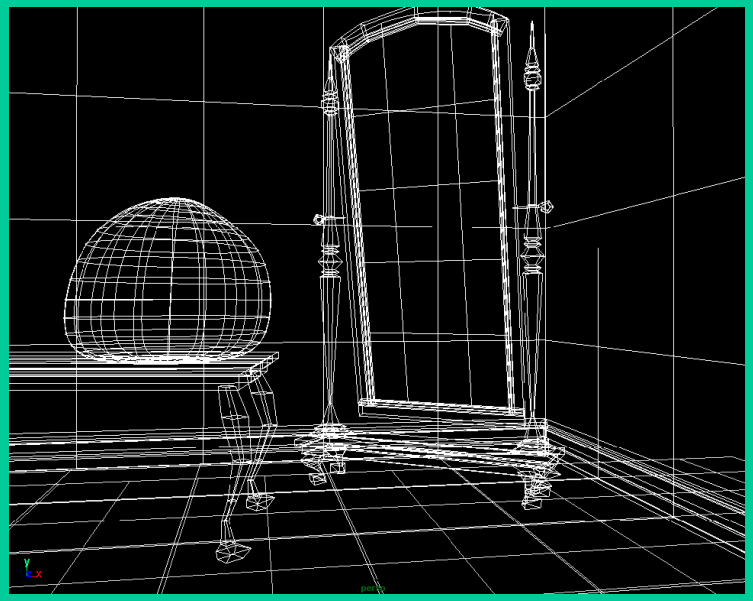
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

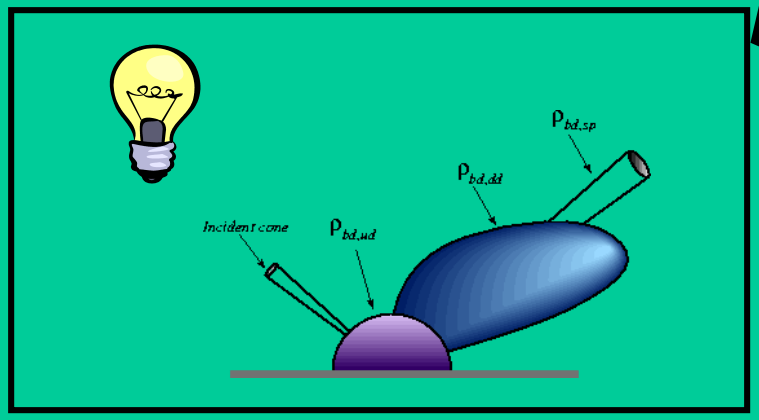
Related disciplines



What is computer vision? (2D->3D)



3D geometry



physics

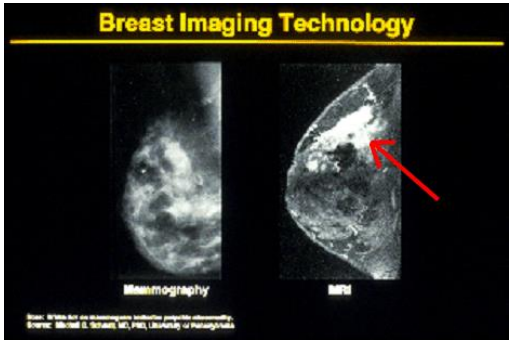
Estimation

Current state of the art

The next slides show some examples of what current vision systems can do

Why computer vision matters

Safety

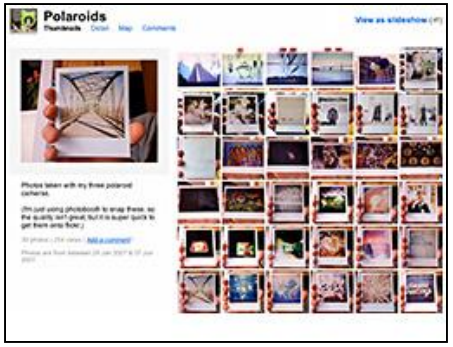


Health

Security

Comfort

Fun



Access

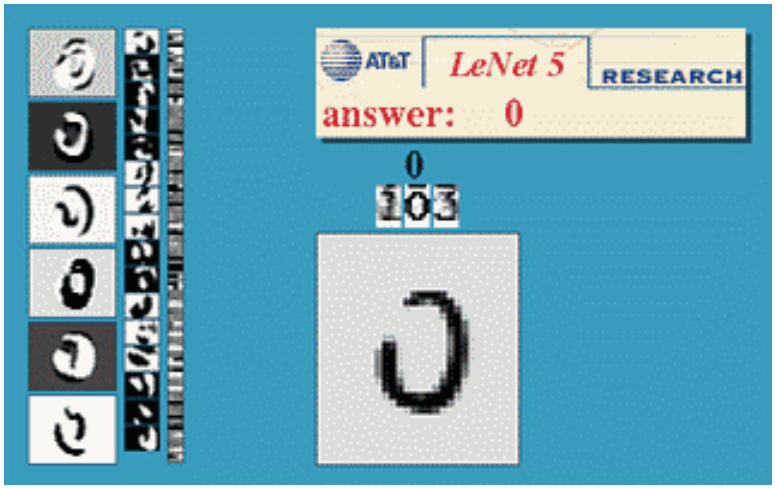
Current state of the art

The next slides show some examples of what current vision systems can do

Optical character recognition (OCR)

Technology to convert scanned docs to text

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

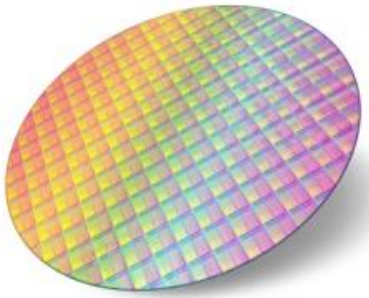


Digit recognition, AT&T labs
http://www.research.att.com/~yann/

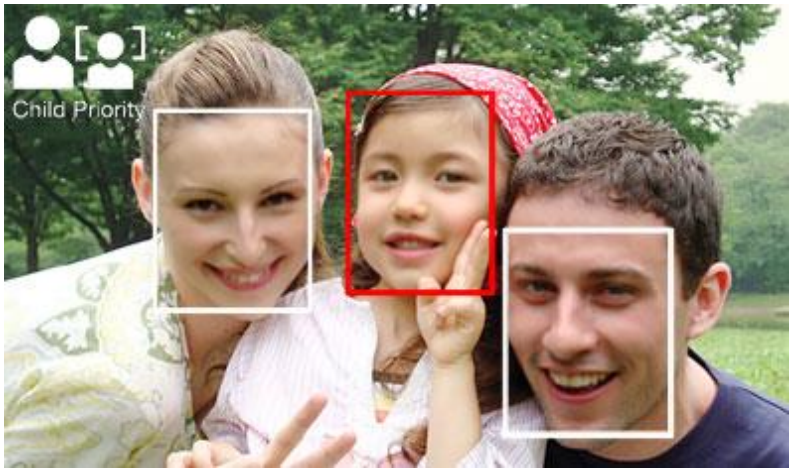
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Machine vision

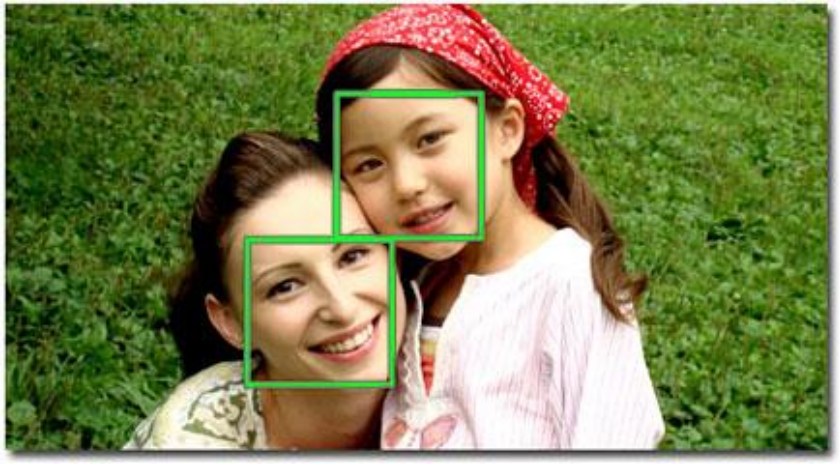
Automated visual inspection



Face detection



Age recognition

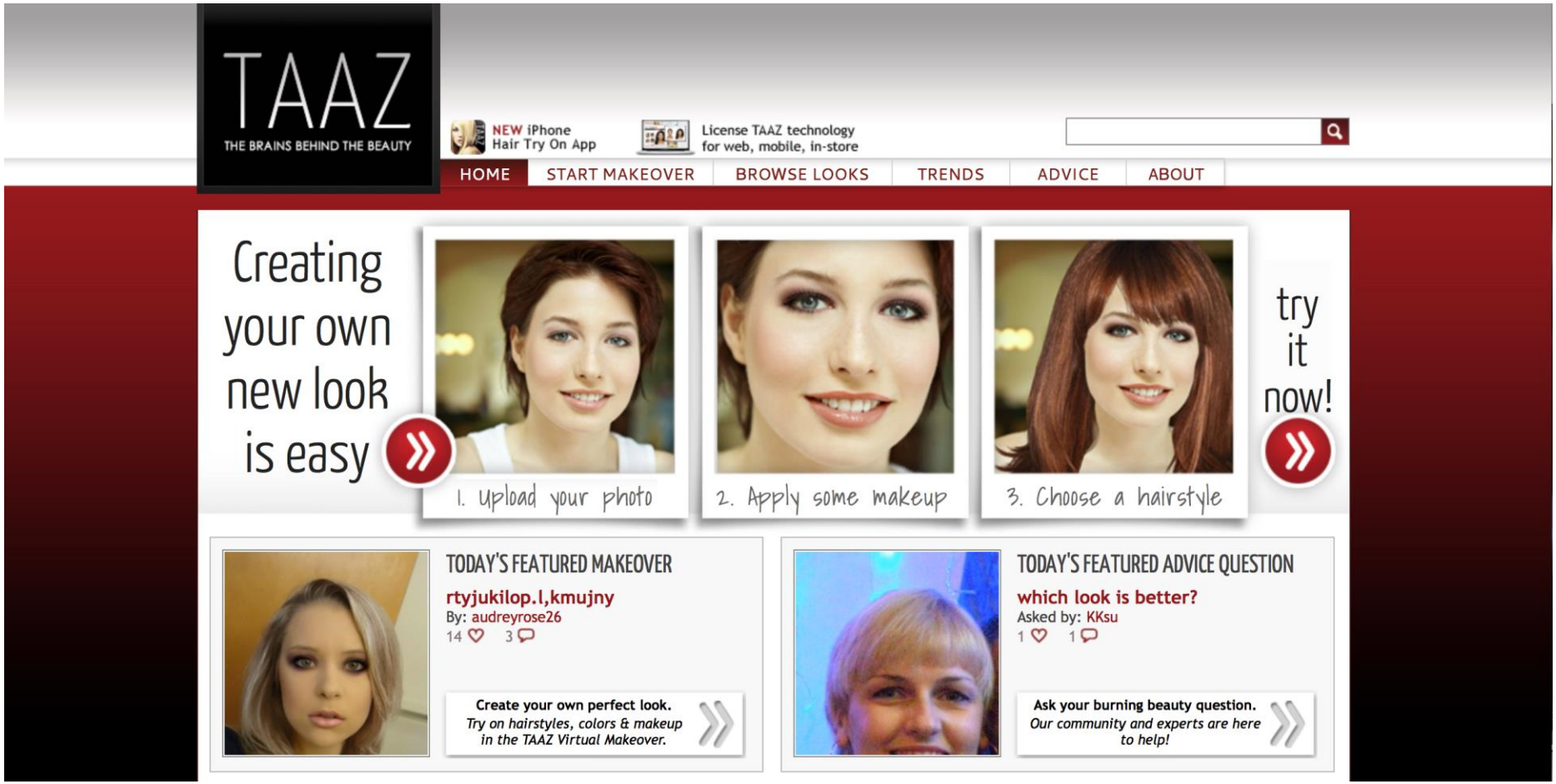


Smile recognition

Many new digital cameras now detect faces

- Canon, Sony, Fuji, …

Face makeovers



Object Recognition

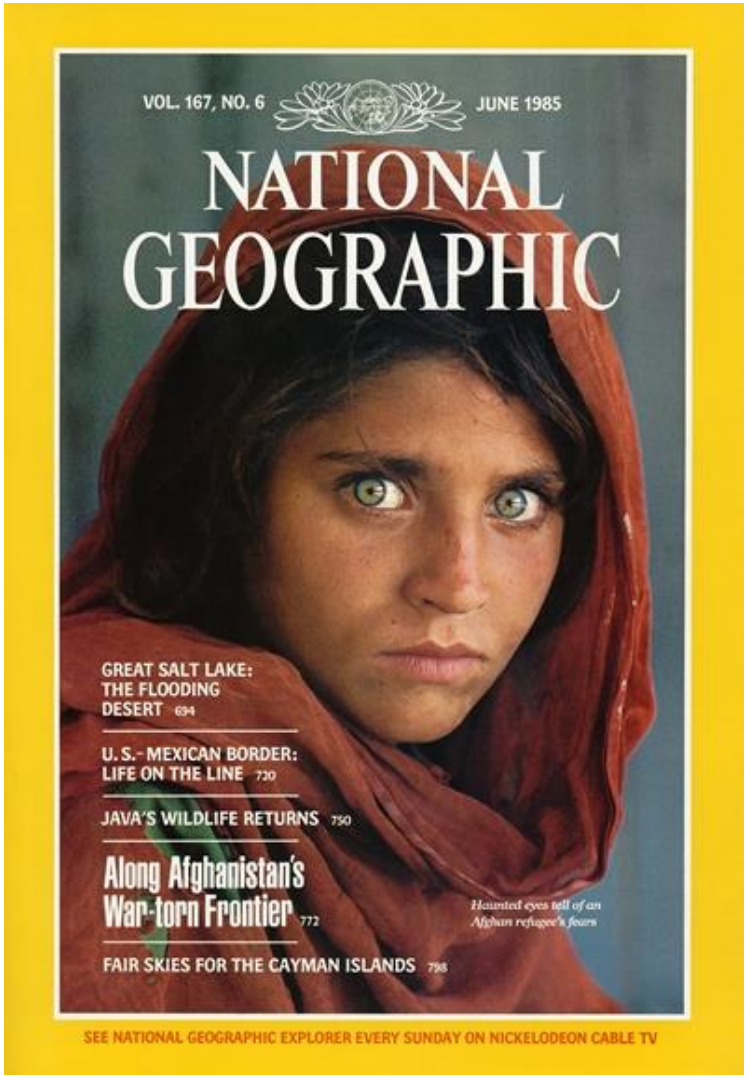
Toshiba Tech IS-910T 2013

DataLogic LaneHawk LH4000 2012

3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

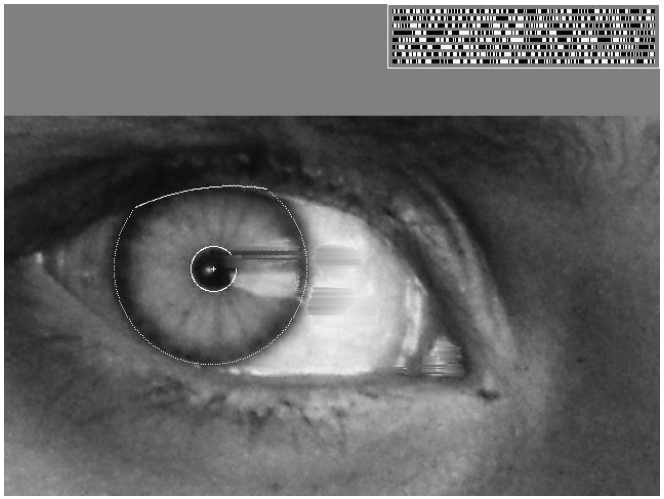
Face recognition

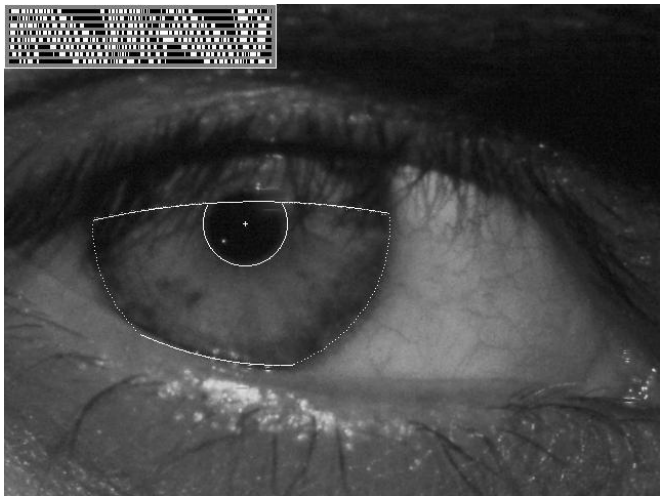


Who is she?

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story





Object recognition (in mobile phones)

This is becoming real:

- Lincoln Microsoft Research
- Point & Find, Nokia

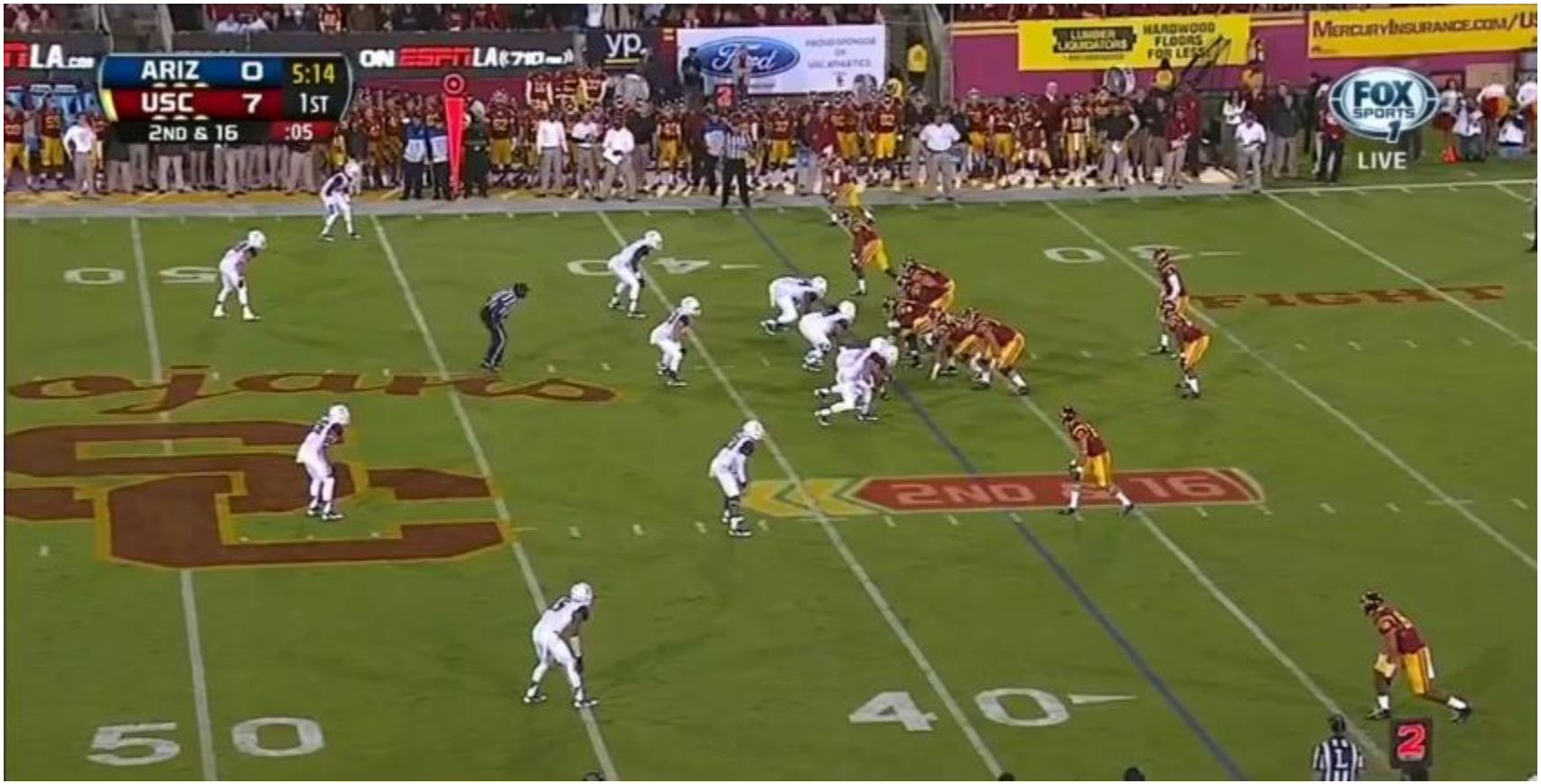
Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic
Click here for interactive demo

Sports



Sportvision first down line
Nice [explanation] on www.howstuffworks.com

Smart cars

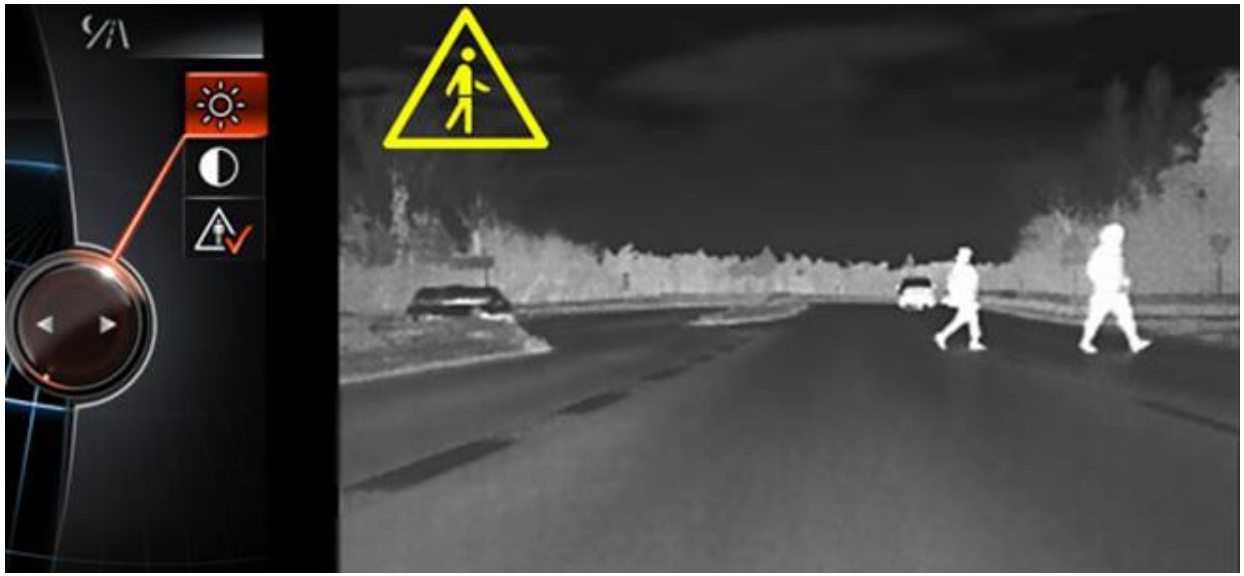
Slide content courtesy of Amnon Shashua

Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models
- By 2010: 70% of car manufacturers.
- Video demo

BMW 5 series

BMW night vision



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!

Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET
Camera tracking a crowd, based on this work.

Virtual Fitting

2015

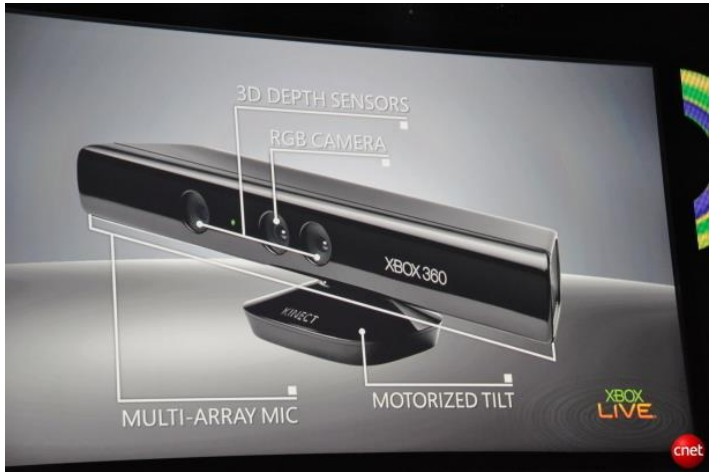
Interactive Games: Kinect

Object Recognition:
http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o

Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg

3D: http://www.youtube.com/watch?v=7QrnwoO1-8A

Robot:
http://www.youtube.com/watch?v=w8BmgtMKFbY



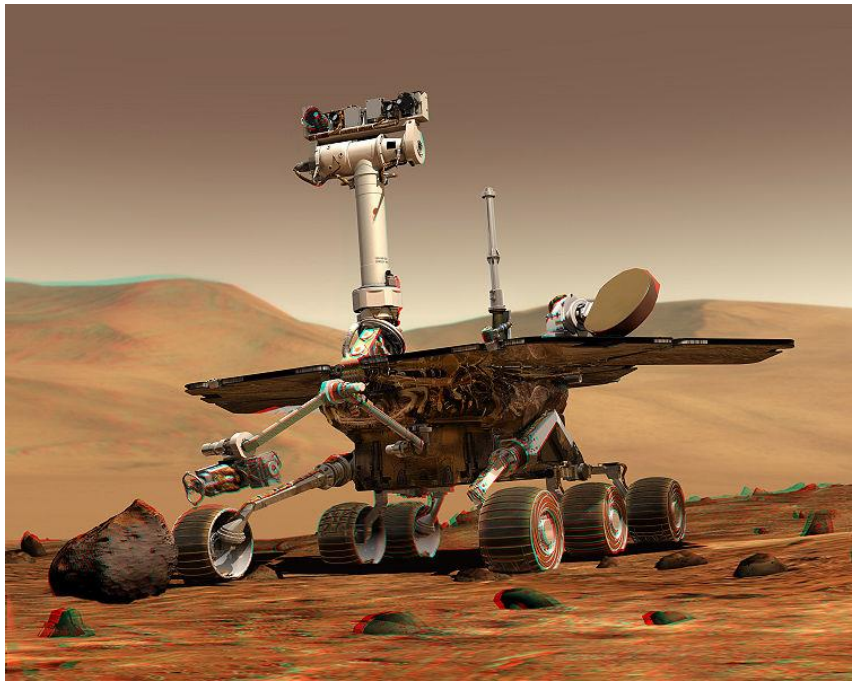
Vision in space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

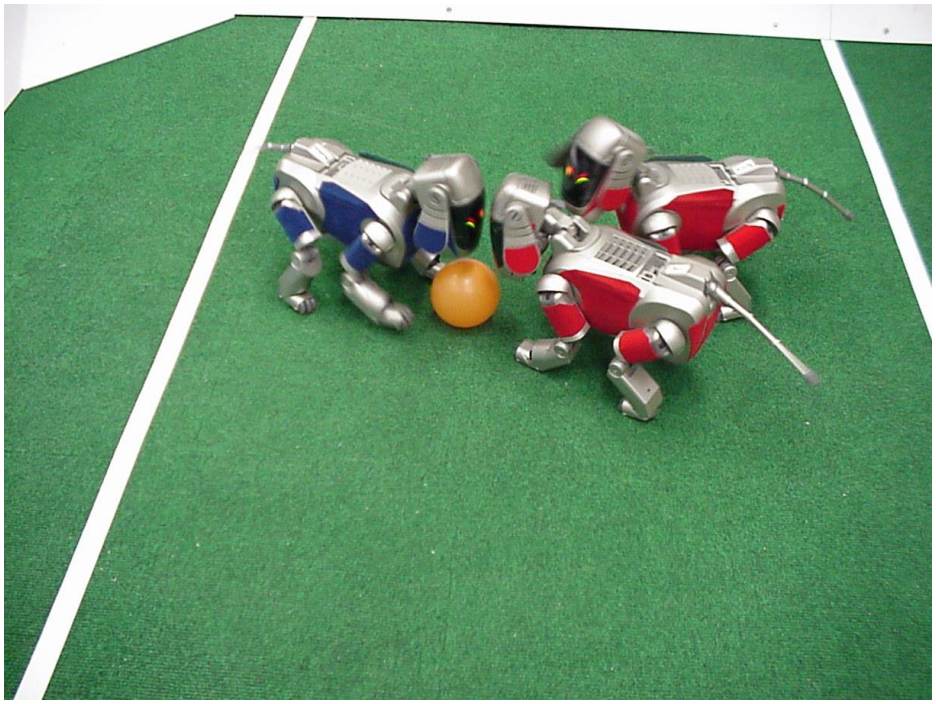
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Robotics

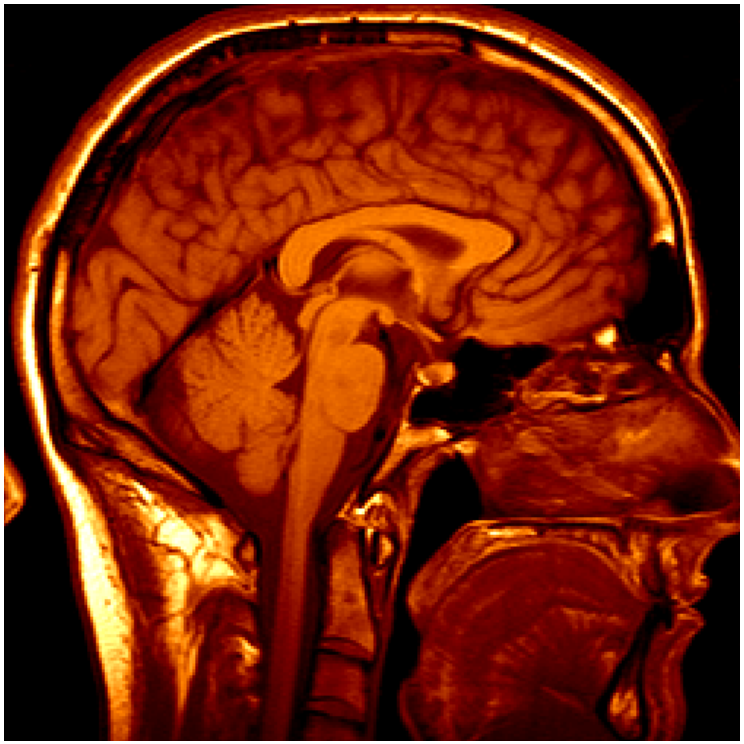


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Challenges

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, F

Challenges: illumination

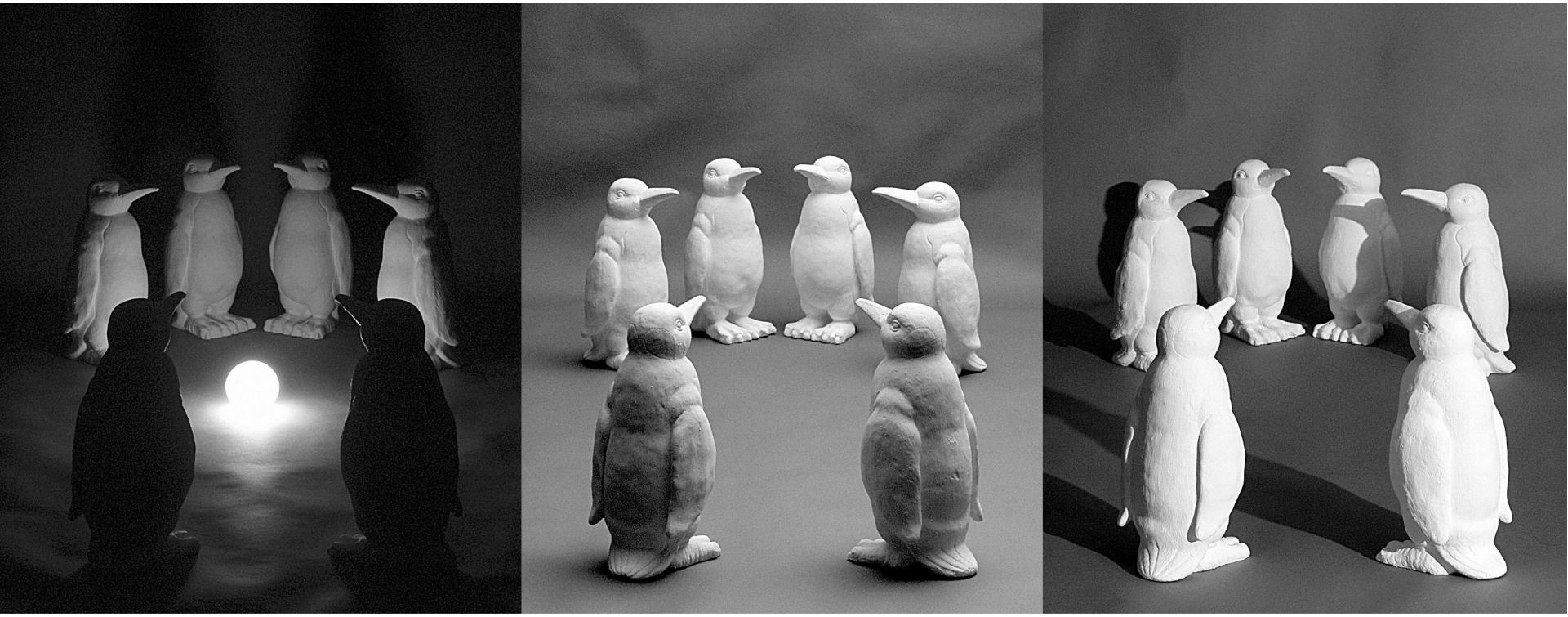
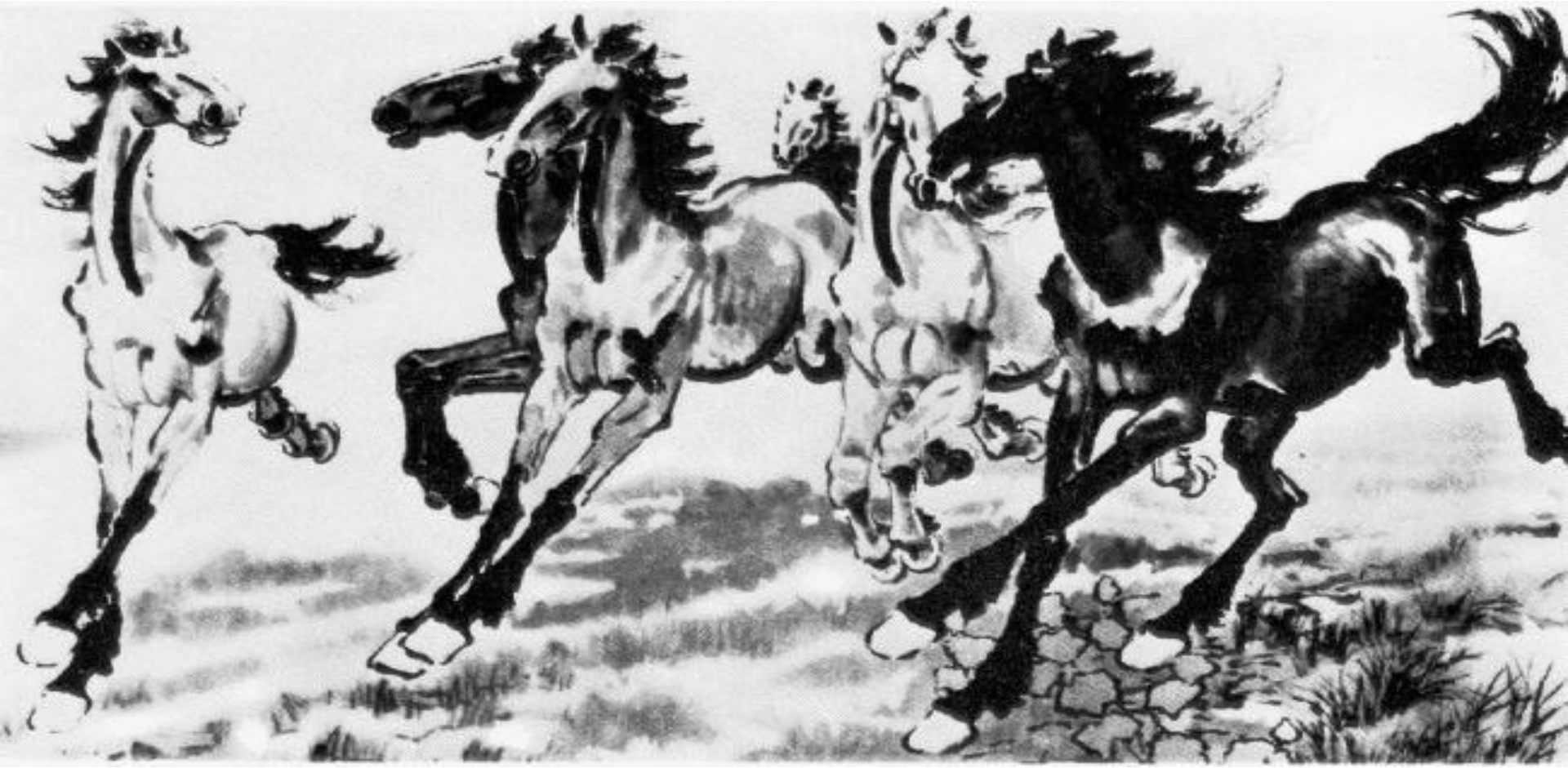


image credit: J. Ko

Challenges: scale

slide credit: Fei-Fei

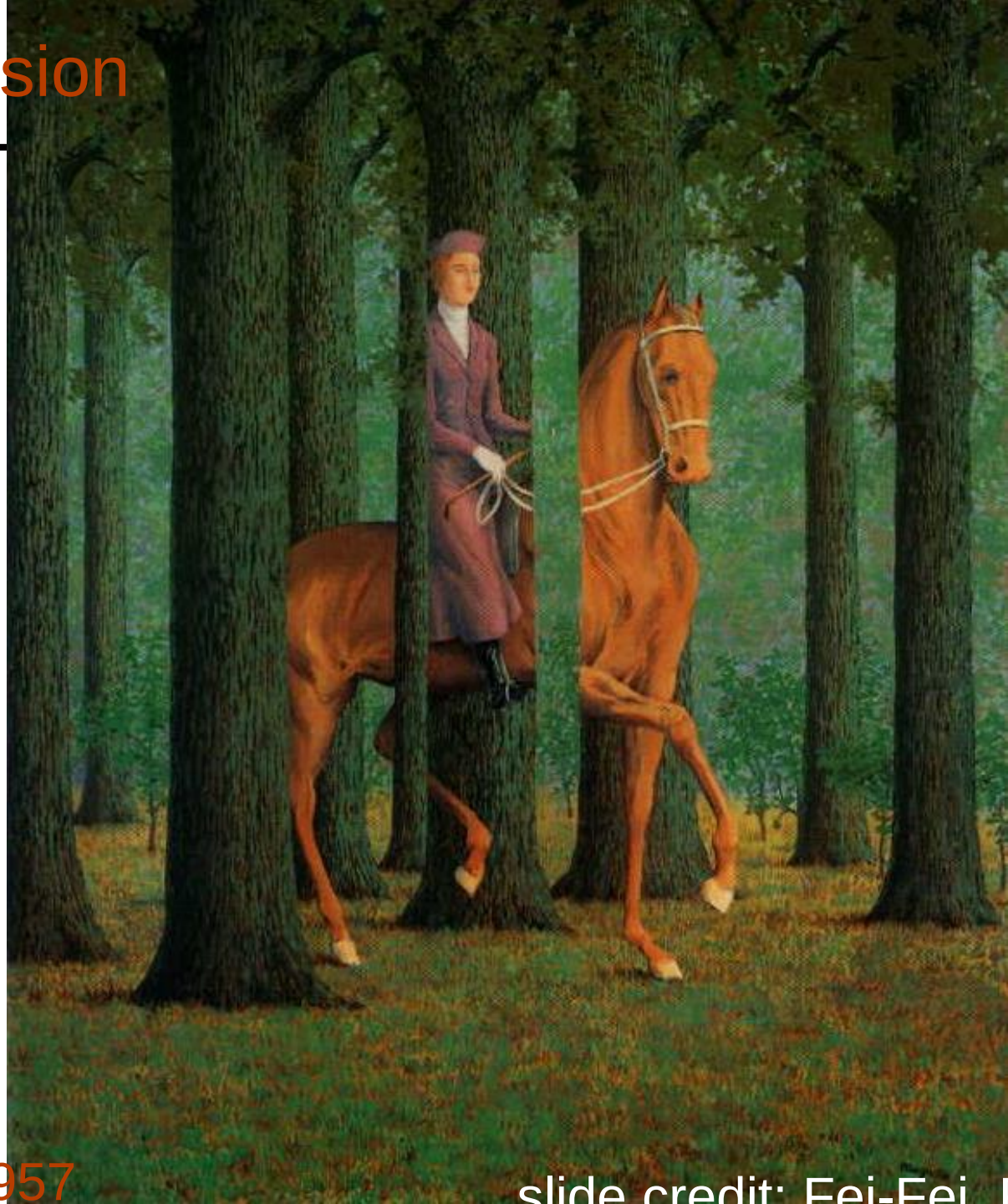
Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei

Challenges: occlusion

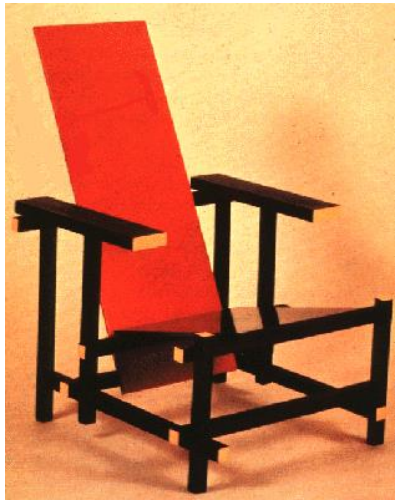


Magritte, 1957

slide credit: Fei-Fei

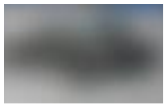
Challenges: background clutter

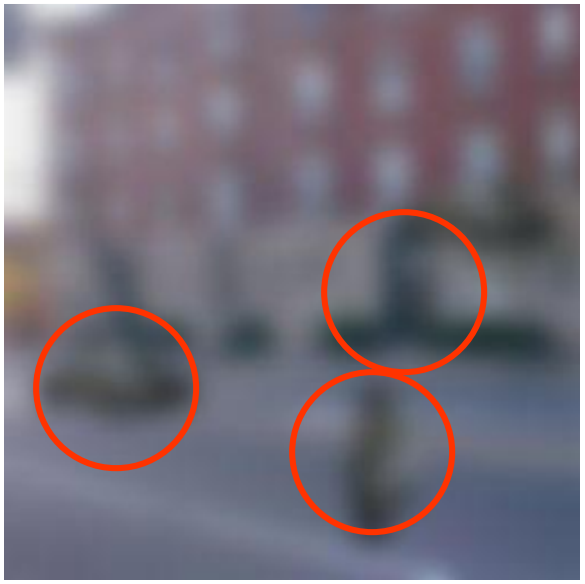
Challenges: object intra-class variation



slide credit: Fei-Fei, F

Challenges: local ambiguity





slide credit: Fei-Fei, F

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!

Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

- Possible solutions
  - Bring in more constraints (or more images)
  - Use prior knowledge about the structure of the world
- Need both exact measurements and statistical inference!

It’s a good time to do computer vision

Industry aggressively hiring CV faculty from universities

(slide already out of date by at least 3 CMU faculty)

NYU

UW

USC

UW

NYU

Dropbox

UCSD

Columbia

UBER

Oculus VR

Google

CMU CMU

CMU CMU CMU

CMU Stanford Toronto UW

aggressively hiring CV graduates, or even students!

facebook research | Research Areas | Publications | People | Programs | Downloads | Careers | Blog

JULY 21, 2017

Advancing computer vision technologies at CVPR 2017

By: Facebook Research

Google Research Blog

The latest news from Research at Google

Google at CVPR 2017

Friday, July 21, 2017

Topics to be covered

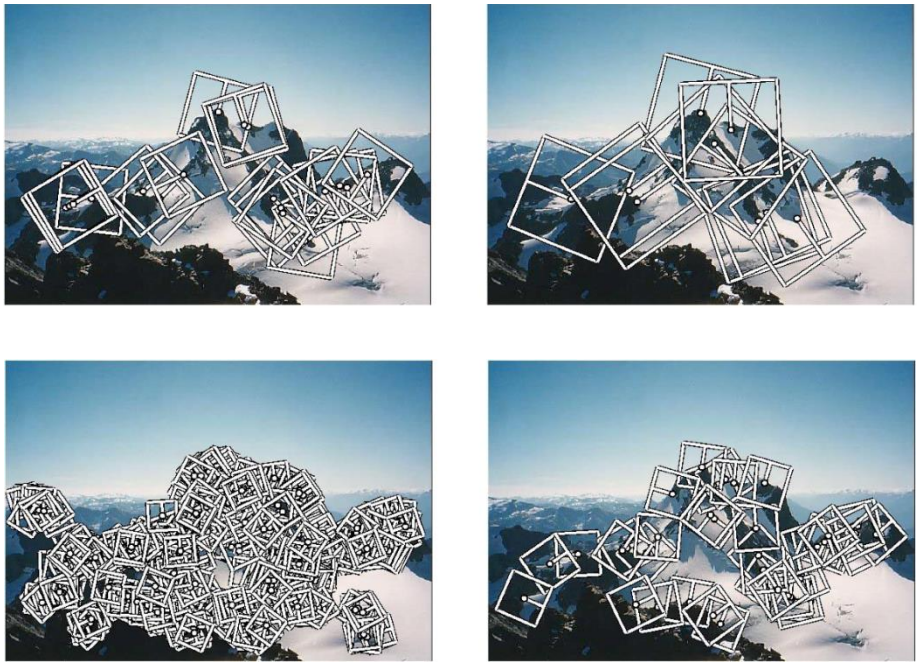
Image processing:

- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.

Topics to be covered

Feature detection and correspondences:

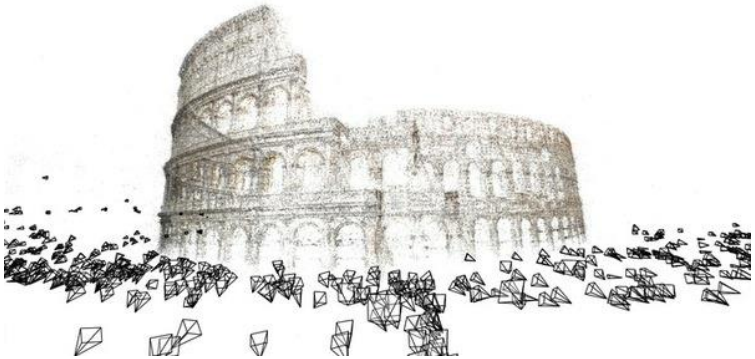
- Corner detection.
- SIFT et al.
- Feature descriptors.
- RANSAC.



Topics to be covered

Transformations and geometry:

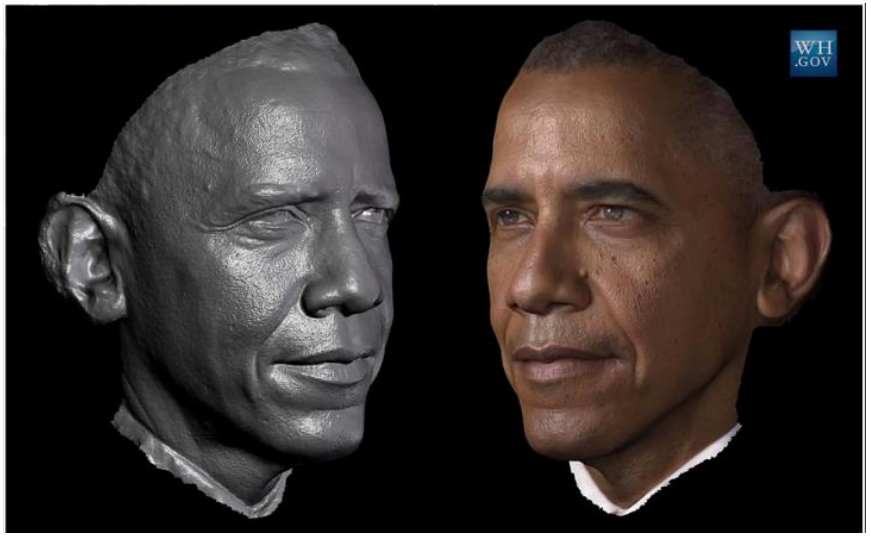
- Homographies and image alignment.
- Camera models.
- Fundamental matrix.
- Epipolar geometry and stereo.
- Structure from motion.



Topics to be covered

Physics-based vision:

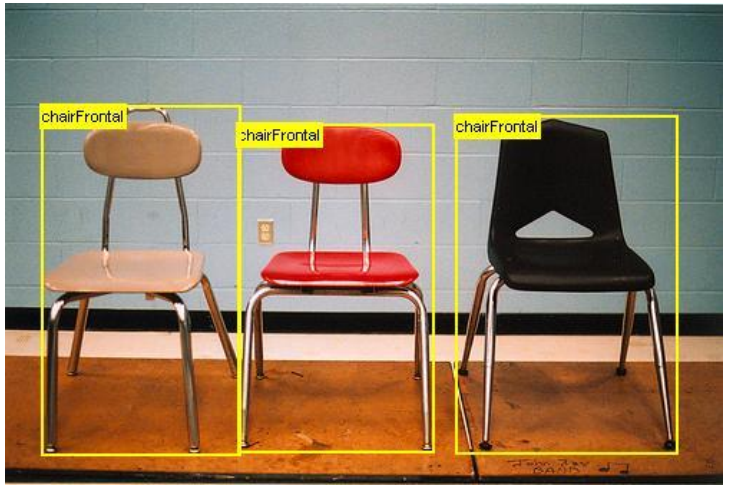
- Reflectance and image formation.
- Radiometry.
- Shape from shading.
- Photometric stereo.
- Color.



Topics to be covered

Objects, faces, and learning:

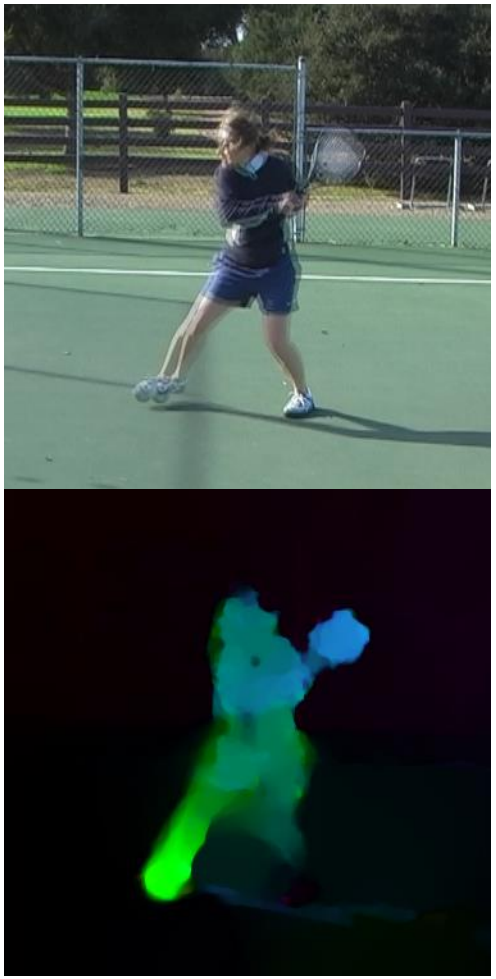
- Basics of probability.
- K-means, KNN, PCA, SVM.
- Bag of words.
- Viola-Jones face detection.
- Perceptron, backpropagation.
- Convolutional neural networks.



Topics to be covered

Dealing with motion:

- Optical flow (LK, HS).
- Image registration.
- Kalman Filtering.
- Tracking (KLT, Mean-Shift).



Grading

Based on

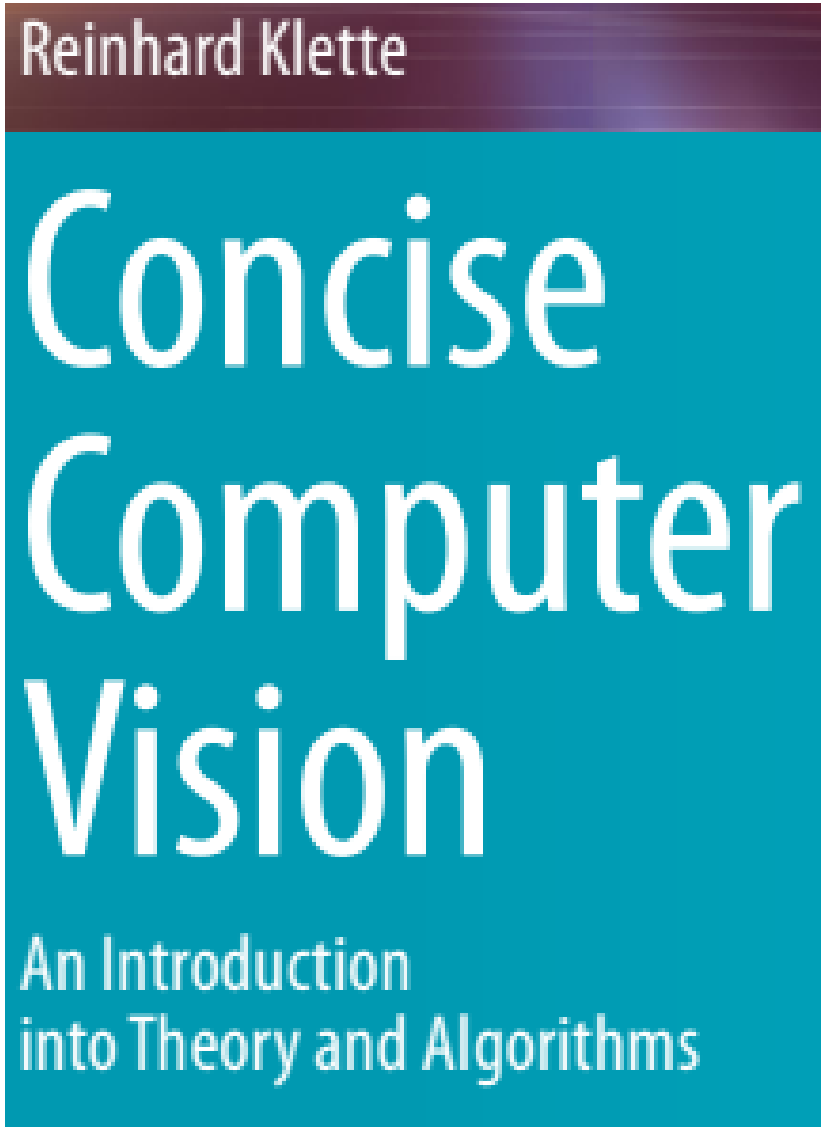
- Projects – 20%
- Final exam - 50-60%
- class participation – 5%
- 2 Seminar/workshop – 15%
- Assignments / Class tests (optional) – 0-10%

Textbook

Computer Vision: Algorithms and Applications

http://szeliski.org/Book/

Textbook


Reinhard Klette
Concise Computer Vision
An Introduction into Theory and Algorithms

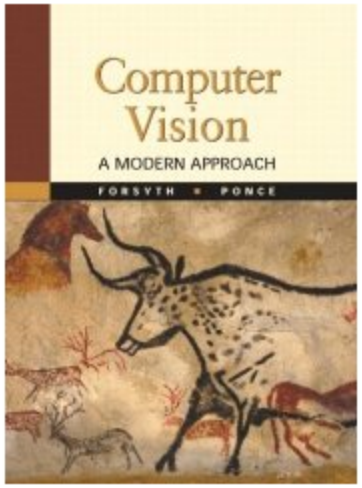
dary Text

Google

m: Computer ... ×

http://www.amazon.com/Computer-Vision-Approach-David-Forsy

Manuscript Central | cvpr ac | EE364a: Lecture Slides | KeepVid: Download ... | Other bookmarks



See larger image

Share your own customer images

Publisher: learn how customers can search inside this book.

Tell the Publisher!
I'd like to read this book on Kindle

Computer Vision: A Modern Approach (Hardcover)

by David A. Forsyth (Author), Jean Ponce (Author)

★★½☆☆ (20 customer reviews)

List Price: ~~$124.80~~

Price: $88.94 & this item ships for **FREE with Super Saver Shipping**. Details

You Save: $35.86 (29%)

Special Offers Available

In Stock.

Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered Wednesday, August 26? Order it in the next 1 hour and 3 minutes, and choose **One-Day Shipping** at checkout. Details

29 new from $74.99 **16 used** from $49.00

General Comments

Prerequisites—*these are essential*!

- Data structures
- A good working knowledge of C and C++ programming
  - (or willingness/time to pick it up quickly!)
- Linear algebra
- Vector calculus