CSE 365: Computer Vision

Prof. Mahmoud Khalil

Summer 2020

Course Team

- Instructor : Dr. Mahmoud Khalil
  - Email: mahmoud.khalil@eng.asu.edu.eg
- Lab TA: Eng. Mohamed Ashraf
  - Email: Muhammed.a.yousuf@eng.asu.edu.eg

Official Course Description

CSE 365: Computer Vision (3 Credit Hours)

Prerequisite: ECE 255

Introduction. The analysis of the patterns in visual images with the view to understanding the objects and processes in the world that generates them. Image representation and processing. Feature extraction and selection. Object recognition and probabilistic inference. Dynamic and hierarchical processing. Multi-view geometry. Projective reconstruction. Tracking and density propagation. Visual surveillance and activity monitoring. Medical imaging. Applications.

Lecture: 3 hours/week, Tutorial: 1 hour/week, Lab: 1 hour/week

Grading Scheme

- Students will be evaluated based on the following:

Component	%
Assignments (2 assignments 5 marks each and 1 Project 10 marks)	20
Quizzes (2 quizzes, 5 marks each)	10
Attendance and participation	5
Midterm exam	25
Final exam	40

Text Books

- Rafael C. Gonzalez and Richard E. Woods, **Digital Image Processing**, 4th Edition, Pearson Education, Inc. 2018., ISBN: 978-9353062989
- Richard Sceliski, **Computer Vision Algorithms and Applications**, Springer, 2011 (available online for free at: http://szeliski.org/Book/), ISBN: 978-1848829343
- Rafael C. Gonzalez , Richard E. Woods and Steven L. Eddins, **Digital Image Processing Using Matlab**, Second edition,, Pearson Education, Inc. 2009., ISBN: 978-0070702622
- **OpenCV**: https://opencv.org/

Visual Sciences

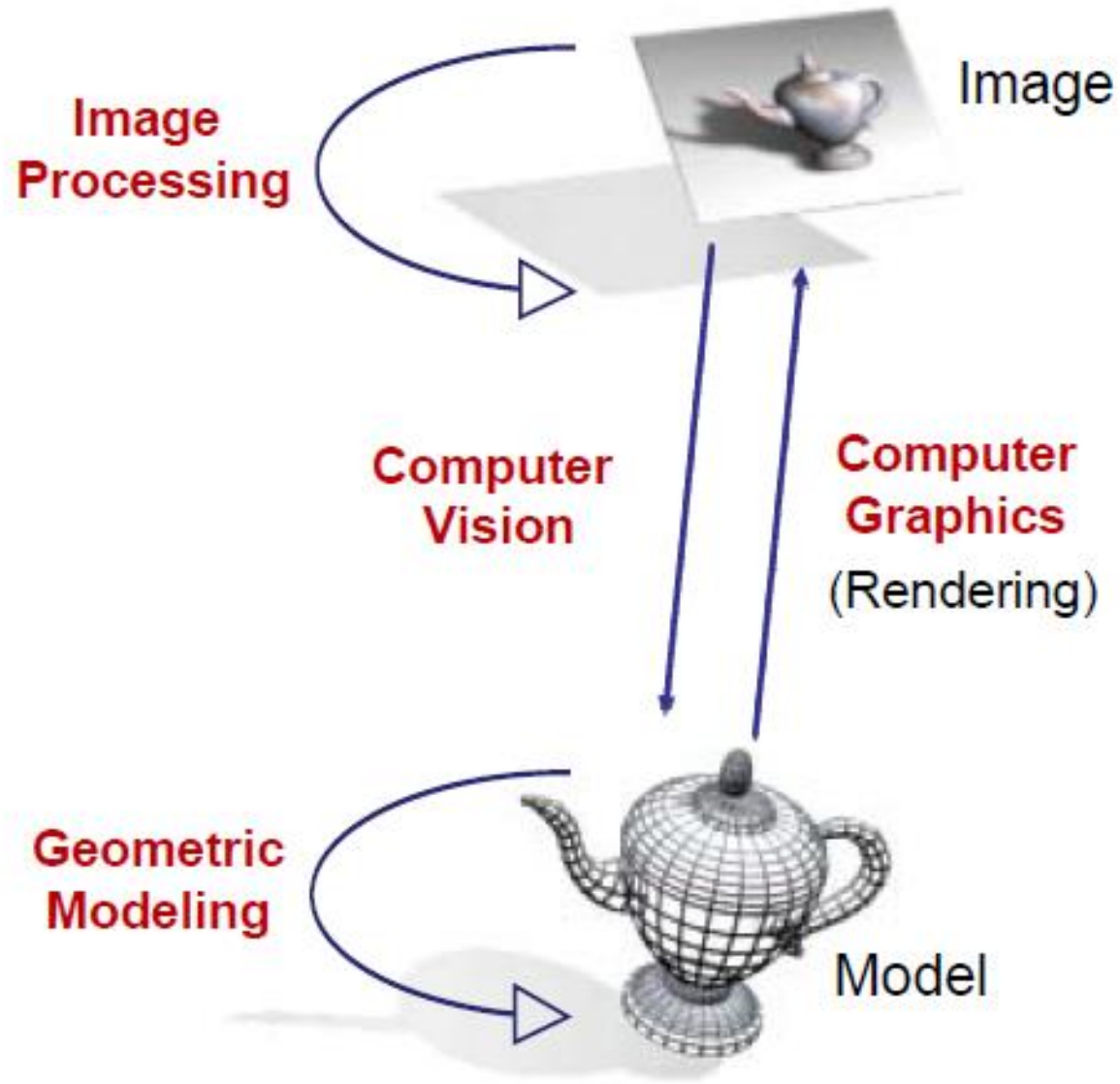
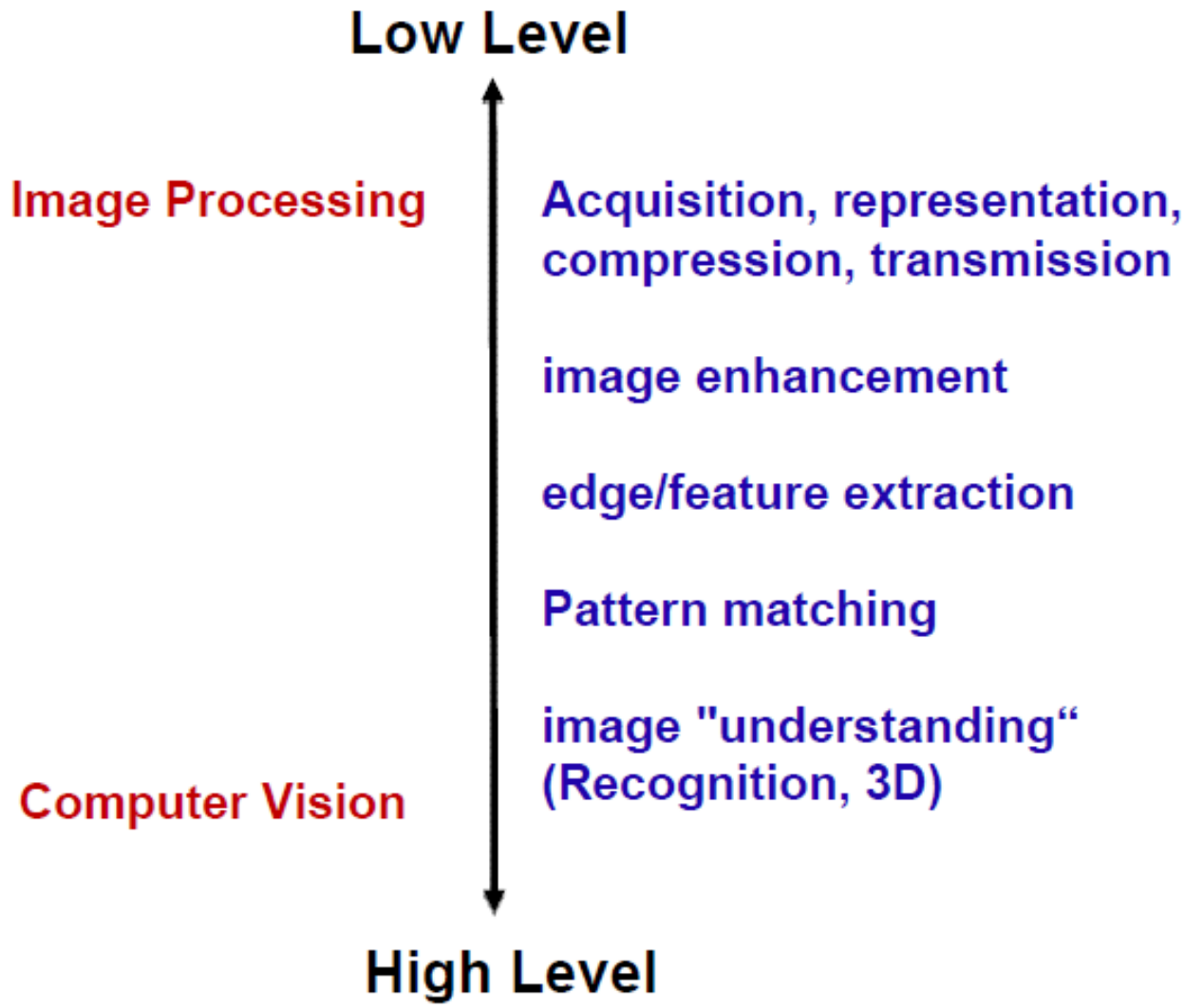


Image Processing - Computer Vision



Course Outlines

FIGURE 1.23
Fundamental steps in digital image processing.

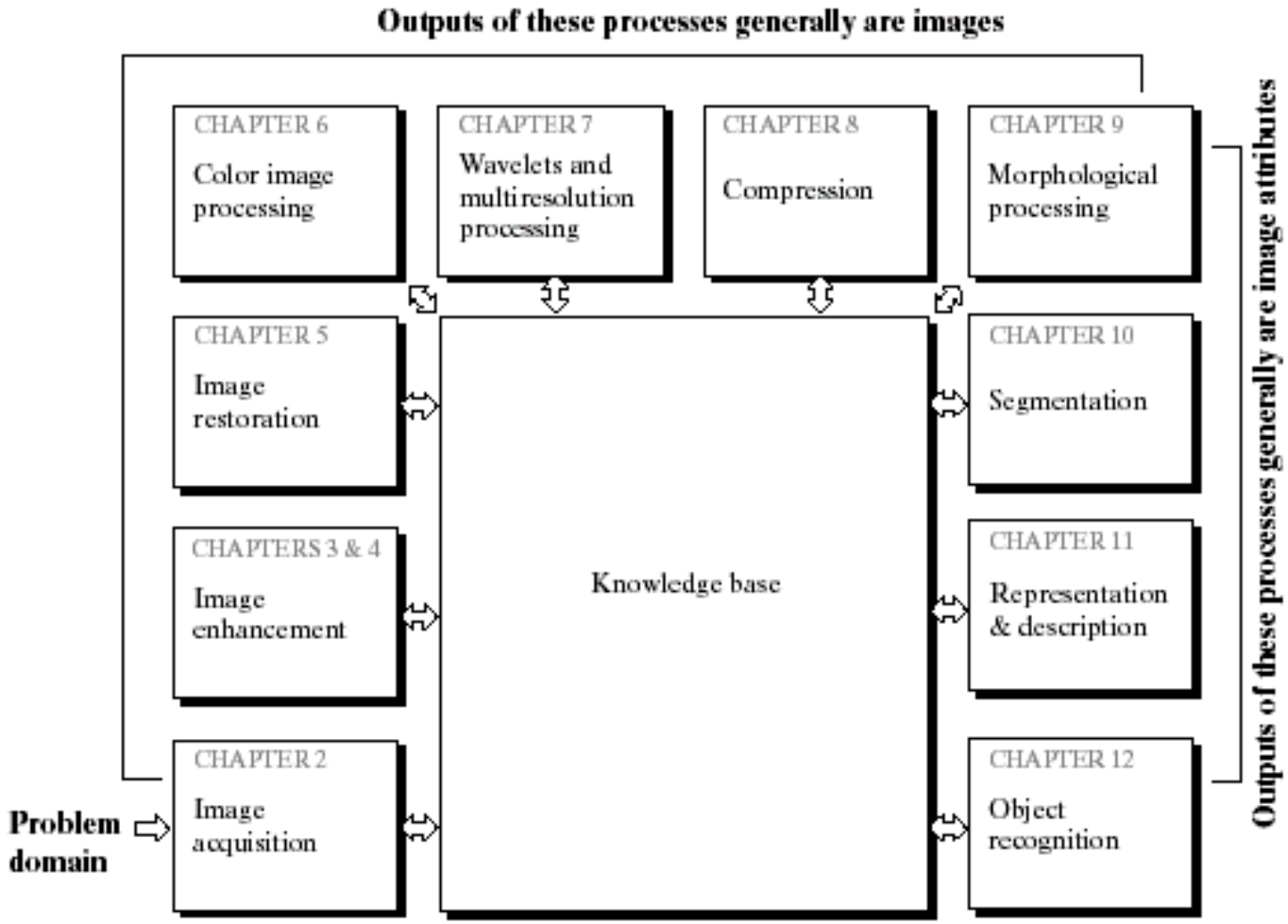
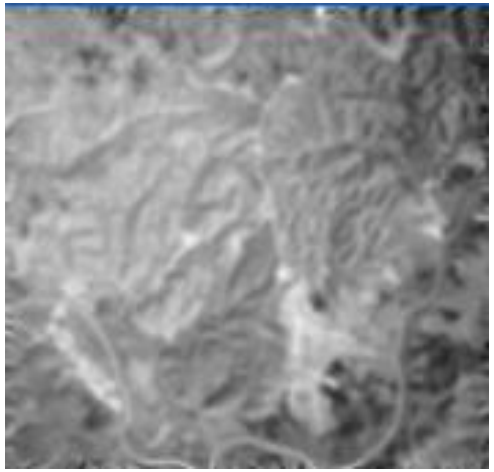


Image Enhancement



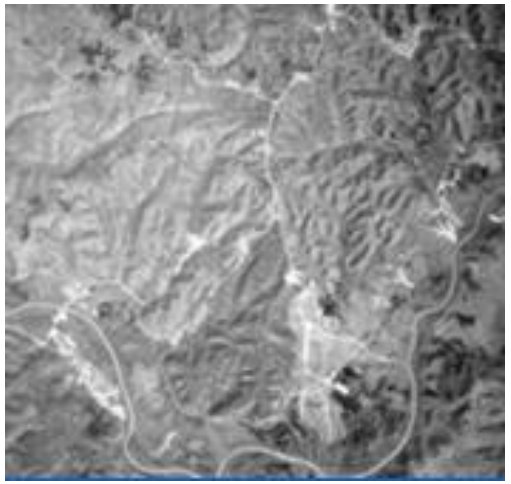
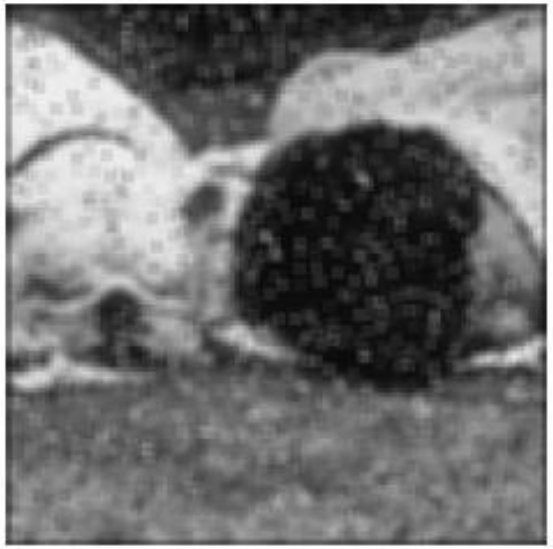


Image Denoising

Salt & Pepper Noise

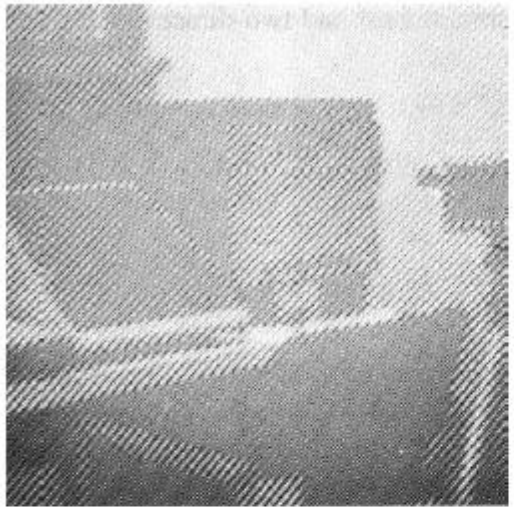
3 X 3 Average



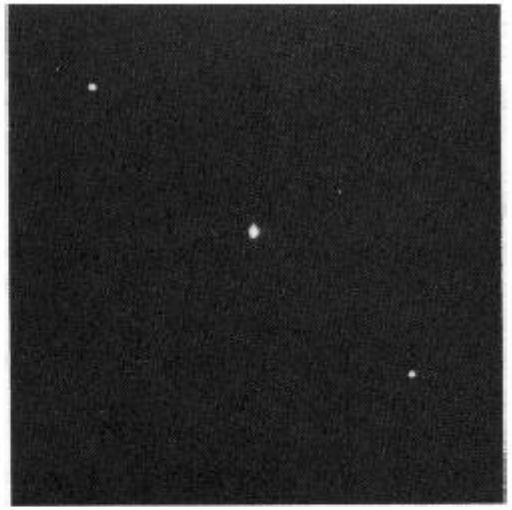
Median

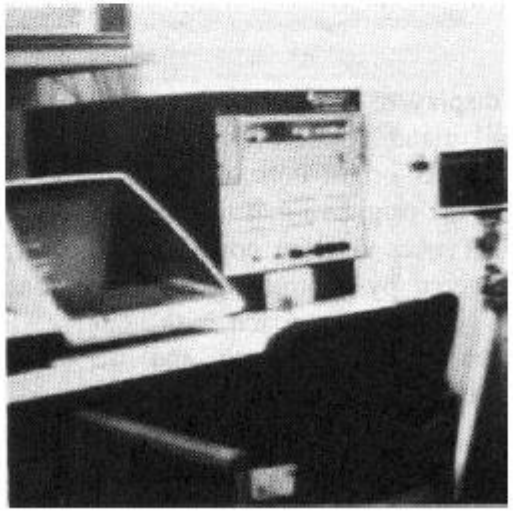
Image Enhancement - Frequency Domain

Original Noisy image



Fourier Spectrum





Edge Detection

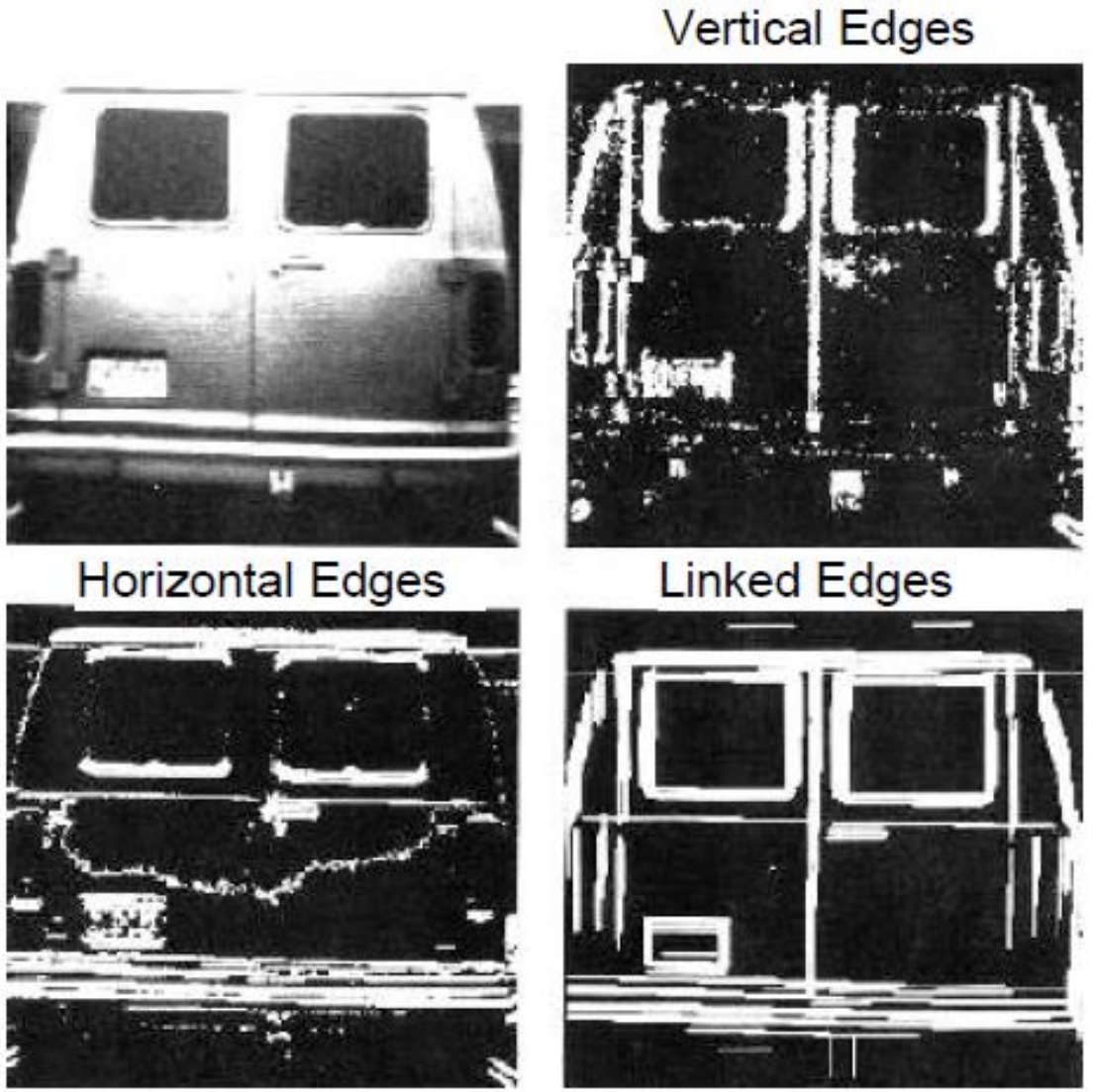
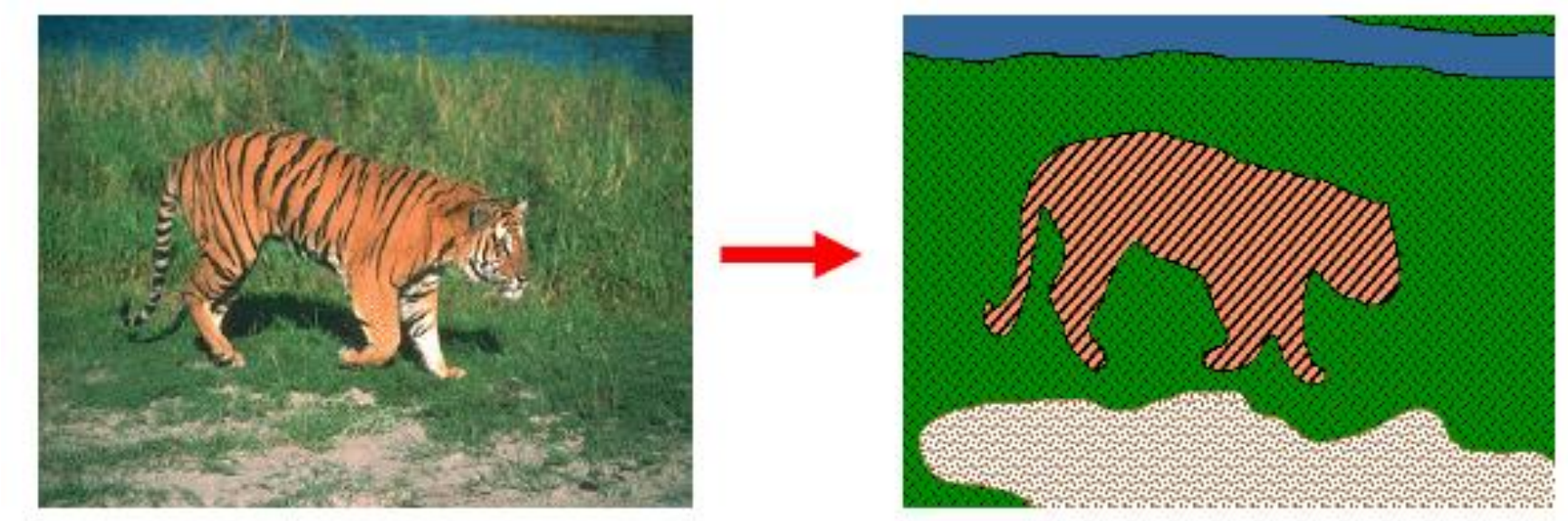


Image Segmentation

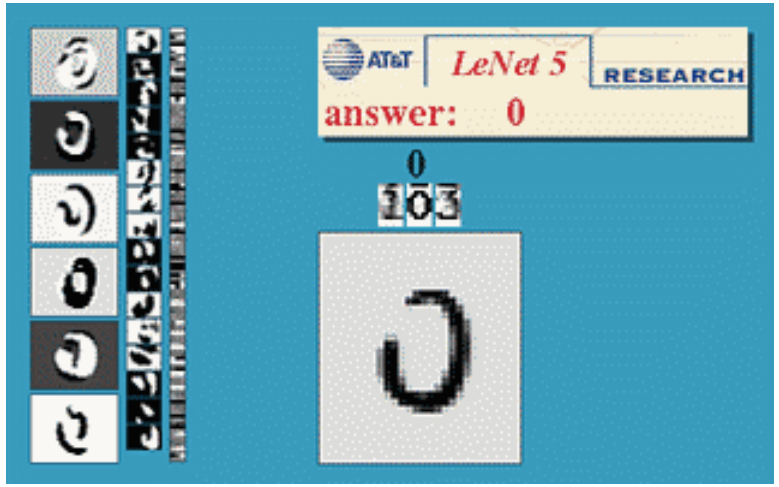
Goal: identify groups of pixels that go together



Optical character recognition (OCR)

Technology to convert scanned docs to text

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



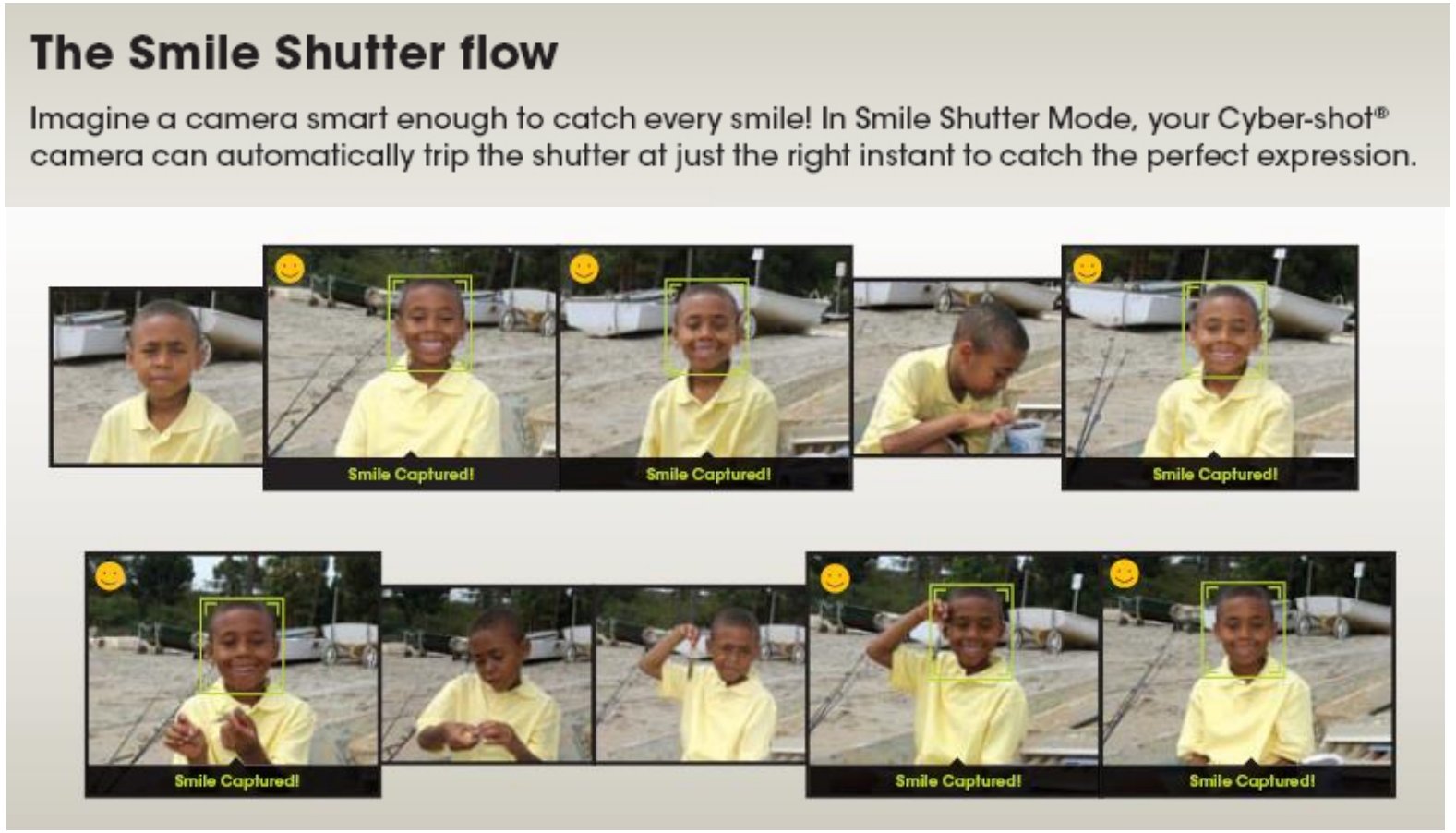
Digit recognition, AT&T labs
http://www.research.att.com/

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

Face detection

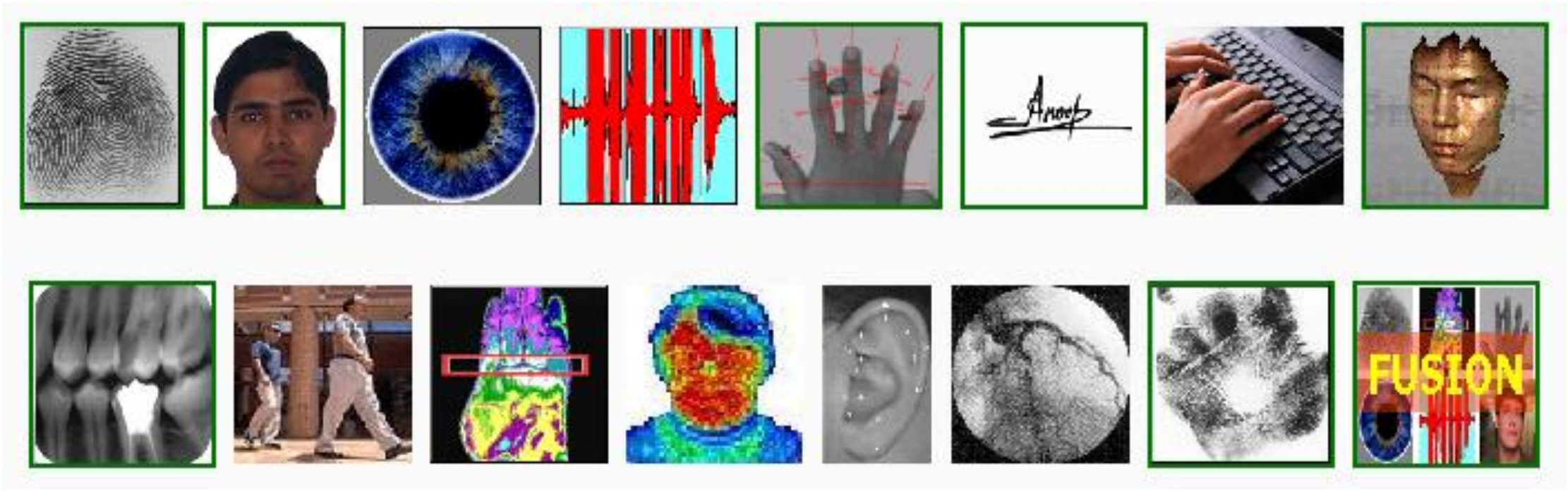
- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Smile detection



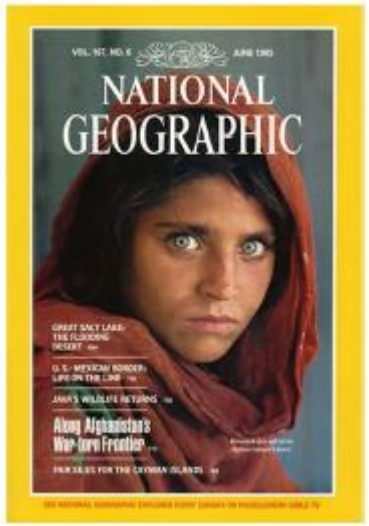
Sony Cyber-shot® T70 Digital Still Camera

Biometrics



Adapted from Anil Jain, Michigan State

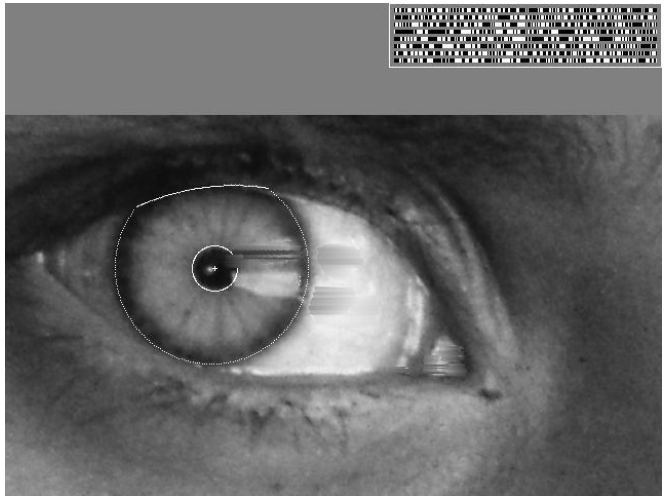
Vision-based biometrics

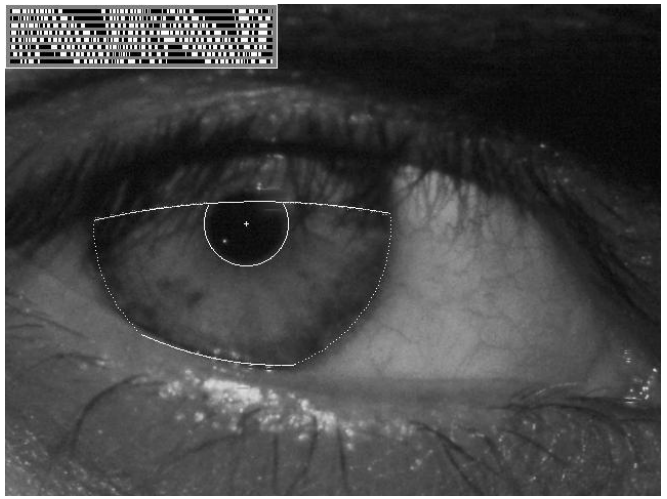


"*How the Afghan Girl was Identified by Her Iris Patterns*"

Read the story

wikipedia





Login without a password…

Fingerprint scanners on many new laptops, other devices

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

Object recognition (in mobile phones)

Point & Find, Google Goggles

Building a Panorama

M. Brown and D. G. Lowe. Recognising Panoramas. ICCV 2003

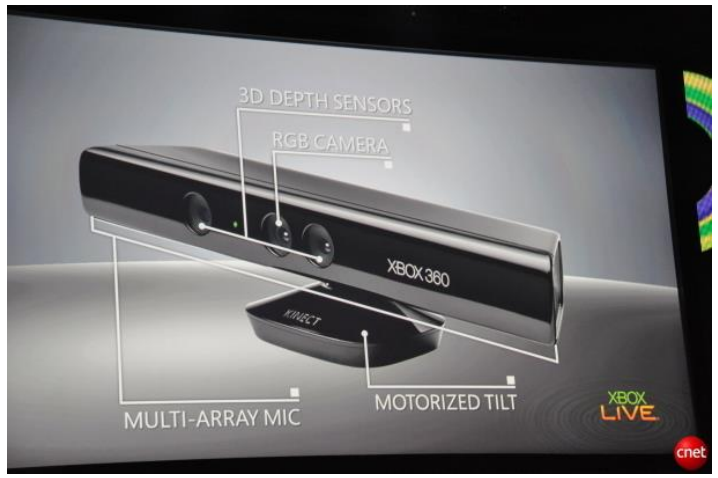
Feature descriptors

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint
    - Up to about 60 degree out of plane rotation
  - Can handle significant changes in illumination
    - Sometimes even day vs. night (below)
  - Fast and efficient—can run in real time

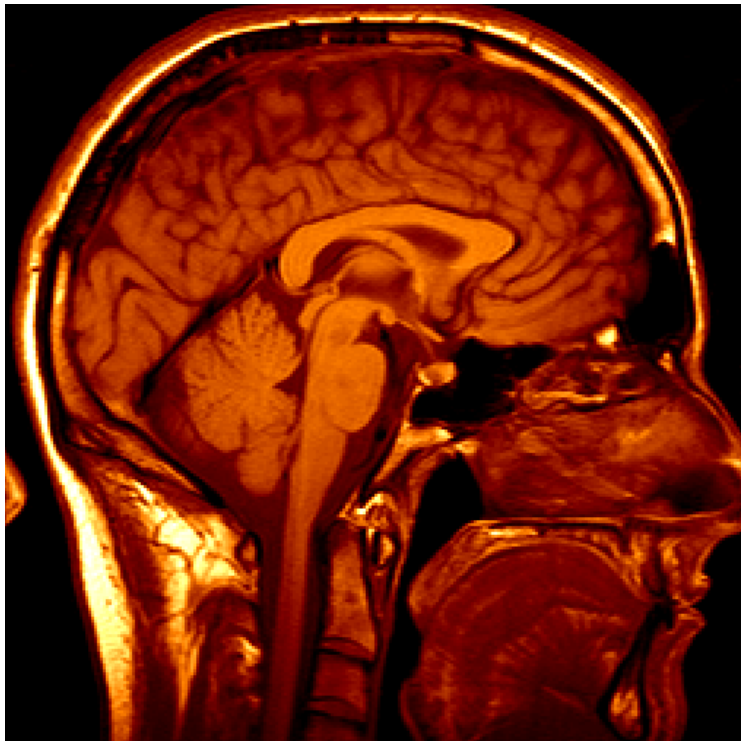
Steve Seitz

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



Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Smart cars

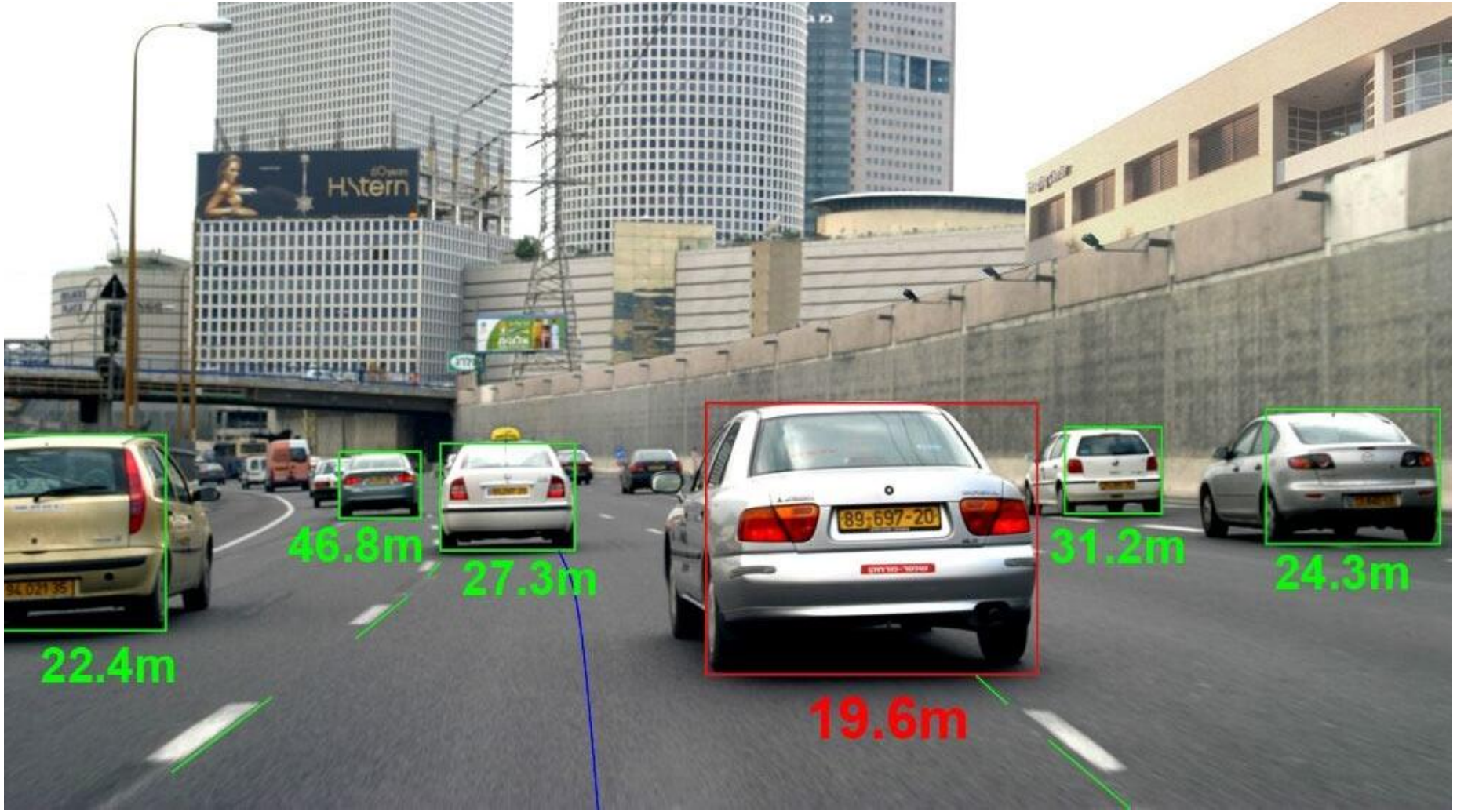
- https://www.mobileye.com/
- Vision systems currently in high-end BMW, GM, Volvo models

Google cars

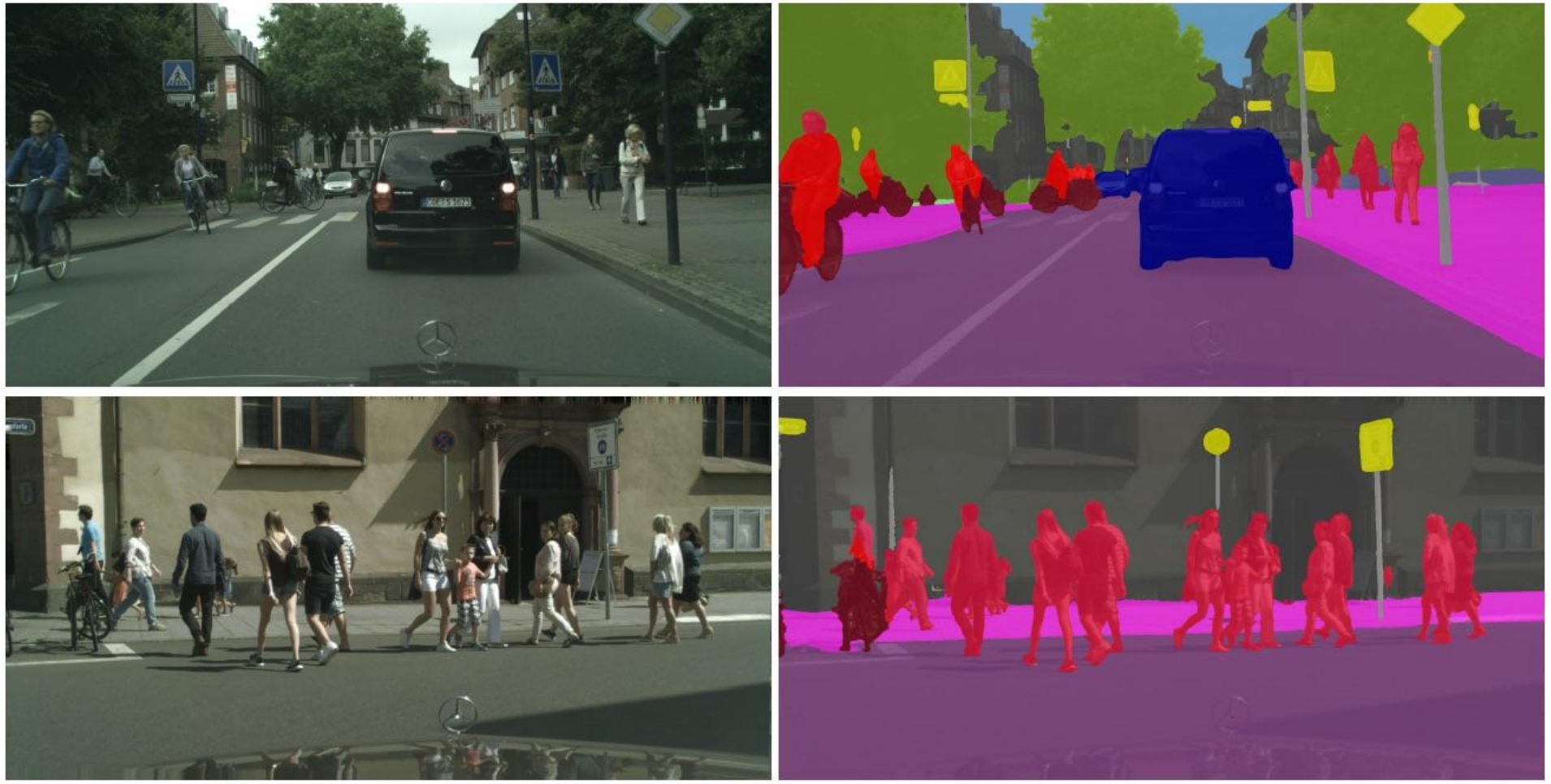
http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence

AutoCars - Uber bought CMU's lab

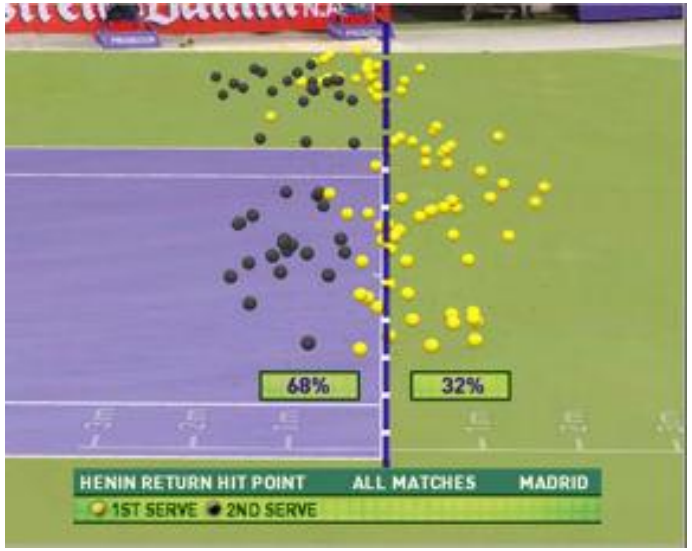
Car Detection and Depth Estimation

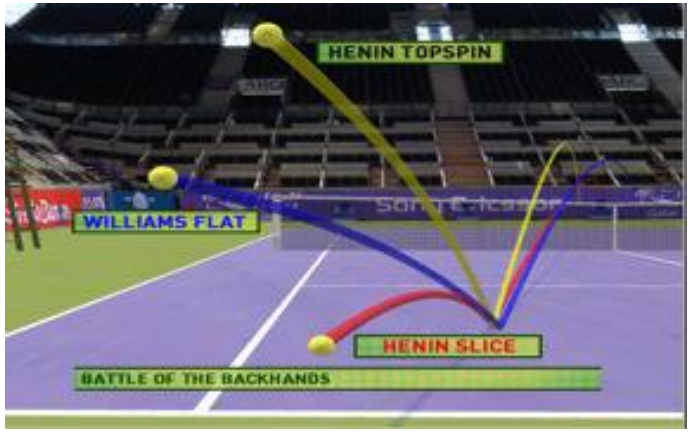


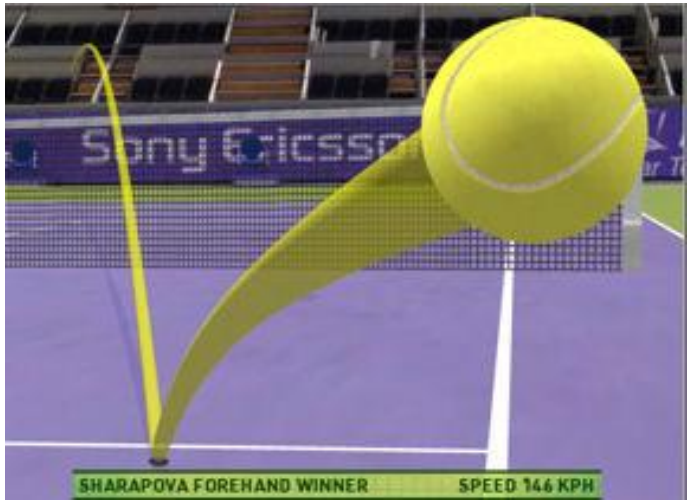
Vision as a Source of Semantic Information



Sports video analysis



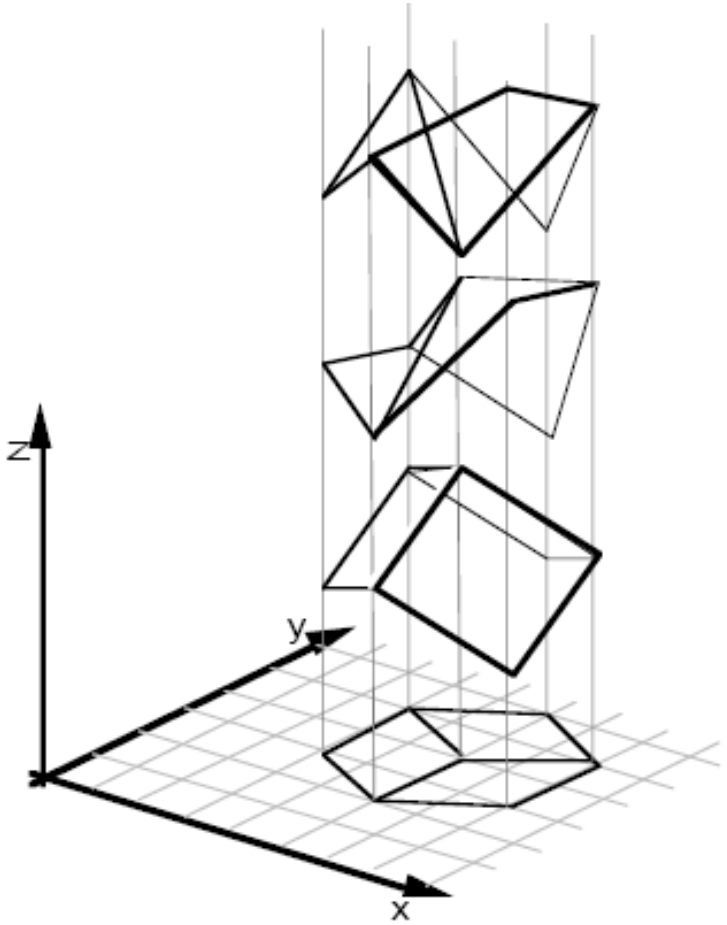




Tennis review system

http://www.hawkeyeinnovations.co.uk

Why is vision so hard?



[Sinha and Adelson 1993]

posed problem

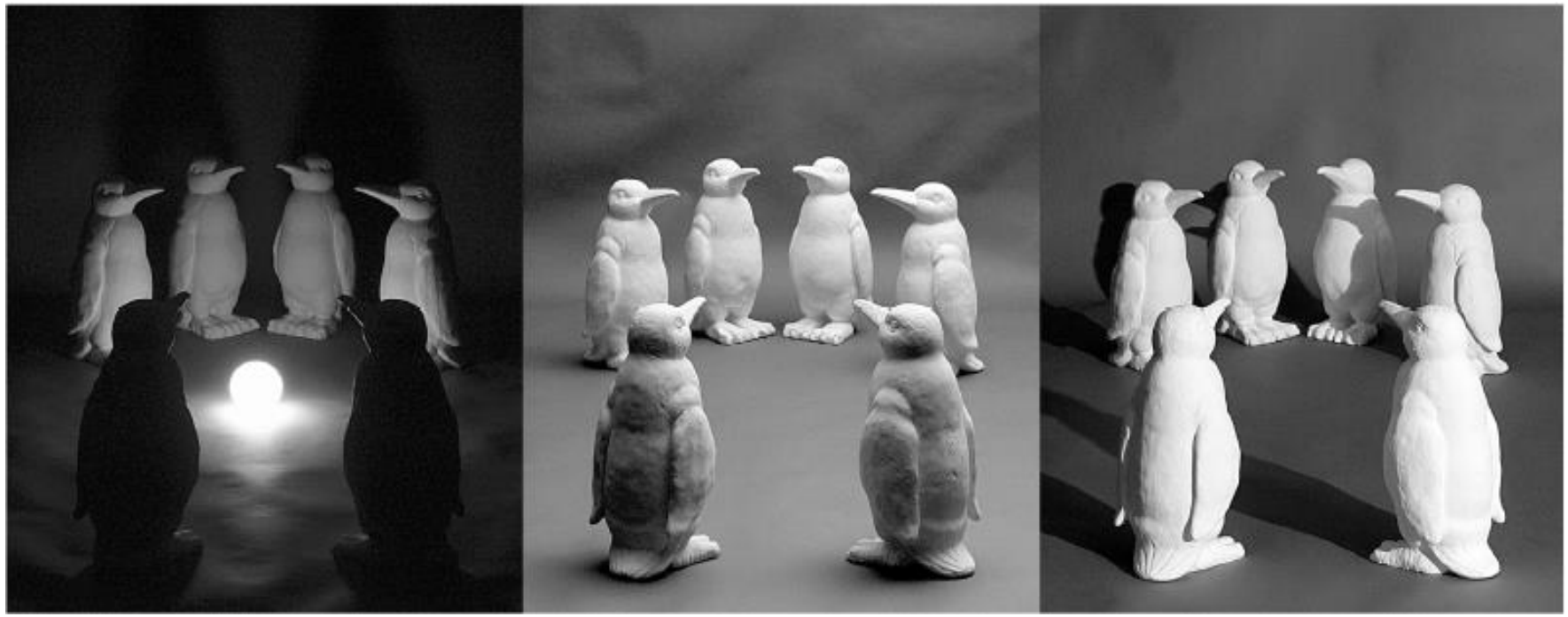
Adapted from Ali Farhadi, U of Washington

Challenges 1: view point variation

Michelangelo 1475-1564

Adapted from L. Fei-Fei, R. Fergus, A. Torralba

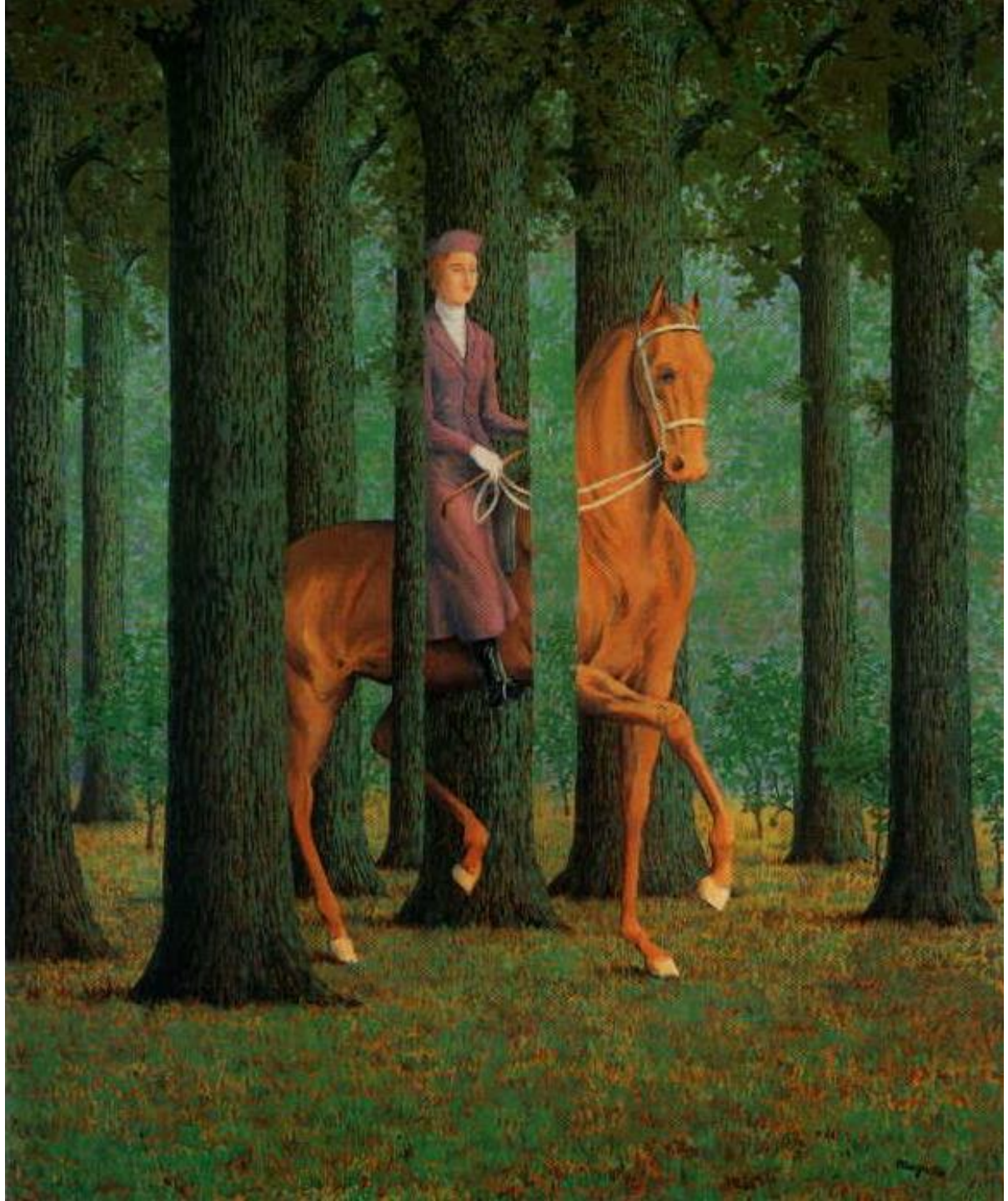
Challenges 2: illumination



Adapted from Fei-Fei Li

Challenges 3: occlusion

Magritte, 1957

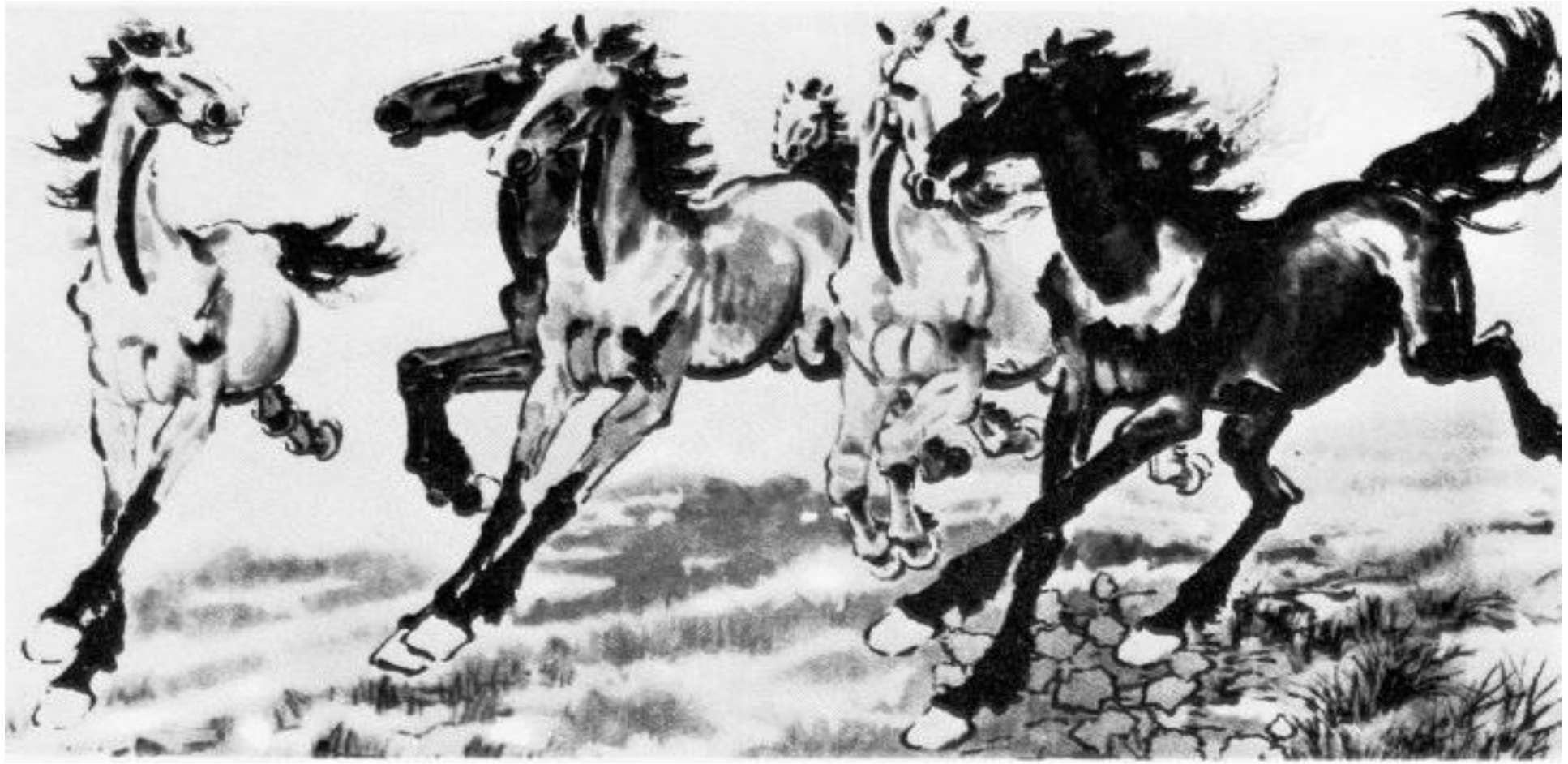


Adapted from L. Fei-Fei,
R. Fergus, A. Torralba

Challenges 4: scale

Adapted from L. Fei-Fei, R. Fergus, A. Torralba

Challenges 5: deformation



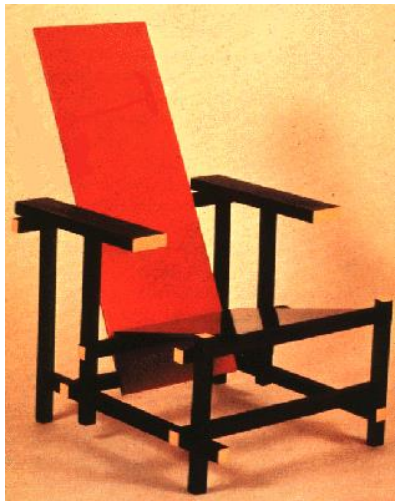
Xu, Beihong 1943

Adapted from L. Fei-Fei, R. Fergus, A. Torralba

Challenges 6: background clutter

Adapted from Fei-Fei Li

Challenges 7: intra-class variation



Adapted from L. Fei-Fei, R. Fergus, A. Torralba

What do computers see?

47	49	51	47	41	41	41	38	42	54	66	66	58	56	53	48	43	43	45	47	50	47	47	47
45	44	39	38	37	48	67	95	138	151	156	157	165	157	125	79	36	38	47	48	48	43	38	36
43	35	31	45	64	109	155	179	178	160	142	132	146	187	195	170	133	86	45	46	51	41	36	32
33	24	24	47	88	149	135	136	160	170	166	135	111	153	169	169	109	113	86	57	49	46	40	36
22	19	22	47	122	131	99	120	204	199	185	150	119	152	159	173	110	80	83	82	63	58	45	42
22	20	24	60	114	108	123	191	215	212	198	169	156	169	168	172	151	115	91	77	82	59	53	53
20	19	29	86	127	87	169	223	219	218	212	182	178	190	194	185	169	108	88	85	74	55	52	51
20	20	26	131	138	129	214	228	224	222	221	206	207	208	203	193	177	136	88	87	72	54	44	42
24	23	28	130	125	152	226	224	222	223	217	218	214	201	185	168	164	114	70	39	45	47	39	34
29	26	25	104	92	123	220	226	230	228	218	213	210	193	152	118	136	97	50	26	39	41	36	33
26	24	25	66	95	140	222	223	228	225	218	208	205	181	140	97	101	121	71	35	78	51	40	37
26	30	24	51	149	179	224	221	218	215	205	204	210	191	140	108	107	127	112	43	46	42	39	40
27	34	30	23	142	198	210	226	233	220	205	204	222	210	175	154	134	125	137	51	54	55	44	34
26	32	29	18	124	197	178	174	140	113	182	183	174	112	98	74	34	69	126	54	53	78	59	41
30	27	26	19	114	197	207	138	73	43	167	191	49	29	139	66	33	76	92	60	85	50	42	40
26	25	23	18	91	198	220	221	184	133	210	214	40	112	210	129	120	105	81	62	60	28	22	30
23	19	16	13	53	201	211	227	220	227	226	216	75	72	196	190	130	58	62	58	32	21	24	26
18	14	12	11	13	93	198	220	226	209	219	218	121	34	148	170	53	37	50	25	17	17	23	24
17	15	14	13	15	25	177	203	189	151	223	219	139	59	33	78	30	39	45	26	22	21	16	38
12	14	17	13	15	11	125	201	149	194	223	203	67	19	15	22	33	43	55	37	29	28	31	68
10	13	14	11	16	15	58	196	170	193	213	175	123	34	19	48	37	93	35	32	30	38	93	118
17	19	19	20	31	35	30	145	191	201	215	182	134	47	66	89	45	196	45	16	52	98	141	149
25	28	34	34	28	32	20	105	216	215	213	187	168	130	73	26	148	195	34	12	21	76	121	123
31	36	30	26	29	42	20	77	220	215	221	213	185	131	37	117	201	85	56	11	16	10	22	38
24	20	21	40	43	42	24	106	190	235	212	188	134	85	138	178	45	89	40	13	19	13	19	21
14	21	41	43	42	32	19	131	207	250	239	197	206	236	220	33	18	94	13	16	18	11	12	17
32	36	46	39	40	27	10	157	250	230	190	156	172	216	250	135	149	50	9	18	16	12	13	18
38	38	38	45	40	29	10	140	240	244	151	50	30	118	229	255	187	11	16	26	20	18	25	29
40	34	33	31	36	27	16	117	237	253	169	60	101	217	245	255	93	3	15	25	21	21	24	27
43	34	34	32	31	21	18	80	232	252	147	85	208	247	252	207	18	13	10	13	19	20	20	21
41	33	33	32	31	18	27	64	220	211	62	71	209	246	250	108	5	19	11	13	16	18	21	20
40	33	33	34	30	17	31	50	182	159	49	45	136	248	208	24	11	13	12	17	11	10	15	19

Stages of computer vision

- Low-level

  image → image

- Mid-level

  image → features / attributes

- High-level

  features → “making sense”, recognition

Low-level

sharpening

blurring

Adapted from Linda Shapiro, U of Washington

Low-level

original image

Canny

edge image

Mid-level

edge image

ORT

data structure

circular arcs and line segments

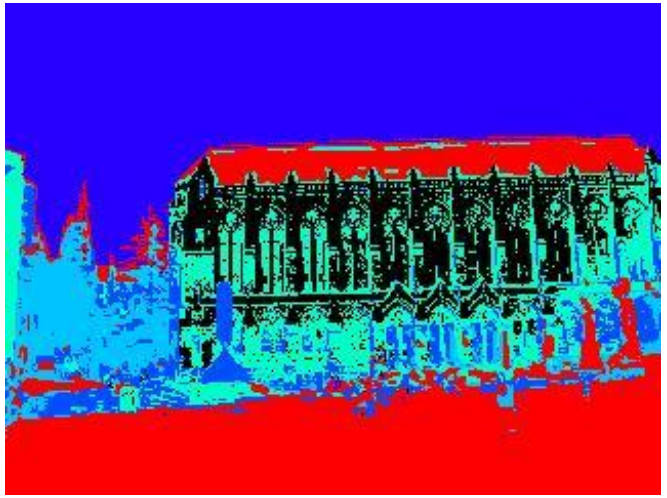
Adapted from Linda Shapiro, U of Washington

Mid-level

original color image

K-means clustering

(followed by connected component analysis)



regions of homogeneous color

data structure

Adapted from Linda Shapiro, U of Washington

Low-level to high-level

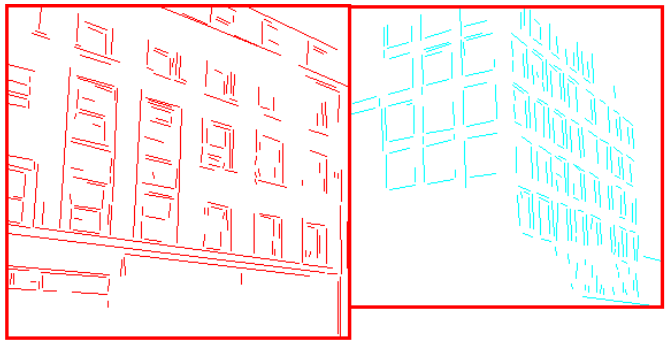
low-level

edge image

mid-level

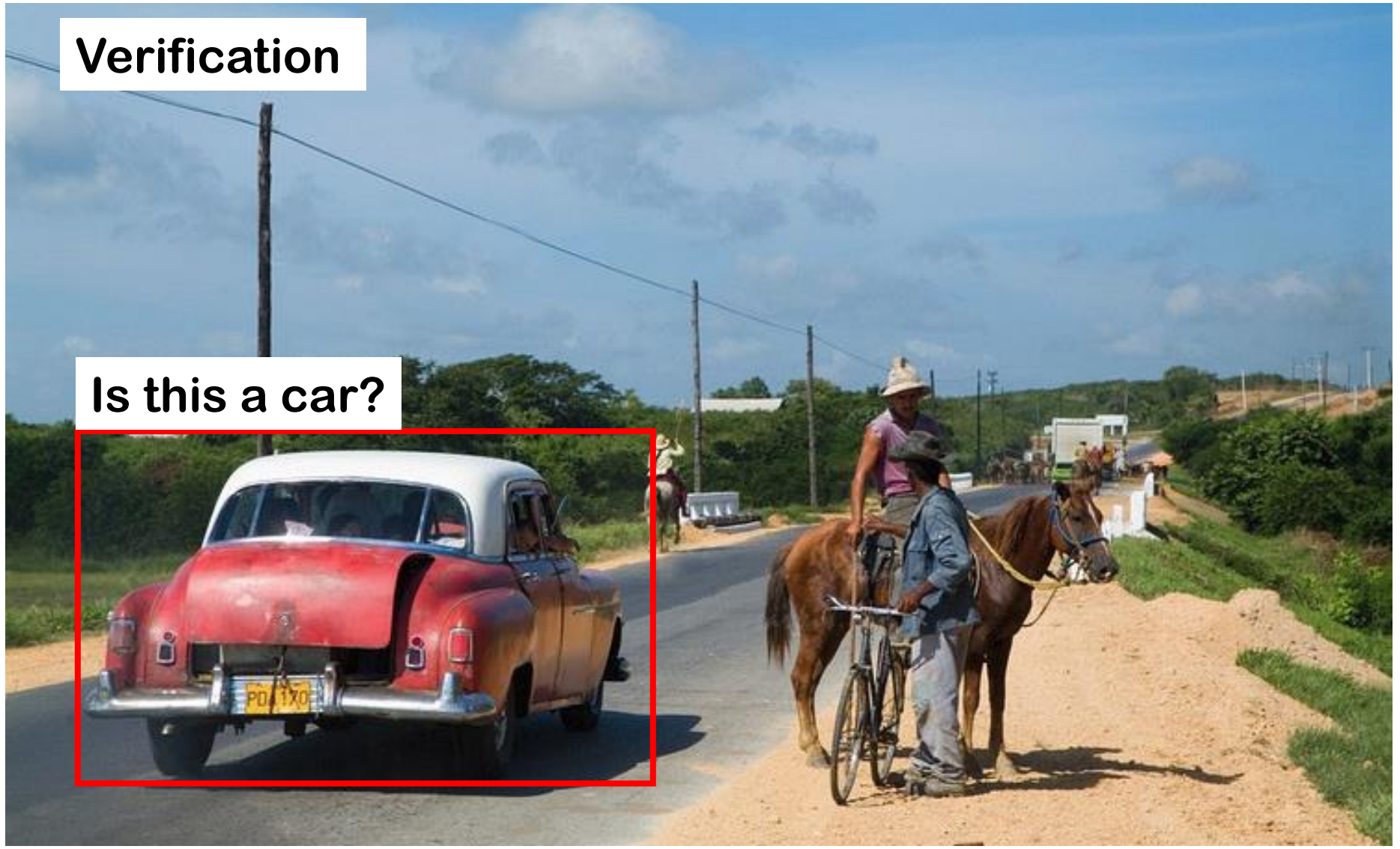
consistent line clusters

high-level



Adapted from Linda Shapiro, U of Washington

Visual recognition

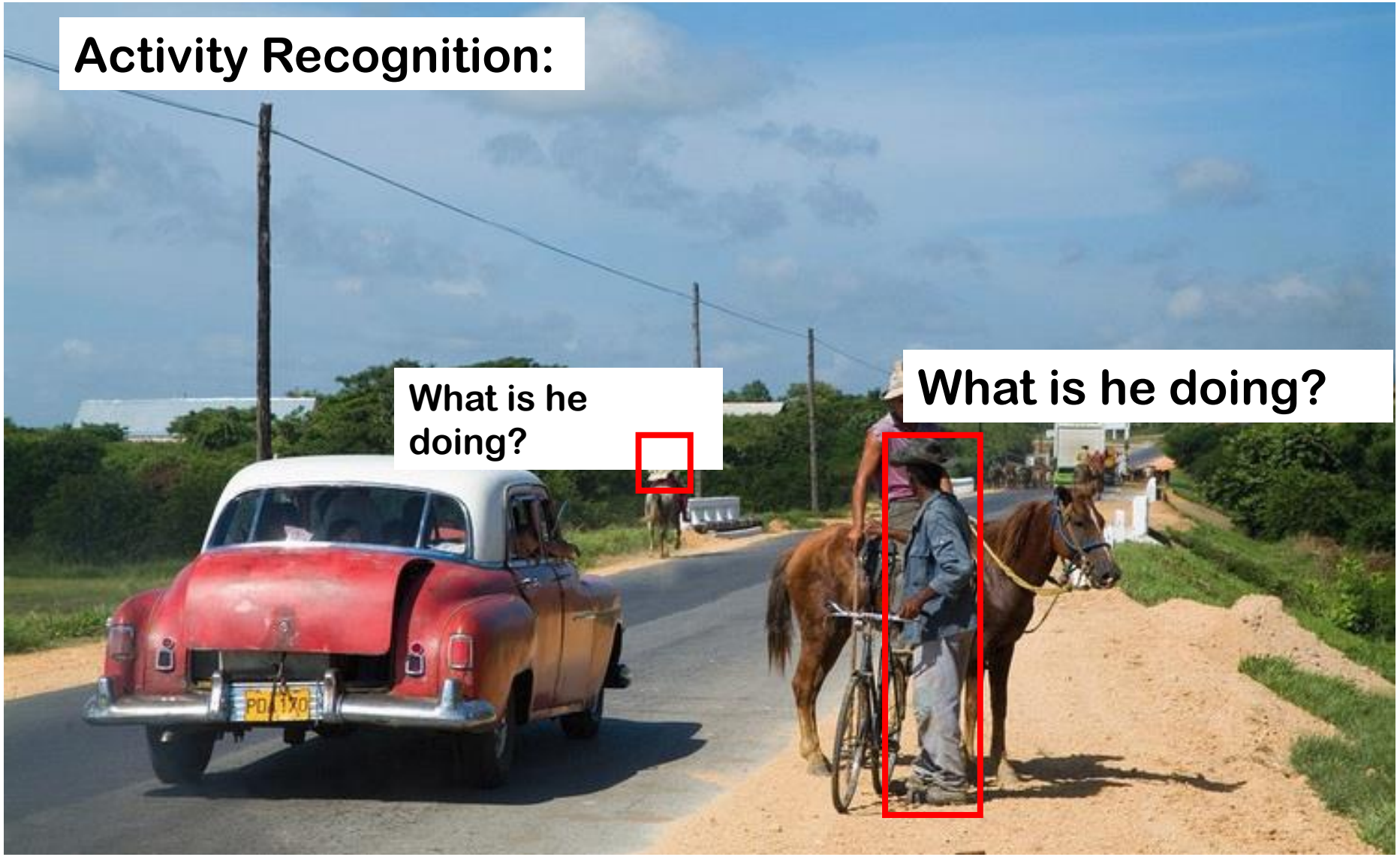


Visual recognition

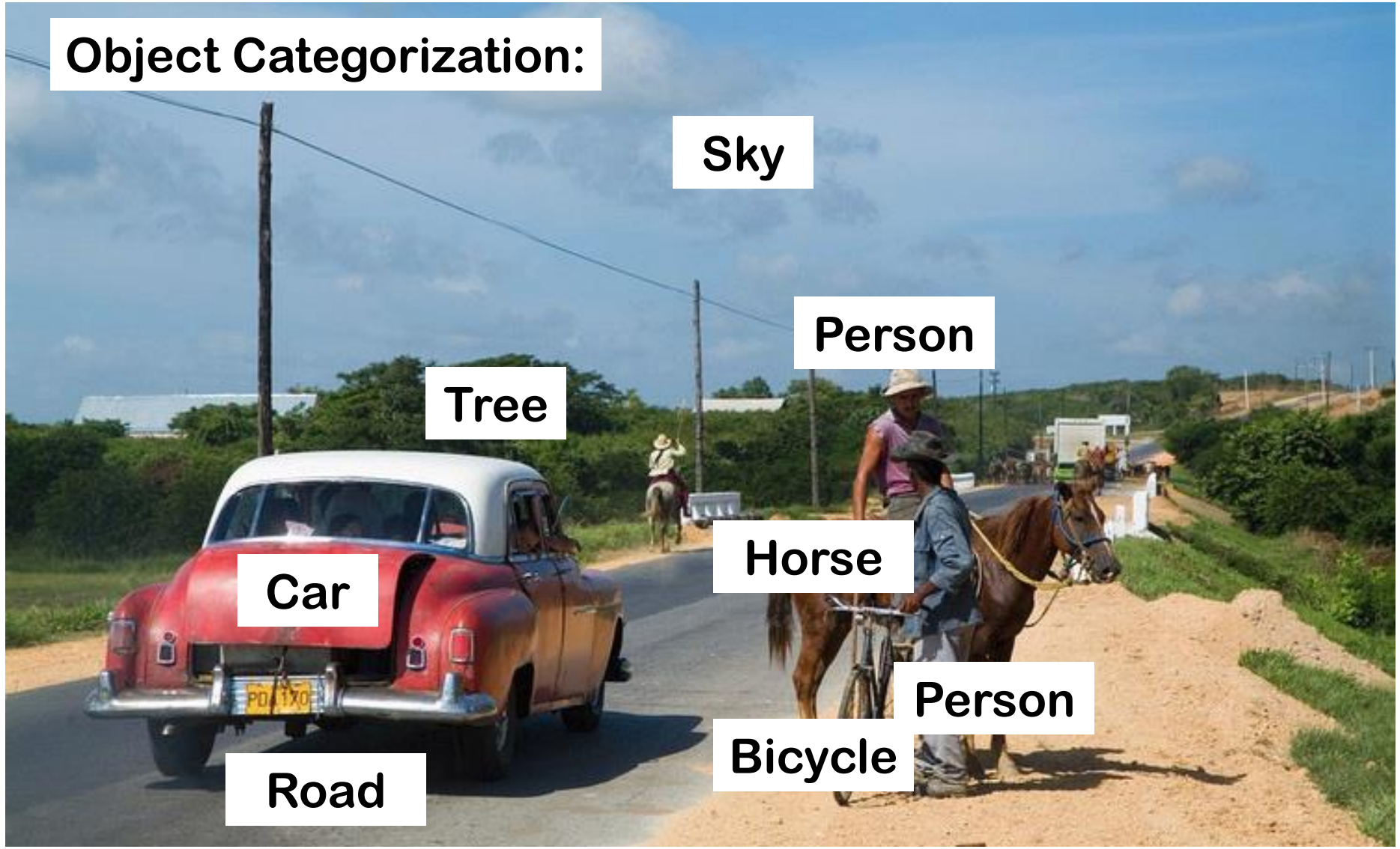
Visual recognition

Visual recognition

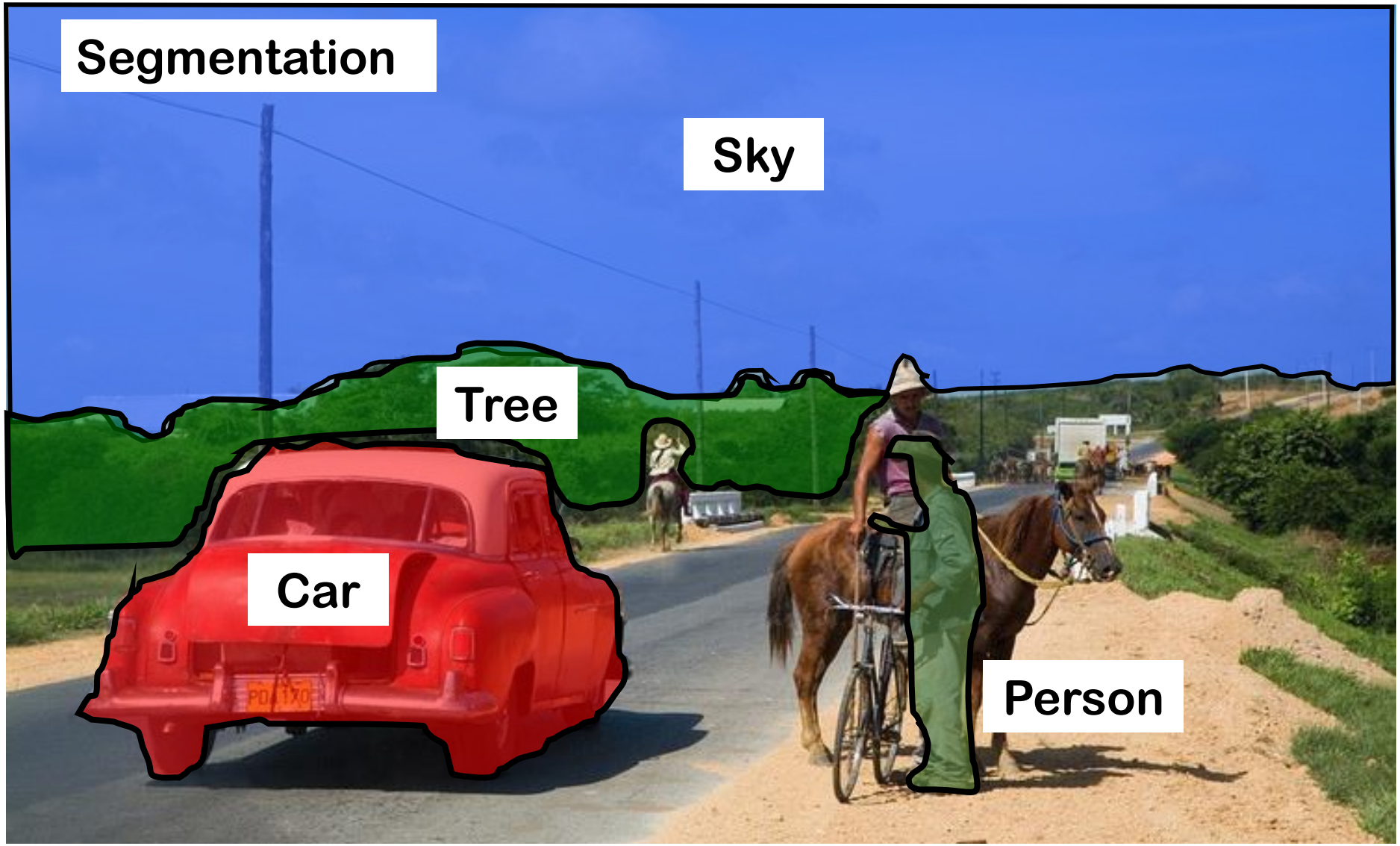
Visual recognition



Visual recognition



Visual recognition



Computer Vision Projects

- Stanford projects

  https://web.stanford.edu/class/ee368/index.html

- Ain Shams Computer Vision Competition projects

http://alyosama.github.io/computer/vision/2018/05/03/CVC18.html

http://alyosama.github.io/computer/vision/2019/05/08/CVC19.html

http://ihub.asu.edu.eg/cvc2020-results.html

OpenCV - Enabling computer vision

- Open Source Computer Vision library
- Cross-platform
- Free for use under open source BSD license
- Can be easily used with Java, Python, C and C++
- Supports Machine Learning libraries such as TensorFlow.
- https://opencv.org