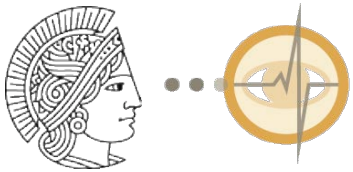
Part Based Object and People Detection

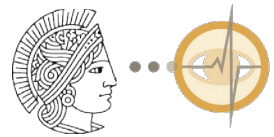
Cognitive Science Summerschool, Aug 27, 2oo9
Part 1: Introduction & Overview

Bernt Schiele

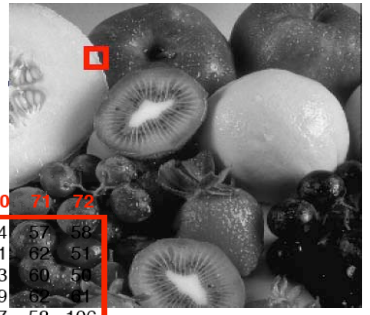
TU Darmstadt, Germany
http://www.mis.informatik.tu-darmstadt.de/
schiele@informatik.tu-darmstadt.de



(Grayscale) Image

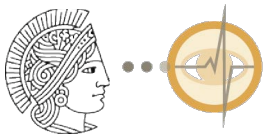


- ‘Goals’ of Computer Vision
  - how can we recognize fruits from an array of (gray-scale) numbers?
  - how can we perceive depth from an array of (gray-scale) numbers?
  - …
- ‘Goals’ of Graphics
  - how can we generate an array of (gray-scale) numbers that looks like fruits?
  - how can we generate an array of (gray-scale) numbers so that the human observer perceives depth?
  - …
- computer vision = the problem of ‘inverse graphics’ …?



	x = 58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
y = 41	210	209	204	202	197	247	143	71	64	80	84	54	54	57	58
42	206	196	203	197	195	210	207	56	63	58	53	53	61	62	51
43	201	207	192	201	198	213	156	69	65	57	55	52	53	60	50
44	216	206	211	193	202	207	208	57	69	60	55	77	49	52	51
45	221	206	211	194	196	197	220	56	63	60	55	46	97	58	106
46	209	214	224	199	194	193	204	173	64	60	59	51	62	56	48
47	204	212	213	208	191	190	191	214	60	62	66	76	51	49	55
48	214	215	215	207	208	180	172	188	69	72	55	49	56	52	56
49	209	205	214	205	204	196	187	196	86	62	66	87	57	60	48
50	208	209	205	203	202	186	174	185	149	71	63	55	55	45	56
51	207	210	211	199	217	194	183	177	209	90	62	64	52	93	52
52	208	205	209	209	197	194	183	187	187	239	58	68	61	51	56
53	204	206	203	209	195	203	188	185	183	221	75	61	58	60	60
54	200	203	199	236	188	197	183	190	183	196	122	63	58	64	66
55	205	210	202	203	199	197	196	181	173	186	105	62	57	64	63

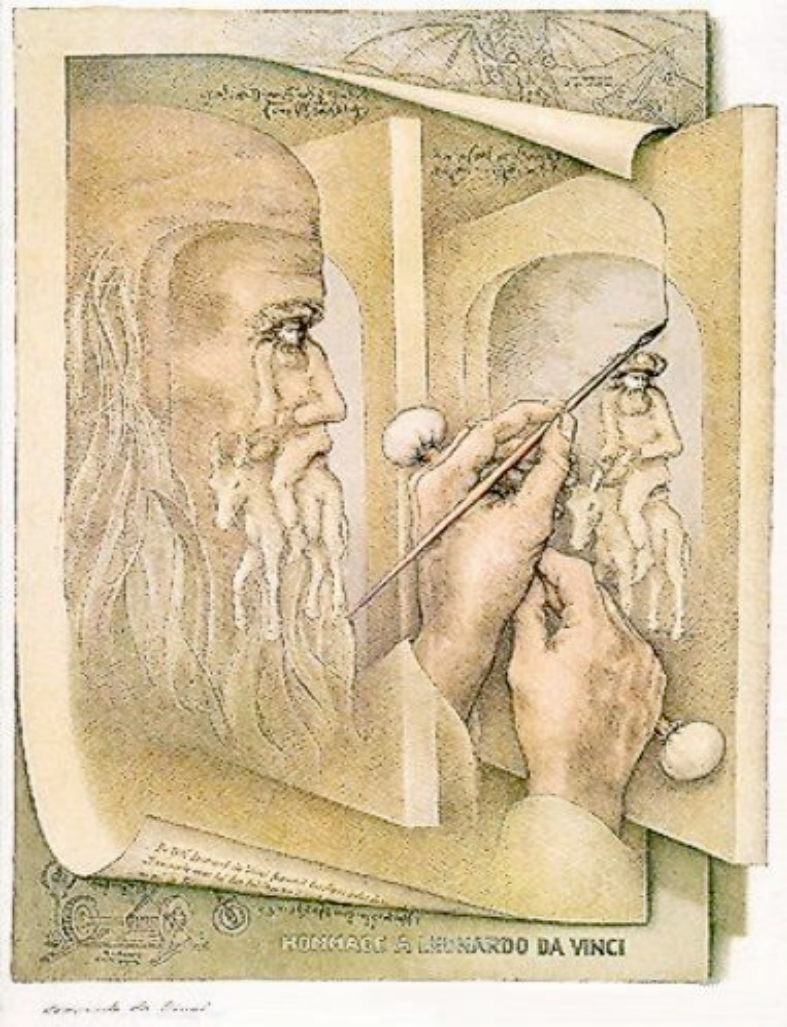
Computer Vision & Object Recognition



- is it more than inverse graphics?
- how do you recognize
  - the banana?
  - the glas?
  - the towel?
- how can we make computers to do this?



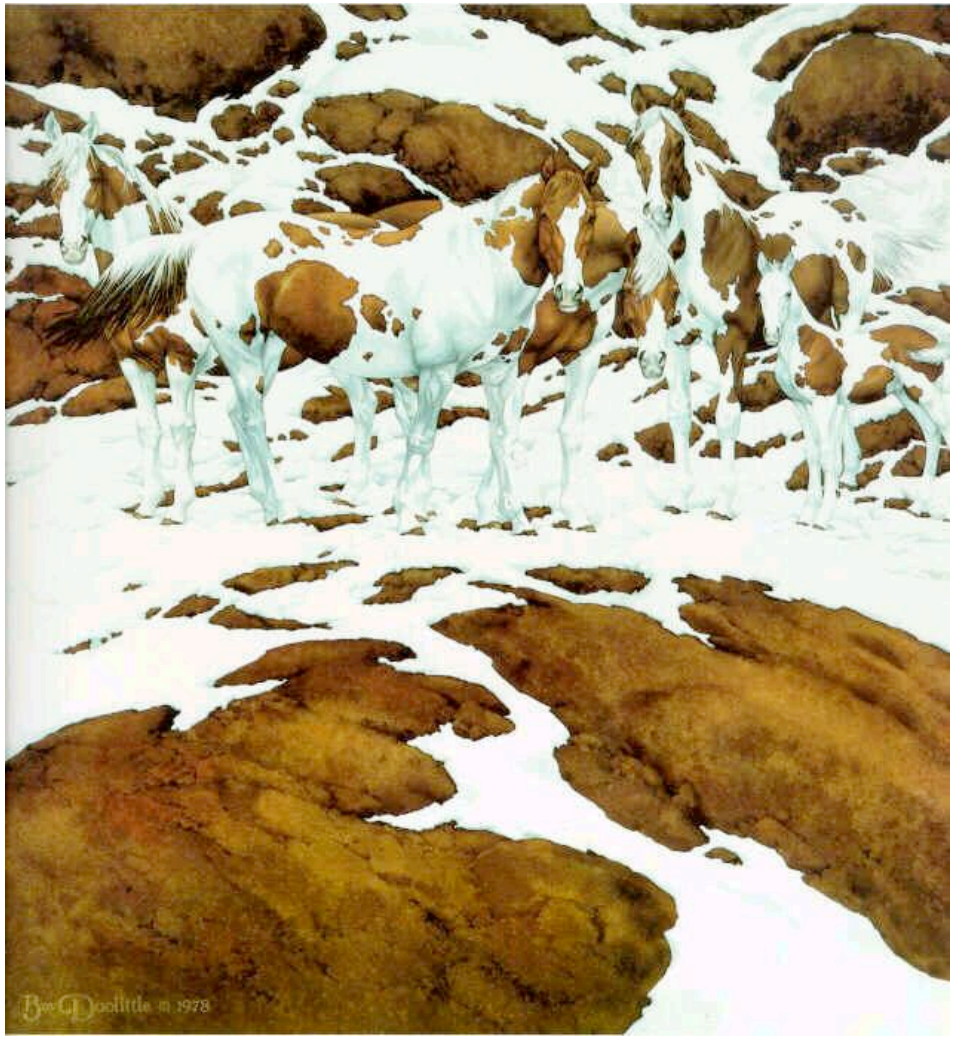
Complexity of Recognition



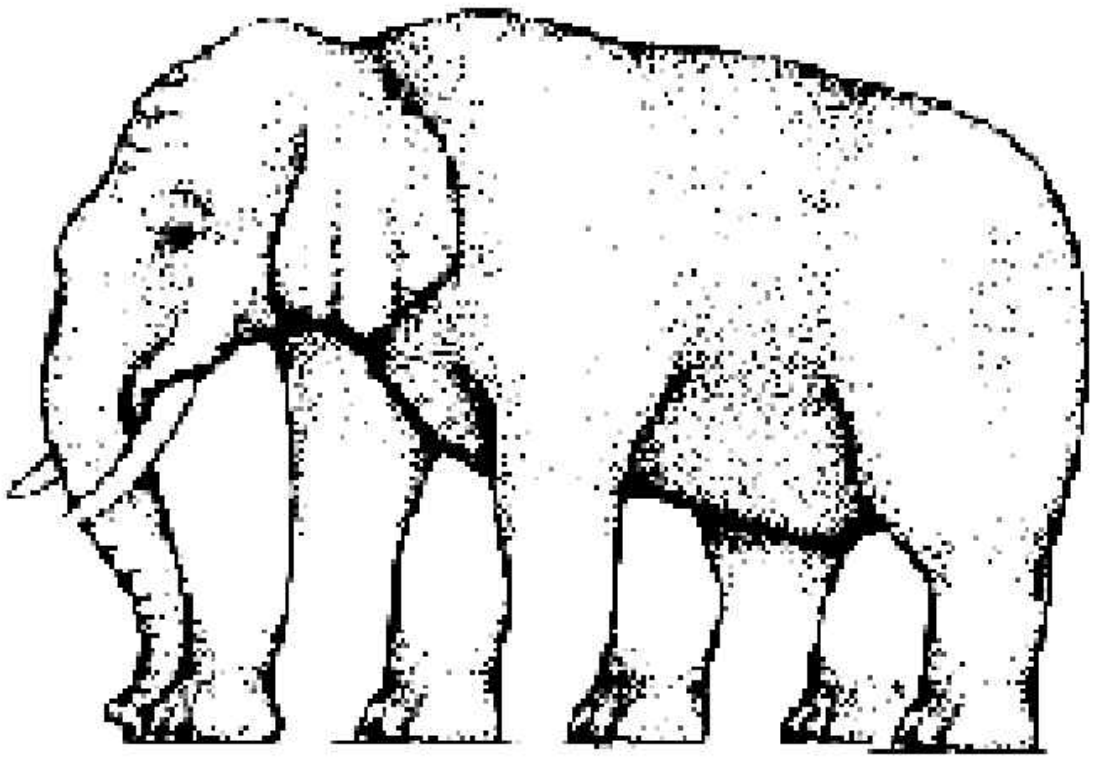
Complexity of Recognition



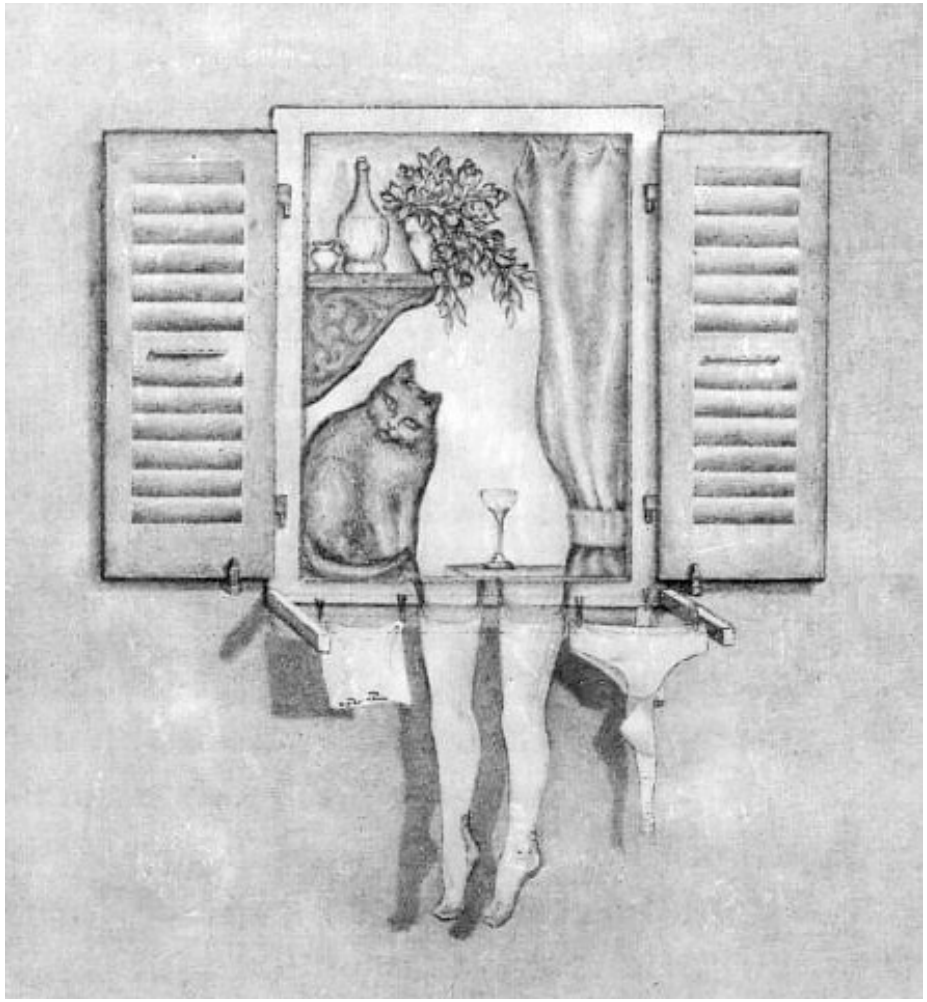
Complexity of Recognition



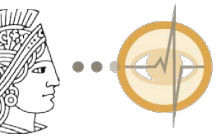
Complexity of Recognition



Complexity of Recognition



Recognition: the Role of Context



- Antonio Torralba

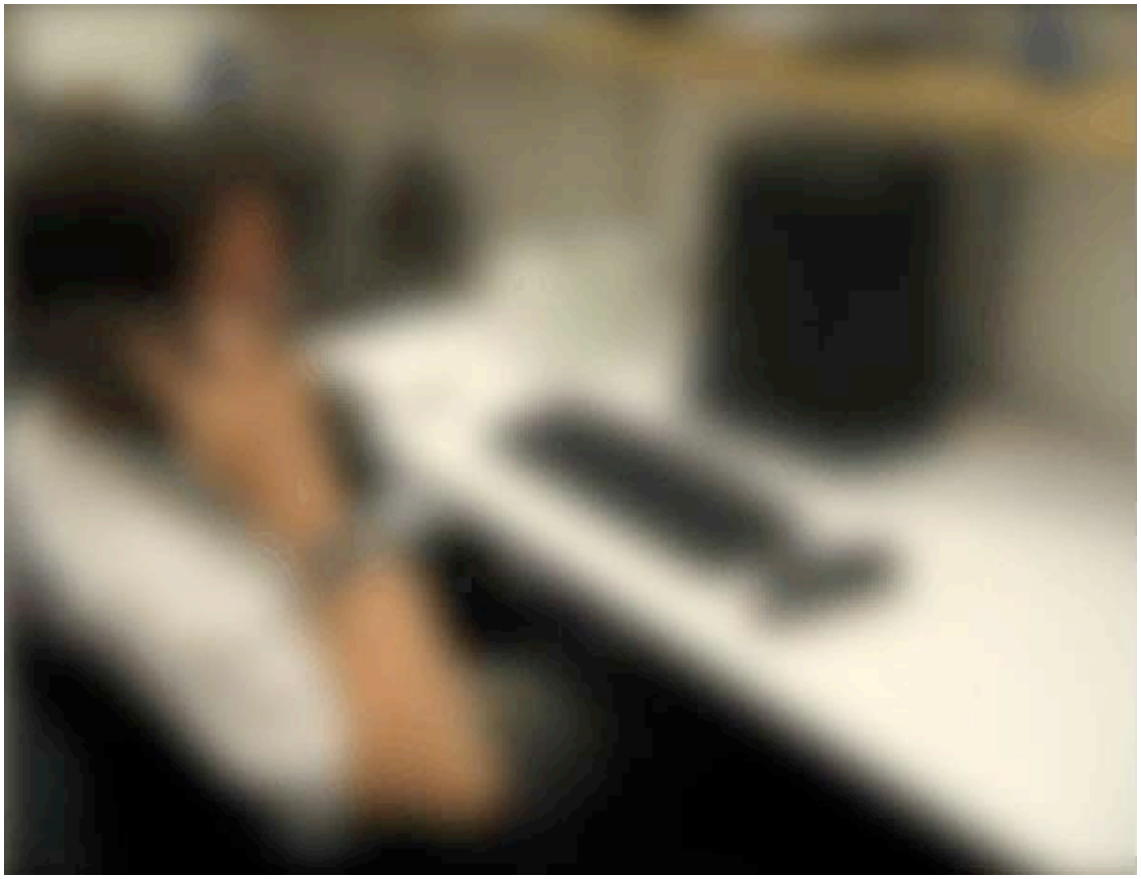




Recognition: the Role of Context



- Antonio Torralba (MIT) & Rob Fergus (NYU)



Recognition: the Role of Context



- Antonio Torralba (MIT) & Rob Fergus (NYU)



Complexity of Recognition



Complexity of Recognition



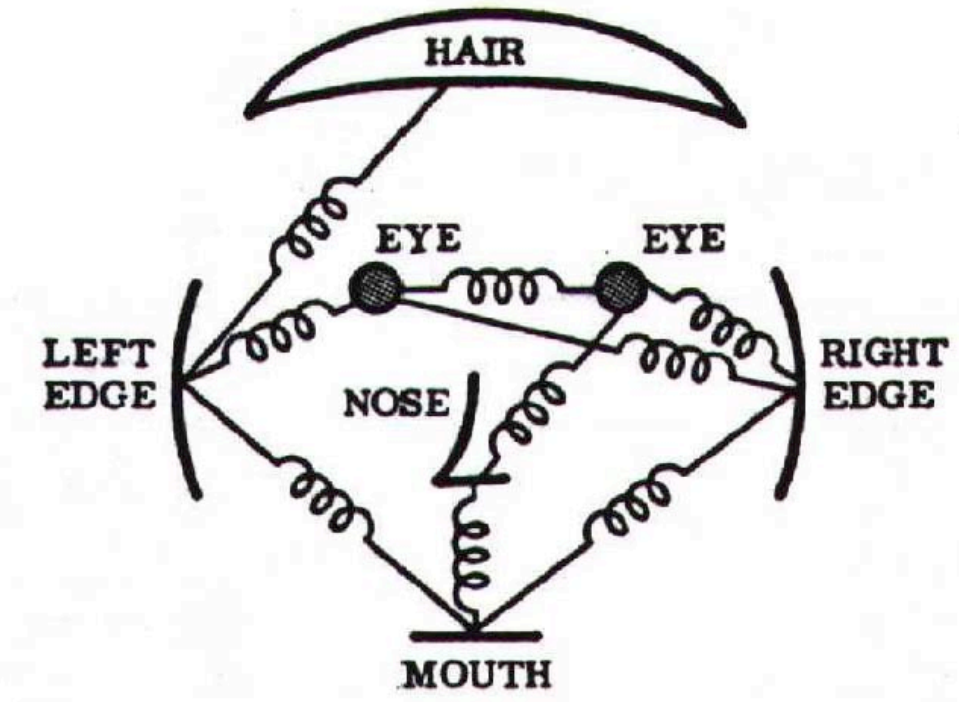
One or Two Faces ?



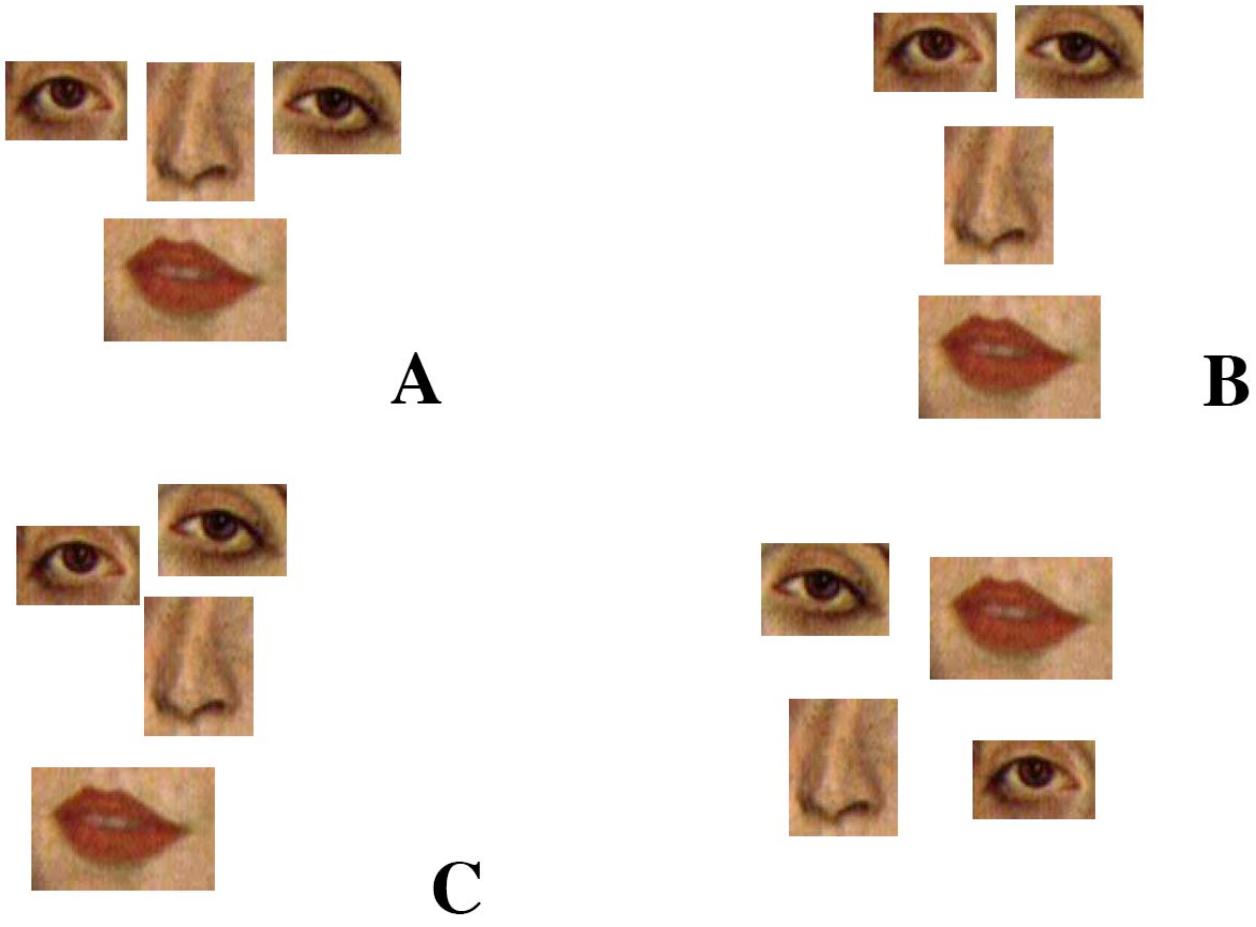
Class of Models: Pictorial Structure

[pictures from Perona]

- Fischler & Elschlager 1973
- Model has two components
  - parts
    (2D image fragments)
  - structure
    (configuration of parts)

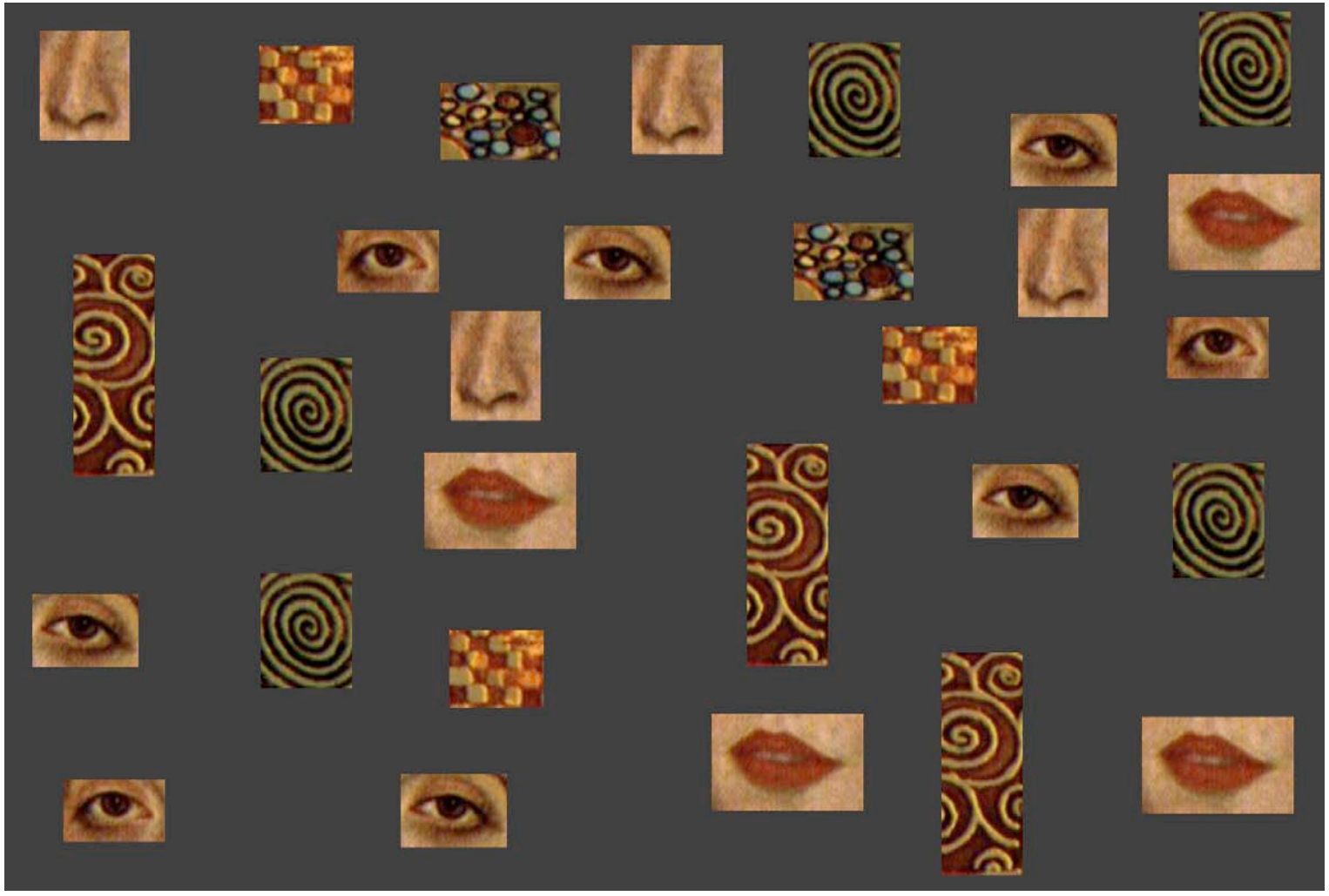


Deformations



Clutter

[slide from Perona]

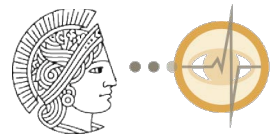


Example





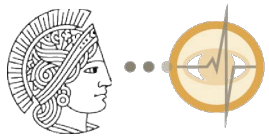
Object Recognition: Focus of today's lecture



- Different Types of Recognition Problems:
  - Object **Identification**
    - recognize your apple, your cup, your dog
  - Object **Classification**
    - recognize any apple, any cup, any dog
    - also called: generic object recognition, object categorization, …
    - typical definition: 'basic level category'

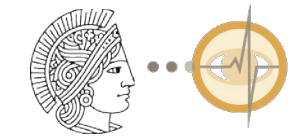


Which Level is right for Object Classes?



- Basic-Level Categories
  - the highest level at which category members have similar perceived shape
  - the highest level at which a single mental image can reflect the entire category
  - the highest level at which a person uses similar motor actions to interact with category members
  - the level at which human subjects are usually fastest at identifying category members
  - the first level named and understood by children
  - (while the definition of basic-level categories depends on culture there exist a remarkable consistency across cultures...)
- Most recent work in object recognition has focused on this problem
  - we will discuss several of the most successful methods in the lecture ;-)

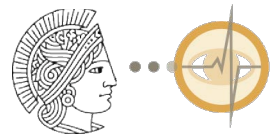
Object Recognition:
Focus of this Computer Vision class



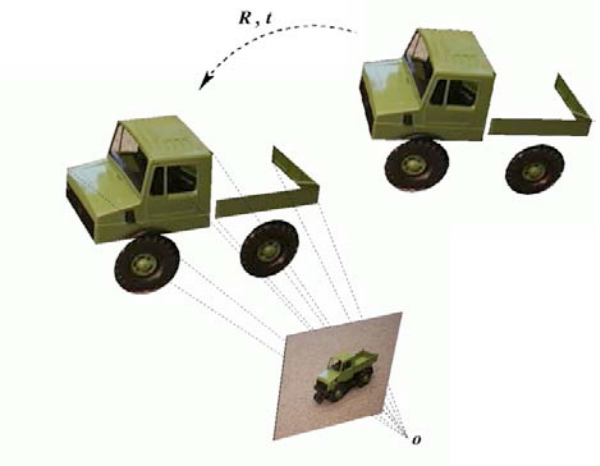
- Recognition and
  - Segmentation: separate pixels belonging to the foreground (object) and the background



Object Recognition:
Focus of this Computer Vision class

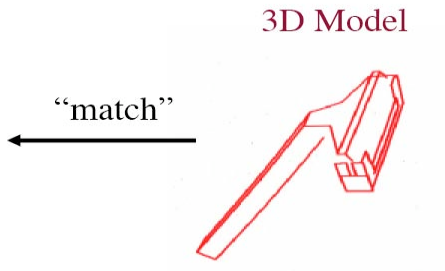


- Recognition and
  - **Localization**: position of the object in the scene, pose estimate (orientation, size/scale, 3D position)



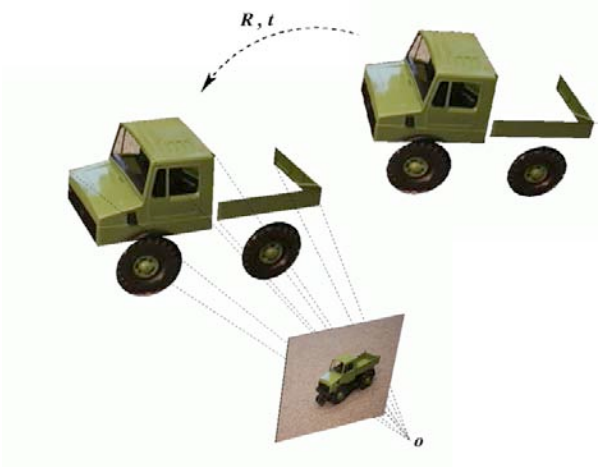
  - Example from David Lowe:

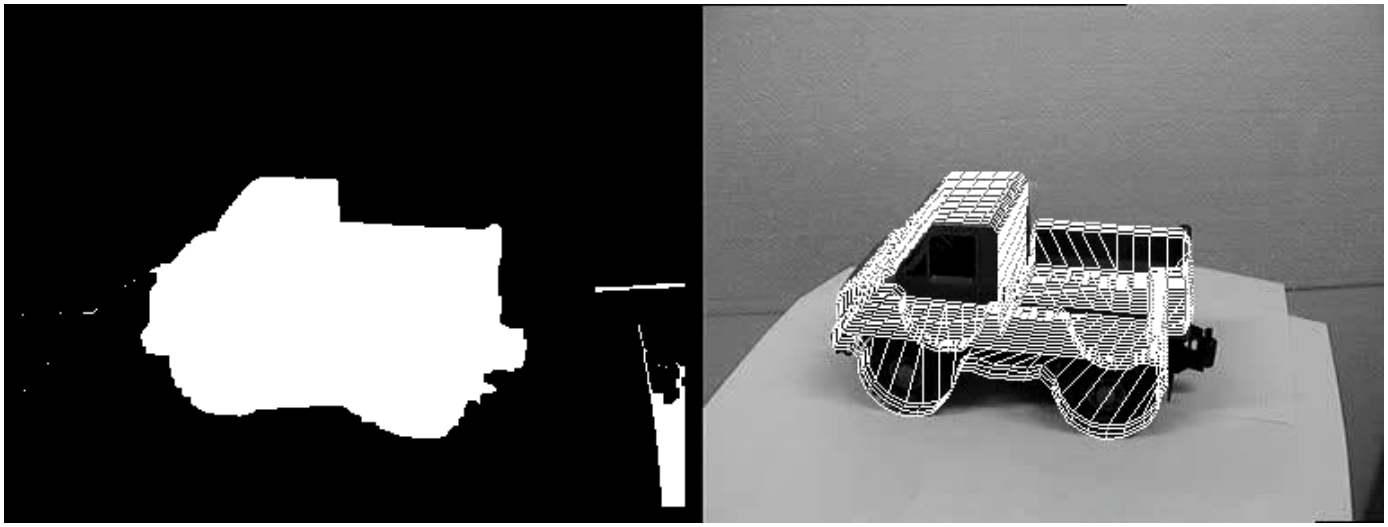




Parameters: 3D position and orientation

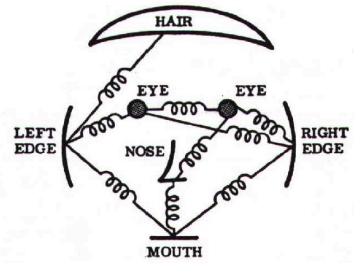
Localization: Example Video 1





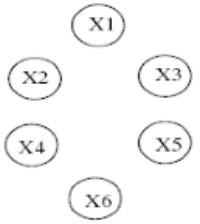
Overview

- Introduction (part 1)
  - why study computer vision in general and object recognition in particular :)

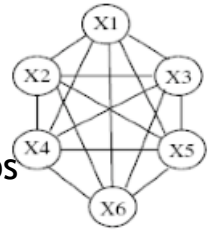


- Object Recognition Methods
  - Bag of Words Models (BoW) (part 2)
    - Model: Histogram of local features
    - e.g. Interest Points (scale invariant)
  - Global Feature Models + Classifier (part 3)
    - e.g. HOG = Histogram of Oriented Gradients
      - global object feature / description
    - e.g. SVM = Support Vector Machines
      - discriminant classifier - widely used
  - Part-Based Object Models (part 4)
    - e.g. Implicit Shape Model (ISM)
    - local parts & global constellation of parts

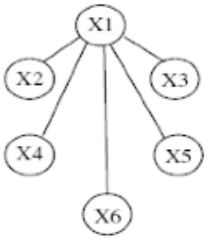
BoW: no spatial relationships



e.g. HOG: fixed spatial relationships



e.g. ISM: flexible spatial relationships



Overview of lecture parts 3 & 4...

TECHNISCHE UNIVERSITÄT DARMSTADT

- Global Feature Based Methods for People Detection (part 3)
  - A Performance Evaluation of Single and Multi-Feature People Detection [Wojek,Schiele@DAGM-08]
  - Pedestrian Detection: A New Benchmark [Dollar,Wojek,Perona,Schiele@CVPR-09]
  - Multi-Cue Onboard Pedestrian Detection [Wojek,Walk,Schiele@CVPR-09]
- Part-Based Model for People & Object Detection (part 4)
  - Detection by Tracking and Tracking by Detection [Andriluka,Roth,Schiele@CVPR-08]
  - Pictorial Structures Revisited: People Detection and Articulated Pose Estimation [Andriluka,Roth,Schiele@CVPR-09]
  - A Shape-Based Object Class Model for Knowledge Transfer [Stark,Goesele,Schiele@ICCV-09]