Course Title:

Image Based Recognition and Classification

Instructors: Josephine Sullivan, Miroslav Kobetski

Lecture 1

Introduction and Motivation

Course details

- Course content
- Marking of the course
- Exercises and Lab

Image Basics

- The digital image
- Image formation
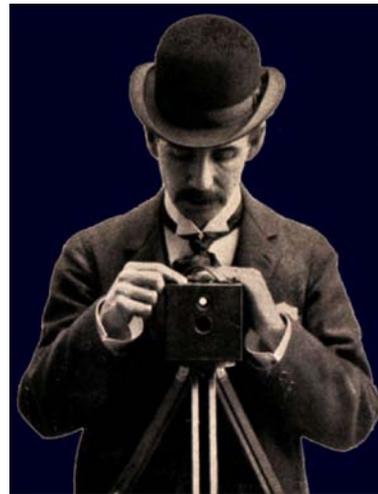- Pixel Description (grayscale and colour)

Introduction and Motivation

Obtaining Images

1670

Camera Obscura

$\longrightarrow$

1860

2000

Camera Phone

It's becoming easier !

Digital Imagery Pervades Society

Even only considering

Personal Digital Cameras 98 million digital cameras were sold world-wide in 2006. Over 2 billion photos have been uploaded onto *Flickr*.

Camera Phones 85% of mobile phones have in-built cameras.

Surveillance Cameras Everywhere An estimated 4 million CCTV cameras in UK.

That's a *huge* amount of data.

Today **almost** all of the analysis, interpretation or cataloging based on semantic content has to be done **manually**.

Digital Images & Computers

Recently computer vision has developed successful techniques for

computer graphics/image processing

geometric measurement

Digital Images & Computers

Modifying the initial image(s): Mosaic generation
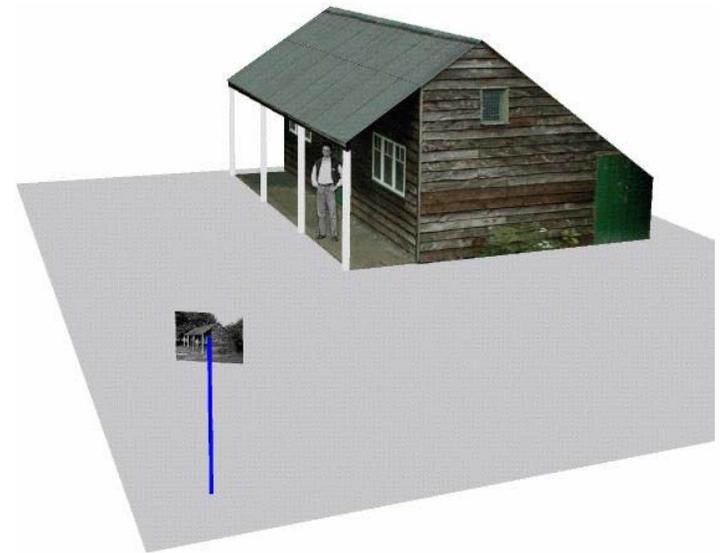

(a) Matier data set (7 images)

Image Editing

Remove unwanted objects from a scene

Digital Images & Computers

Visualization, building 3D models of real world objects from images

Inserting artificial objects to real images

Digital Images & Computers

However, the more exciting and perhaps necessary goal of computer vision is the:

> Automatic **description** and **recognition** of the world, via visual images, in terms **humans** care about.

Journey from

Measurement and Manipulation
to
Perception

What & Where

Given an image what do humans see or notice ?

Query Task

Does the image contain a car ?
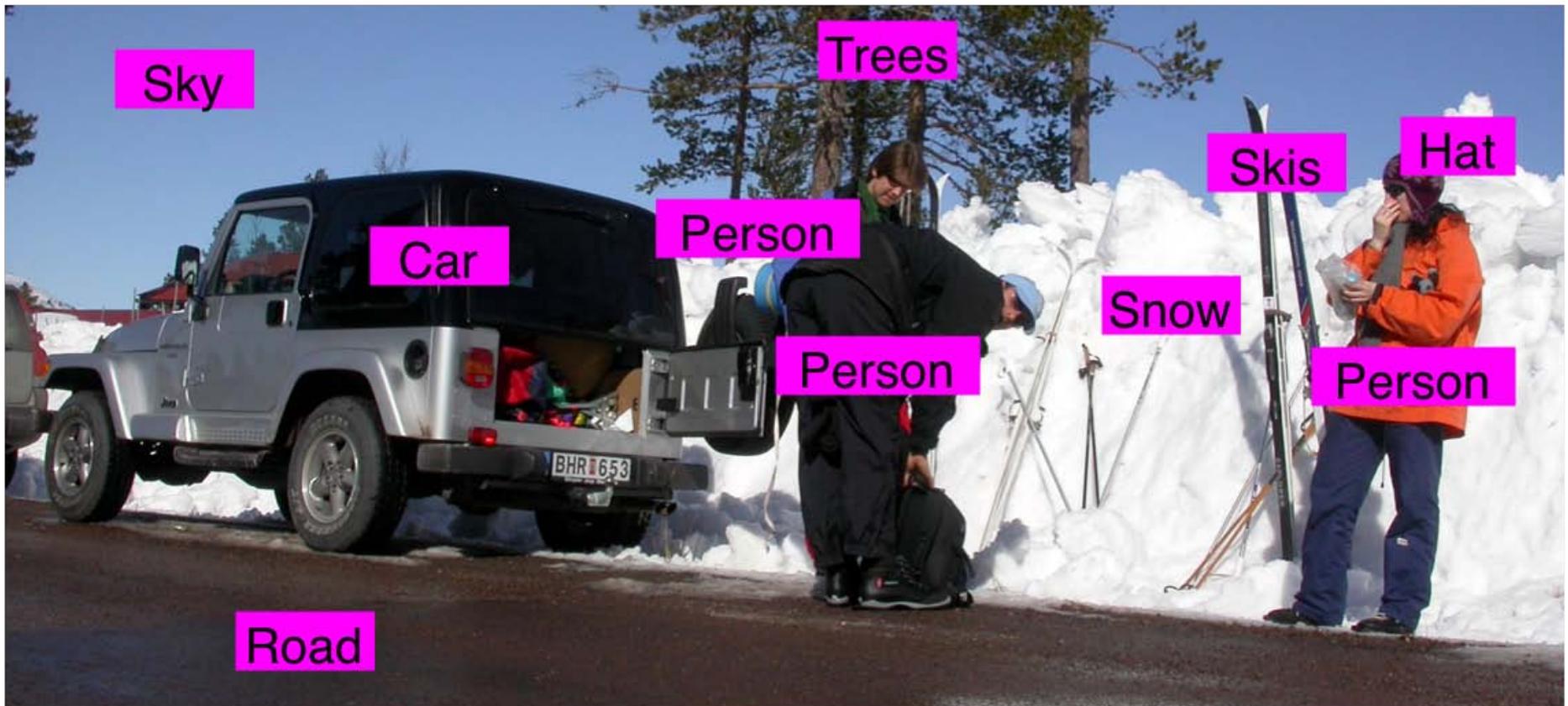
Detection

Where are the people ?

Object Categorization

What type of objects present and where ?

Scene Understanding

It is a very sunny day at a ski resort. Three friends are unpacking (packing) their car. They are either about to embark on a skiing trip or have just returned from one.

A picture is worth a thousand words !

Potential Application Domains

Home Robots

Evolution Robotics

Cataloging and Searching Photo and Video Collections

Surveillance

Presently, computers cannot replicate human *image analysis*

Why Not ??

Let's focus on the task of object recognition.

Challenges:

View Point Variation

Illumination Variation

Occlusion

Deformation

Intra-class variation

Current State-of-the-Art

- Detection of specific objects in stereo-typical poses

    faces, cars, pedestrians

- Classification of the scene environment

    indoor, urban/city, forest, building

What has been achieved

Handwritten Digit Recognition

Typical: 98-99% correct classification

Finding Frontal Faces

Finding People
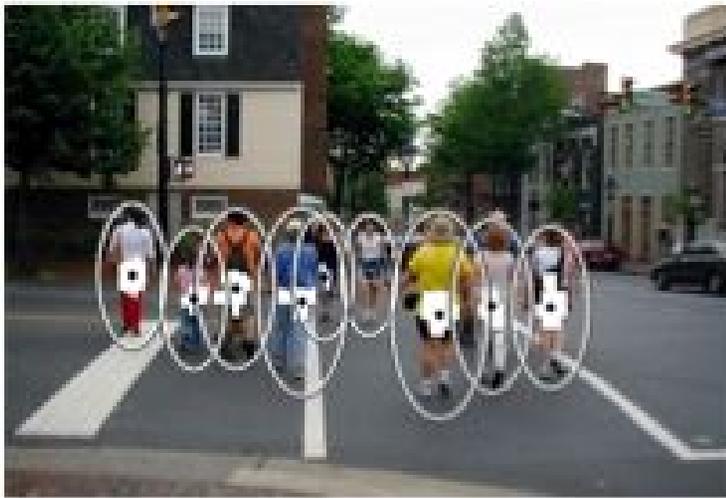
Finding skeletal joints: Kinect

depth image ➡ body parts ➡ 3D joint proposals
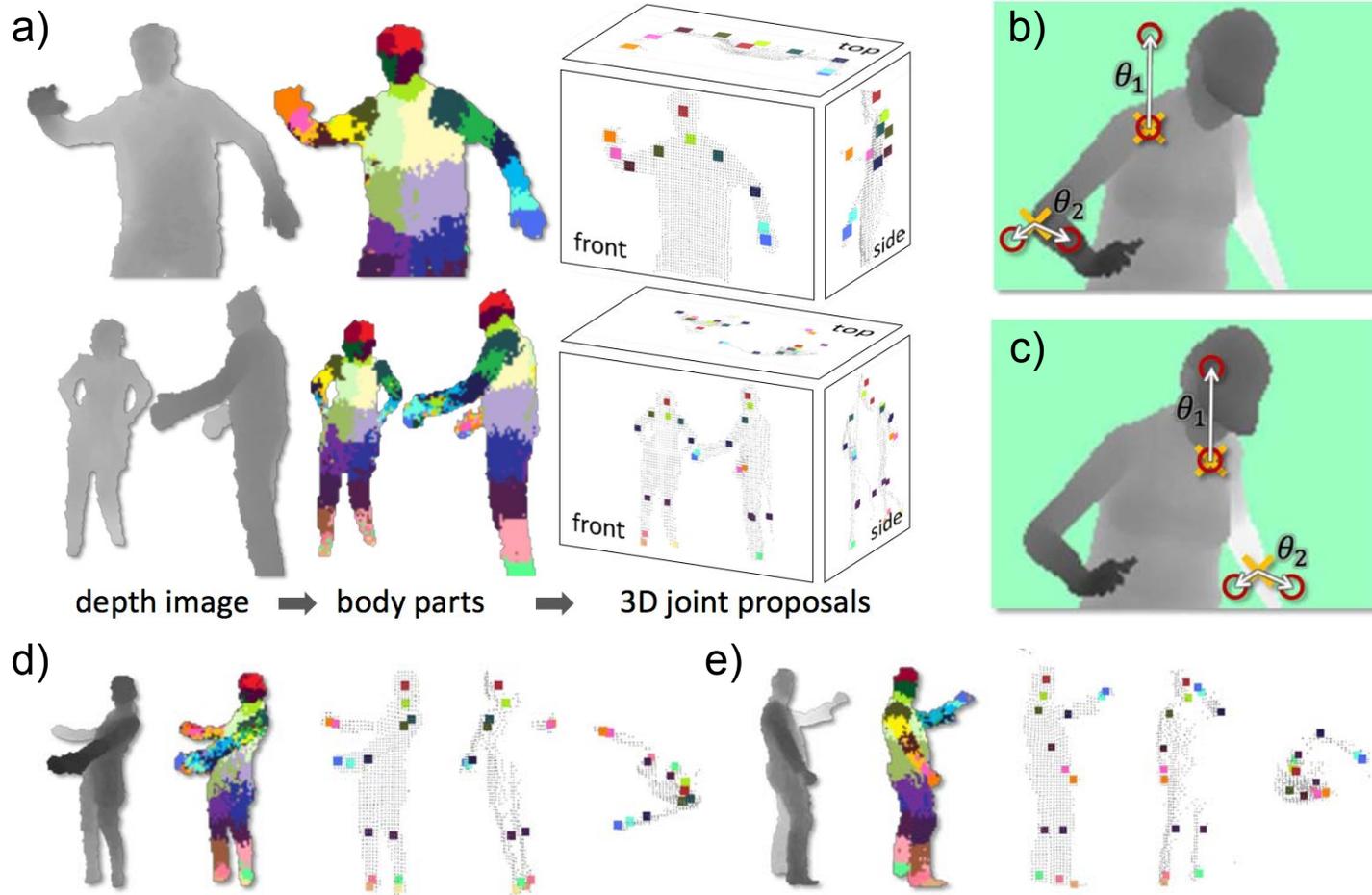
Recovery of Scene Layout

| Input | Location | Colour | Texture | Perspective | All Cues |
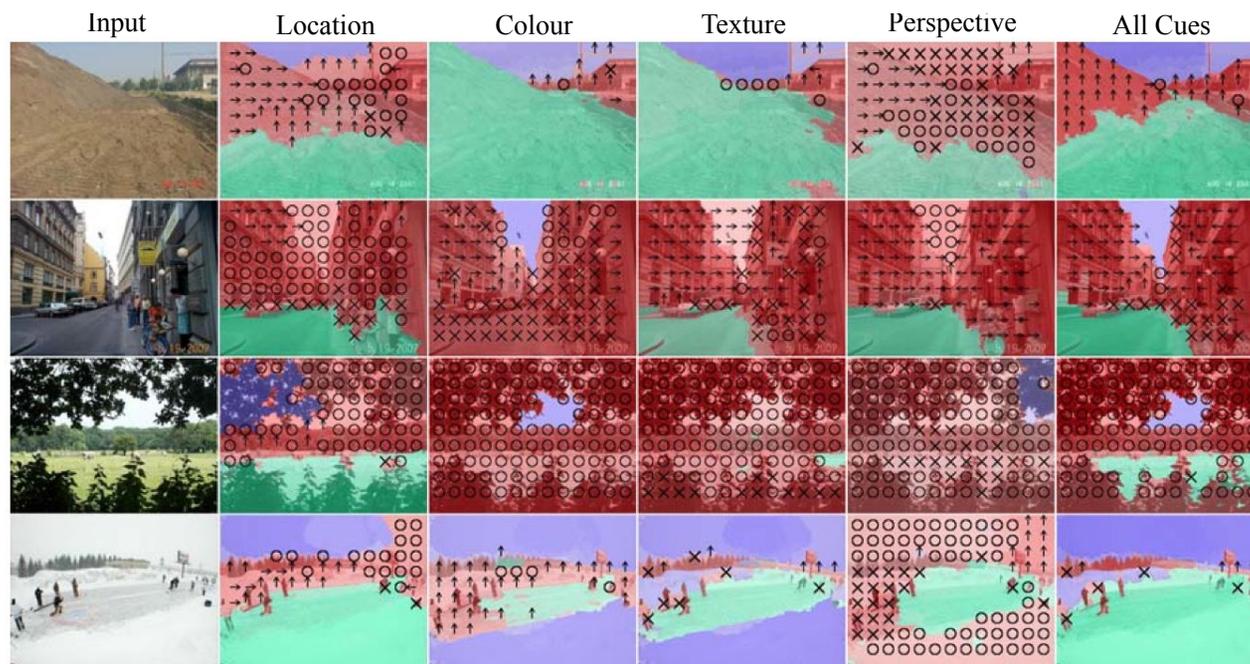
- For each pixel return label of Sky, Support surface, Vertical surface.

- Vertical surfaces are then sub-classified

  - into planar objects at different orientations (shown by arrows) **and**
  - non-planar objects which are either porous ('o') or non-porous ('x').

Semantic Labelling of an image

The Common Denominator

These applications **do not rely** on traditional vision methods

Physics-Based Methods - Understanding and exploiting the physics of imaging i.e., how the light from different sources reflects from different surfaces of differing shapes and materials.

Geometry-Based Methods - Recovery of 3D structure from 2D images.

They exploit in some form

Learning-Based Methods - Solve problem by reference to training data (e.g., have a face image if it looks like an image which I know is a face.)

Trend fueled by the rapid growth of computational power, memory and the abundance of visual images and video and the web.

Why must computers learn ?

Unclear how to model object categories, so we must learn what distinguishes them rather than manually specify the difference.
How do computers learn ? . . .

The Main Issues

Representation

How to represent an object and/or object category

Classification Design

How to form the classifier, given the training data

Recognition

How the classifier is used on novel data

Representation I

An image is an array of numbers

| 34 | 45 | 53 | 55 | 69 | 79 | 91 | 95 | 105 | 197 | 254 | 250 | 254 | 254 | 254 | 254 | 254 | 254 | 254 | 254 |
0	11	20	39	59	58	62	73	67	92	213	255	254	254	254	254	254	254	254	254
5	5	0	11	30	16	39	87	67	27	167	255	254	254	254	254	254	254	254	254
0	0	10	12	8	5	73	172	172	140	204	255	254	254	254	254	254	254	254	254
5	0	17	0	0	20	123	237	249	255	246	250	254	254	254	254	254	254	254	254
0	16	9	0	0	48	200	255	242	255	255	255	255	255	254	252	251	252	253	254
7	0	0	5	23	175	234	250	243	250	253	254	251	251	252	252	253	253	254	254
0	0	17	0	17	198	255	248	250	246	255	245	254	255	255	255	255	253	252	250
0	16	2	14	69	125	247	255	255	247	255	249	253	253	253	254	253	253	252	251
26	15	1	109	181	102	148	235	254	240	249	252	250	250	250	250	251	252	254	255
0	0	44	203	249	169	69	208	255	248	255	255	255	255	255	255	255	253	251	250
4	47	156	232	255	245	115	166	244	253	249	245	244	247	252	255	255	254	251	249
114	193	251	253	247	255	191	88	153	185	207	182	200	209	224	240	251	255	255	255
193	255	255	213	147	131	97	63	59	77	86	81	88	110	123	112	156	199	250	245
228	178	151	113	9	4	17	40	43	32	42	68	65	53	36	74	70	75	121	215
171	52	33	0	13	0	0	31	44	29	32	55	61	72	71	107	91	55	62	165
47	0	17	11	40	28	22	33	52	68	76	80	78	101	119	110	124	63	74	170
21	22	19	30	26	45	59	60	64	77	85	84	93	101	125	120	117	30	56	213
35	40	27	51	52	51	57	66	66	55	48	49	92	108	108	101	52	0	18	195
27	19	52	89	56	31	19	34	45	41	40	47	67	66	39	15	18	45	51	159

Representation II

Feature Extraction – Summarize the pattern of the number array by applying measurements i.e., oriented edge filters

| 34	45	53	55	69	79	91	95	105	197	254	250	254	254	254	254	254	254	254	254
0	11	20	39	59	58	62	73	67	92	213	255	254	254	254	254	254	254	254	254
5	5	0	11	30	16	39	87	67	27	167	255	254	254	254	254	254	254	254	254
0	0	10	12	8	5	73	172	172	140	204	255	254	254	254	254	254	254	254	254
5	0	17	0	0	20	123	237	249	255	246	250	254	254	254	254	254	254	254	254
0	16	9	0	0	48	200	255	242	255	255	255	255	255	254	252	251	252	253	254
7	0	0	5	23	175	234	250	243	250	253	254	251	251	252	252	253	253	254	254
0	0	17	0	17	198	255	248	250	246	255	245	254	255	255	255	255	253	252	250
0	16	2	14	69	125	247	255	255	247	255	249	253	253	253	254	253	253	252	251
26	15	1	109	181	102	148	235	254	240	249	252	250	250	250	250	251	252	254	255
0	0	44	203	249	169	69	208	255	255	248	255	255	255	255	255	255	253	251	250
4	47	156	232	255	245	115	166	244	253	249	245	244	247	252	255	255	254	251	249
114	193	251	253	247	255	191	88	153	185	207	182	200	209	224	240	251	255	255	255
193	255	255	213	147	131	97	63	59	77	86	81	88	110	123	112	156	199	250	245
228	178	151	113	9	4	17	40	43	32	42	68	65	53	36	74	70	75	121	215
171	52	33	0	13	0	0	31	44	29	32	55	61	72	71	107	91	55	62	165
47	0	17	11	40	28	22	33	52	68	76	80	78	101	119	110	124	63	74	170
21	22	19	30	26	45	59	60	64	77	85	84	93	101	125	120	117	30	56	213
35	40	27	51	52	51	57	66	66	55	48	49	92	108	108	101	52	0	18	195
27	19	52	89	56	31	19	34	45	41	40	47	67	66	39	15	18	45	51	159

$$\Rightarrow \mathbf{f} = (f_1, f_2, \cdots)$$

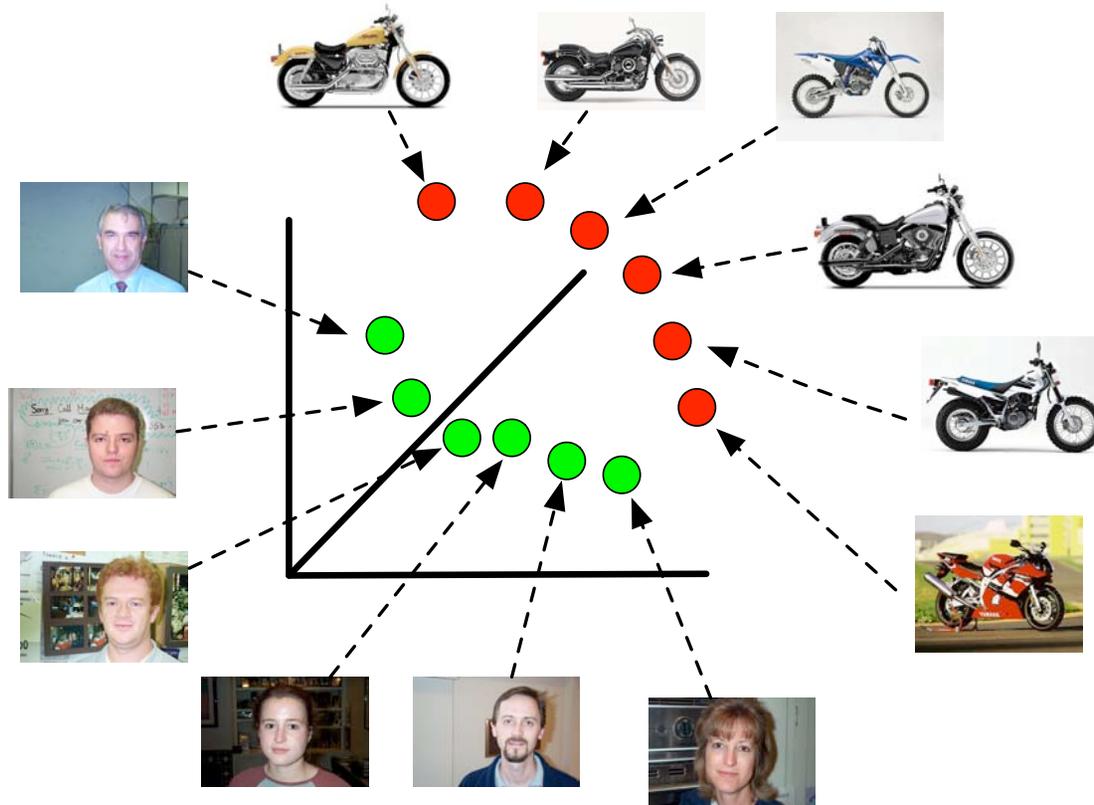# Representation III

Feature Vector
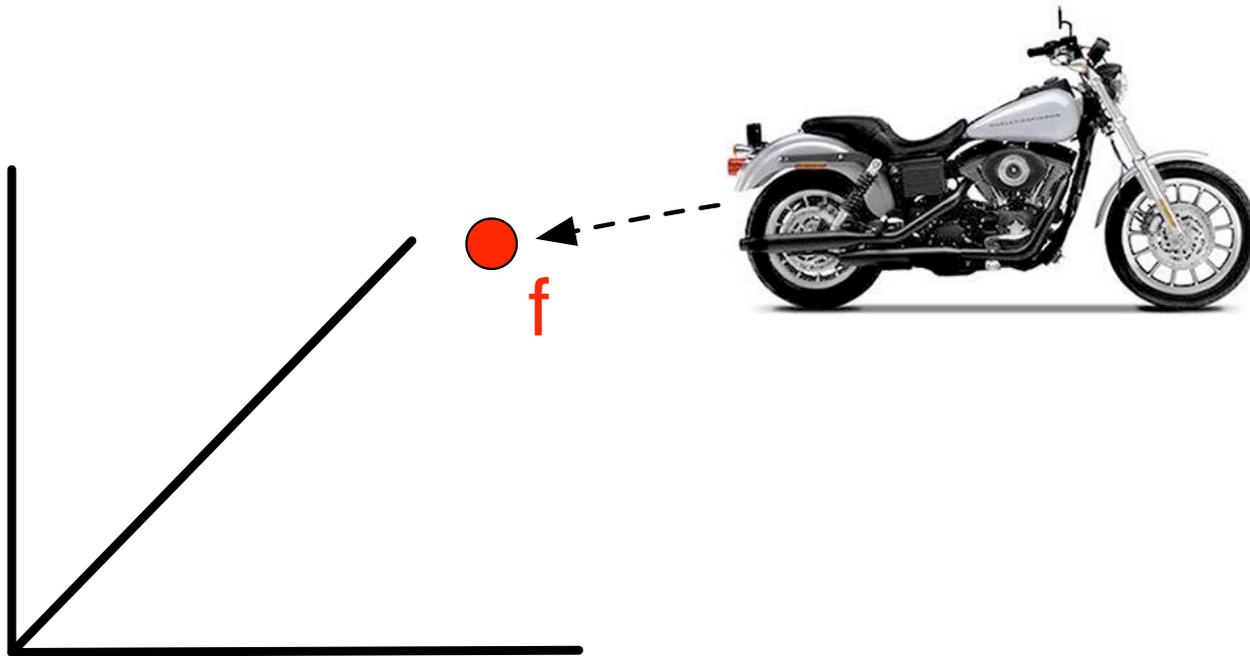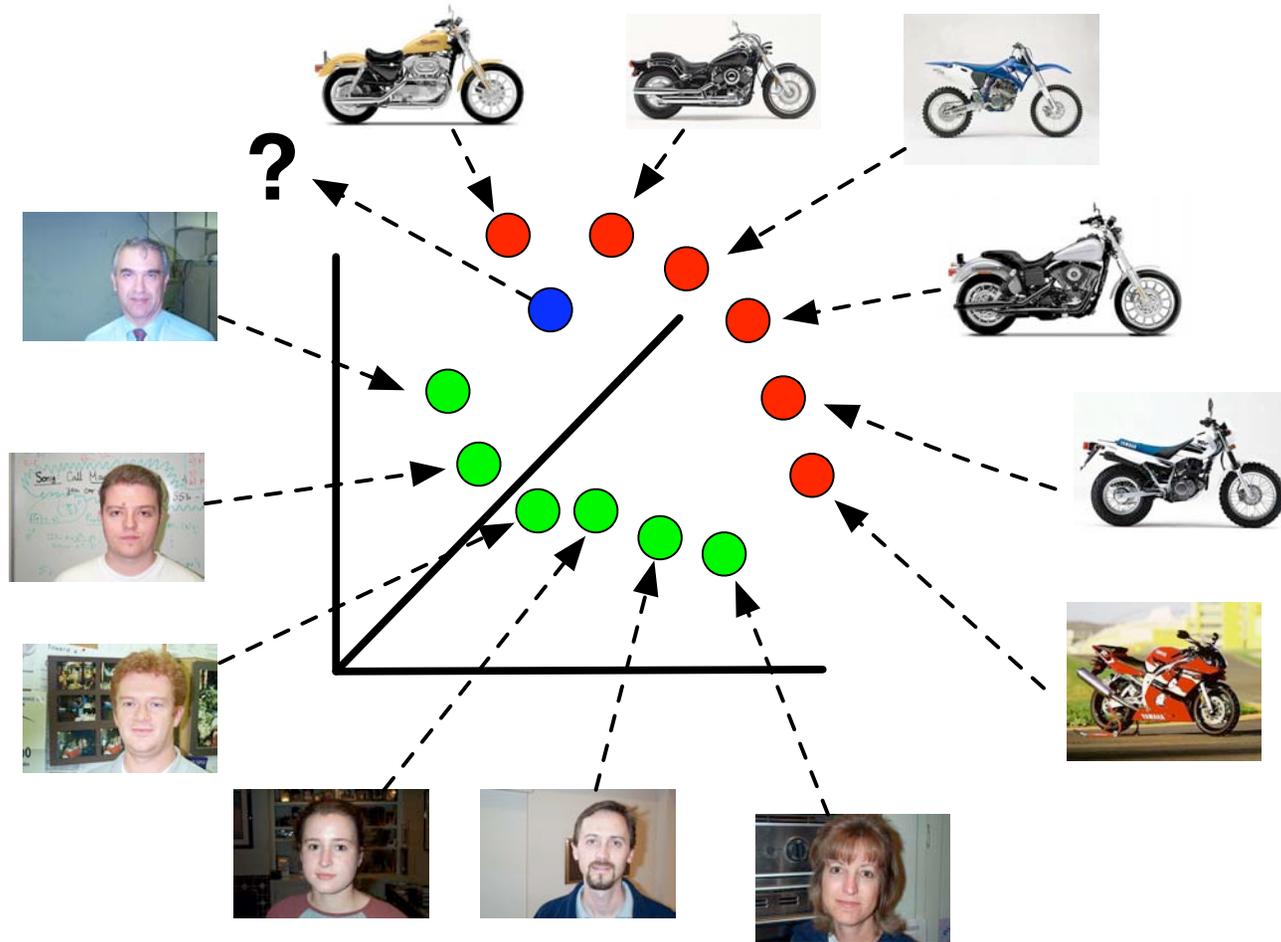
Learning from examples
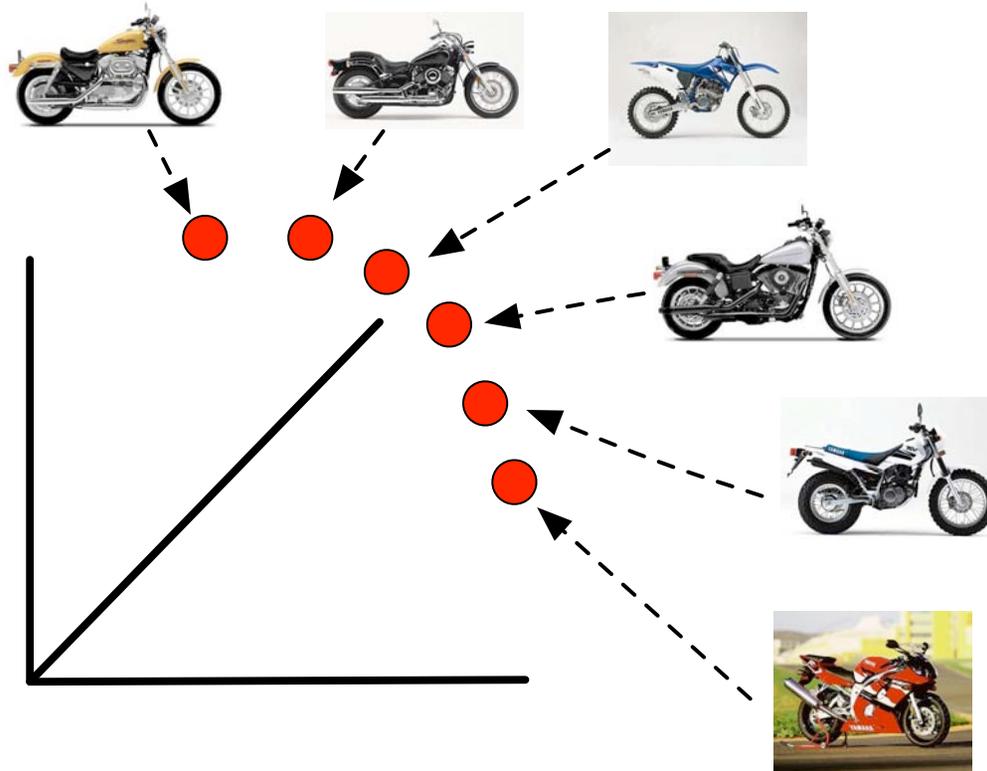
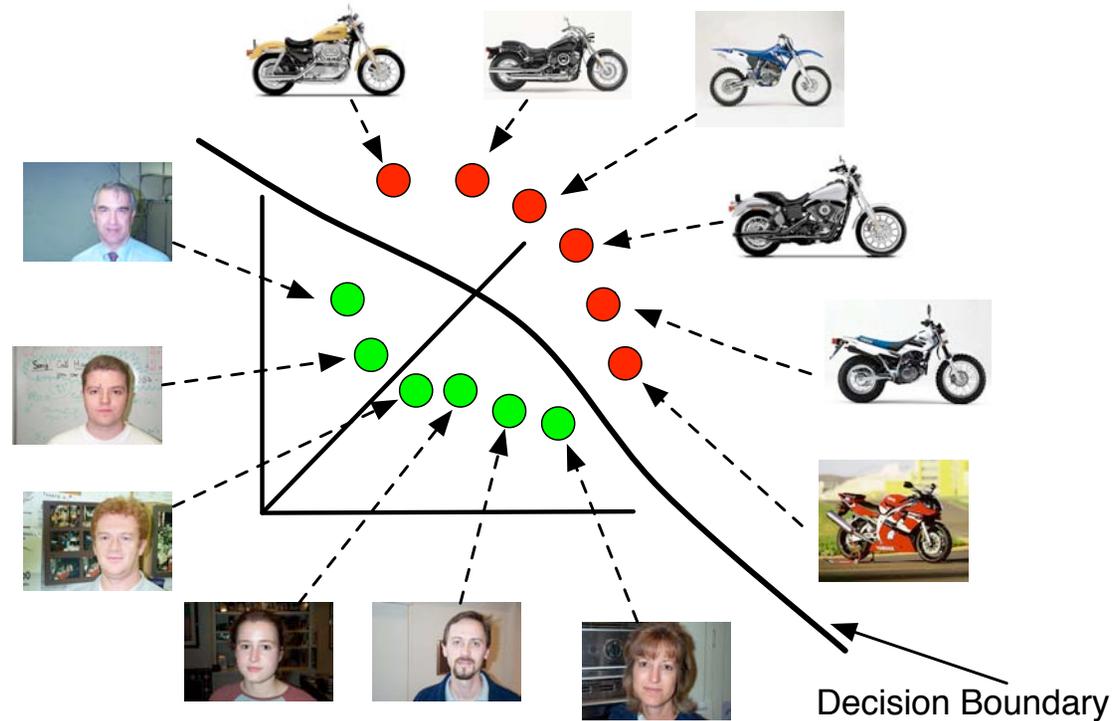
Many feature vectors from a category

Learning from examples

Want different categories to occupy different parts of the space.

Is it a bike or a face ?

# Discriminative Categorization

Decision Boundary

Generative Categorization

In this case consider:

- an object class forms a *volume* in the high dimensional feature space.

- Model the shape of this volume and construct the likelihood $p(\mathbf{f}|\text{object})$.

- This problem involves the combination of *empirical data analysis*, *mathematical* and *statistical* modeling in high dimensional spaces.

Generative Categorization
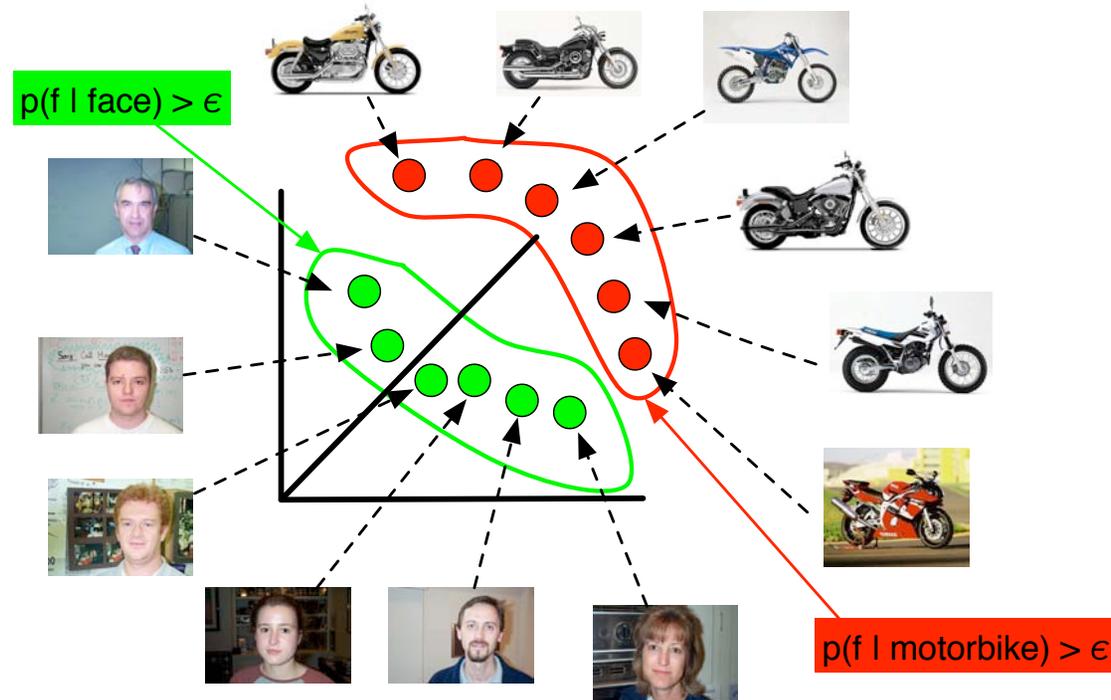
Decision boundary:
$$p(\mathbf{f}|\text{face})p(\text{face}) = p(\mathbf{f}|\text{motorbike})p(\text{motorbike}).$$