Computer Vision – TP10
Pattern Recognition

Miguel Tavares Coimbra

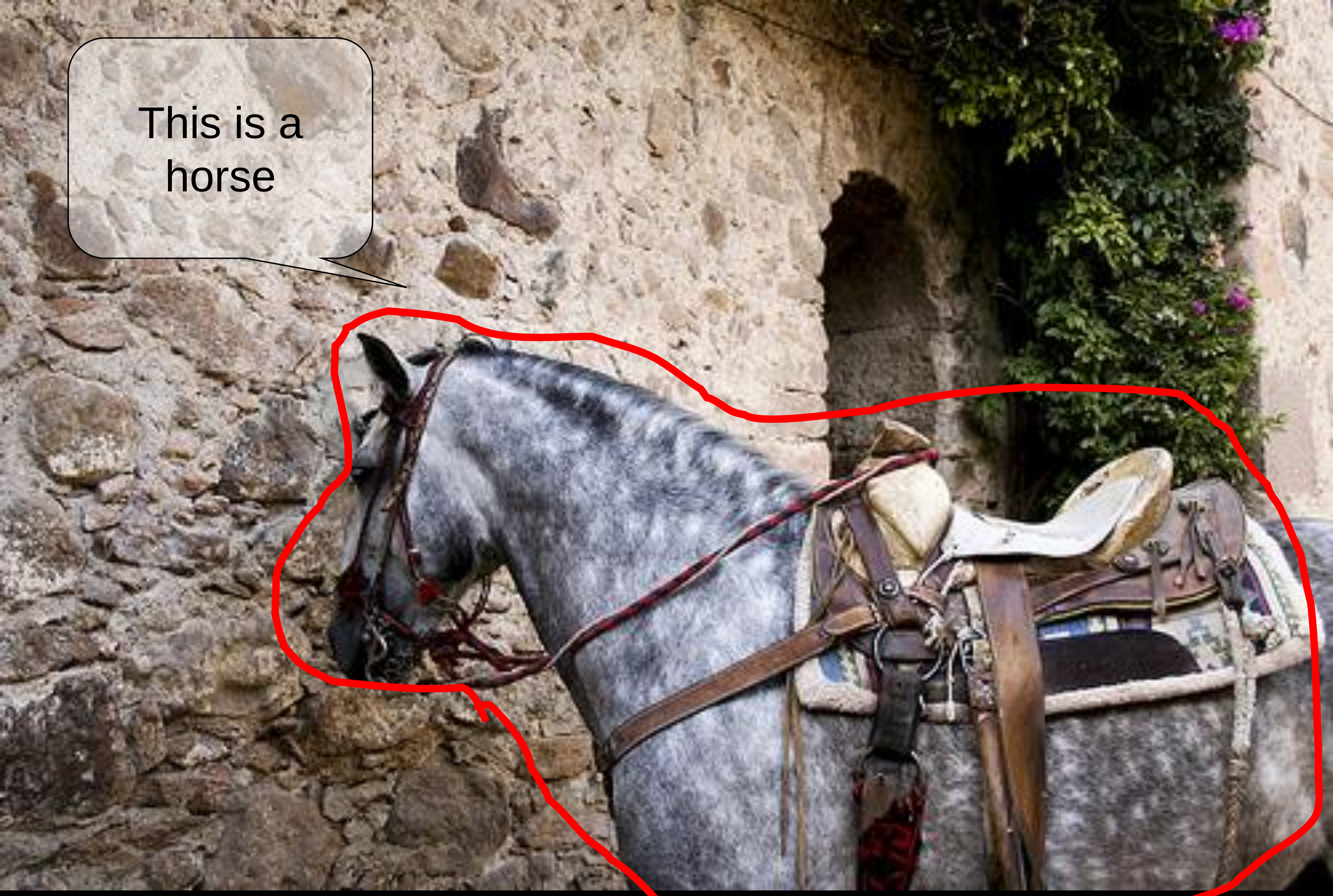
Outline

- Introduction to Pattern Recognition
- Statistical Pattern Recognition
- Visual Features

Topic: Introduction to Pattern Recognition

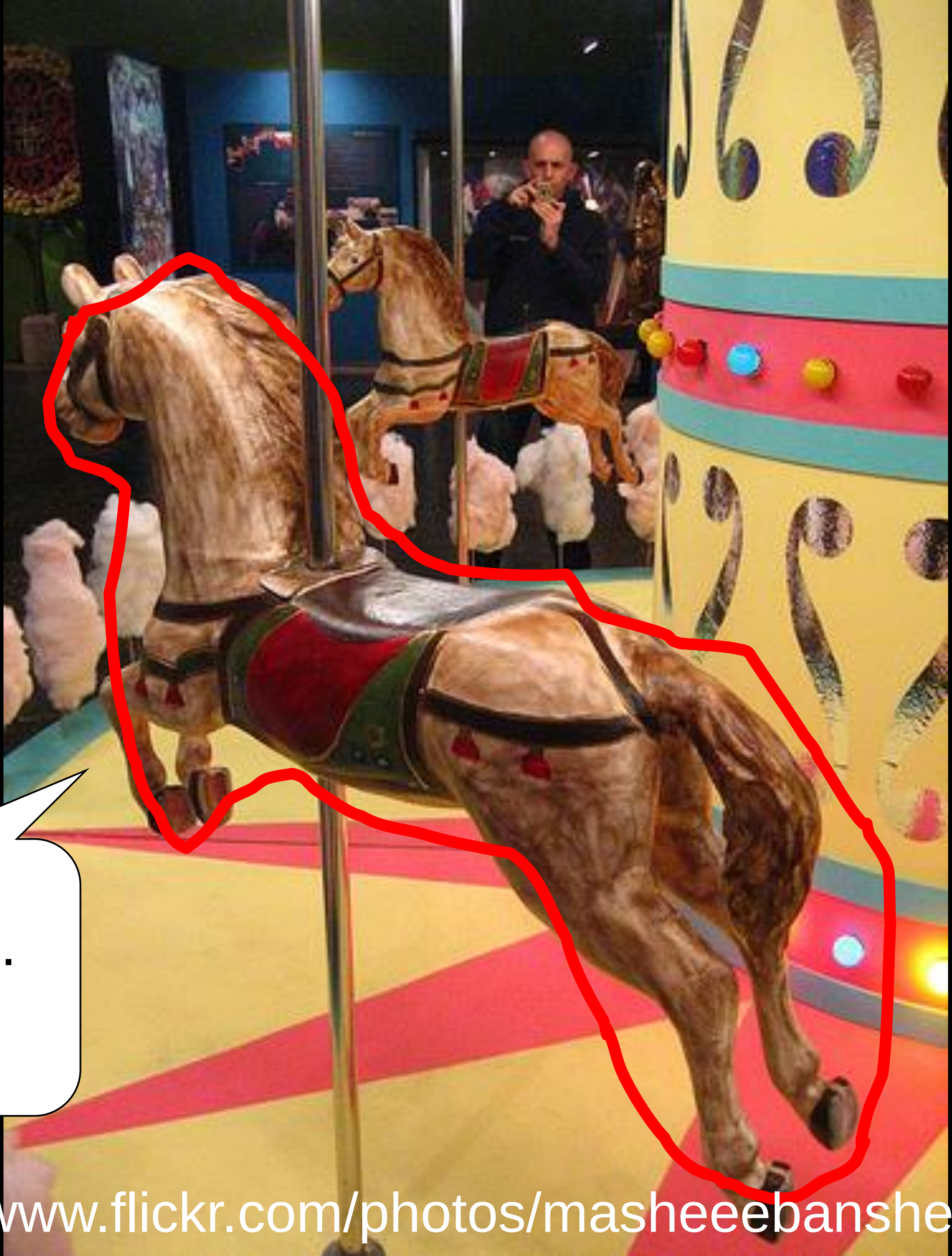
- Introduction to Pattern Recognition
- Statistical Pattern Recognition
- Visual Features

U. PORTO FC



http://www.flickr.com/photos/kimbar/2027234083/

http://www.flickr.com/photos/genewolf/2031802050/

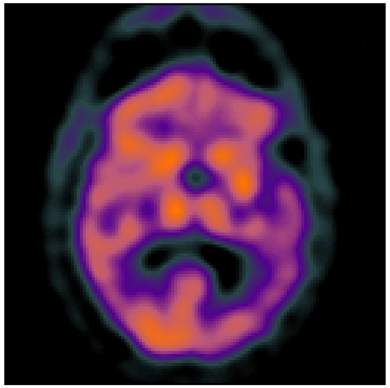
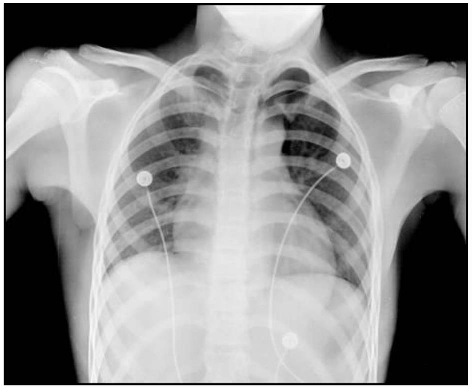
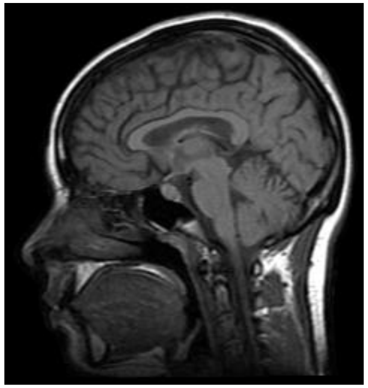


This is a...
Horse?

http://www.flickr.com/photos/masheeebanshee/413465808/

Decisions

- I can **manipulate** images
- I want to make **decisions**!

- Classify / Identify **features**
- Recognize **patterns**

One definition

- Pattern recognition

  "the act of taking in raw data and taking an action based on the category of the data"

  Wikipedia

- How do I do this so well?
- How can I make machines do this?

The problem

Do you ‘see’ a horse?

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

U. PORTO FC

Mathematics

- We only deal with numbers.
  - How do we represent knowledge?
  - How do we represent visual features?
  - How do we classify them?
- Very complex problem!!
  - Let's break it into smaller ones...

Typical PR system

Sensor

Gathers the observations to be classified or described

Feature Extraction

Computes numeric or symbolic information from the observations;

Classifier

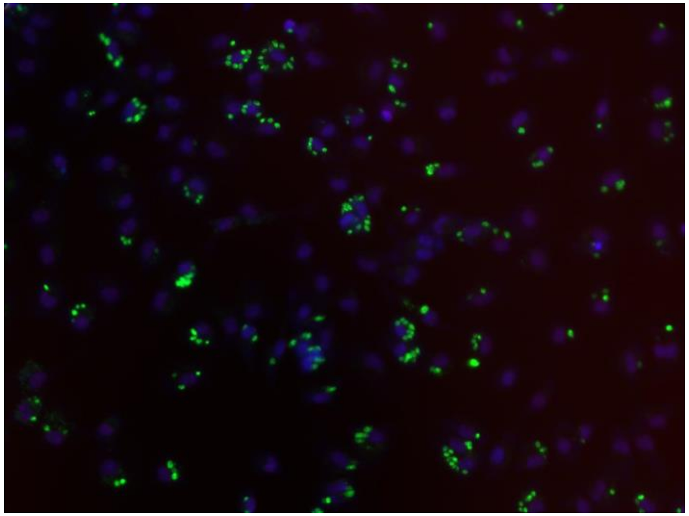
Does the actual job of classifying or describing observations, relying on the extracted features.

Sensor

- In our specific case:
  - Image acquiring mechanism
  - Let’s assume we don’t control it

One observation = One Image

Video = Multiple Observations



Feature Extraction

- What exactly are features?
  - Colour, texture, shape, etc
  - Animal with 4 legs
  - Horse
  - Horse jumping
- These vary a lot!
- Some imply some sort of 'recognition'

  e.g. How do I know the horse is jumping?

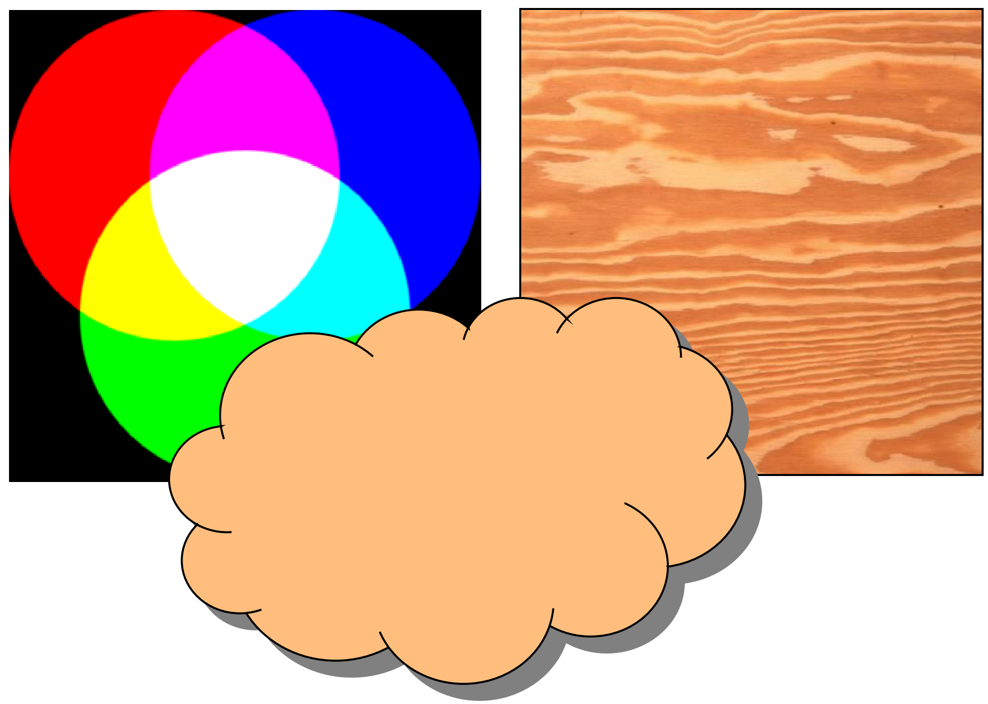
U. PORTO FC

Broad classification of features

- Low-level
  - Colour, texture
- Middle-level
  - Object with head and four legs
  - Object moving up
  - Horse
- High-level
  - Horse jumping
  - Horse competition

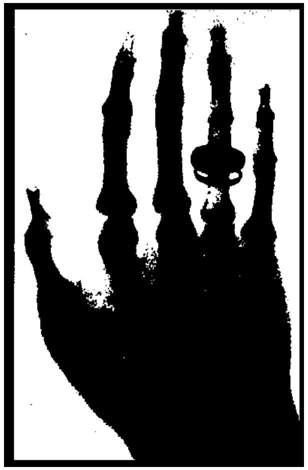
Low-level features

- Objective
- Directly reflect specific image and video features
  - Colour
  - Texture
  - Shape
  - Motion
  - Etc.



Middle-level features

- Some degree of subjectivity
- They are typically one solution of a problem with multiple solutions
- Examples:
  - Segmentation
  - Optical Flow
  - Identification
  - Etc.



High-level features

- Semantic Interpretation
- Knowledge
- Context
- Examples:
  - This person suffers from epilepsy
  - The virus attacks the cell with some degree of intelligence
  - This person is running from that one

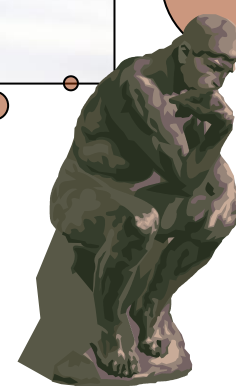
How do humans do this so well?

The semantic gap

- Fundamental problem of current research!

Low-level:
-Colour
-Texture
-Shape
-…

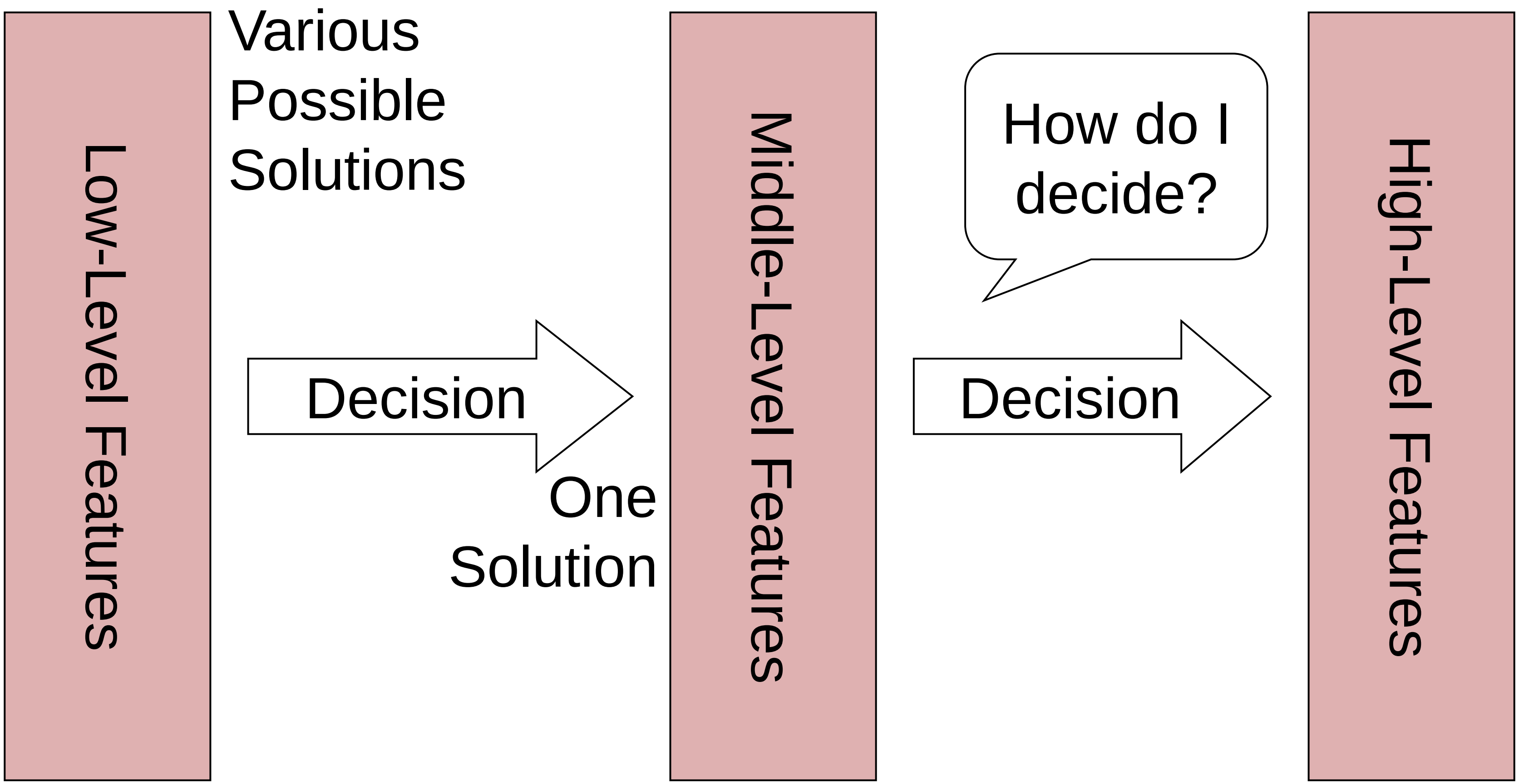
High-level:
-Interpretation
-Decision
-Understanding
-…



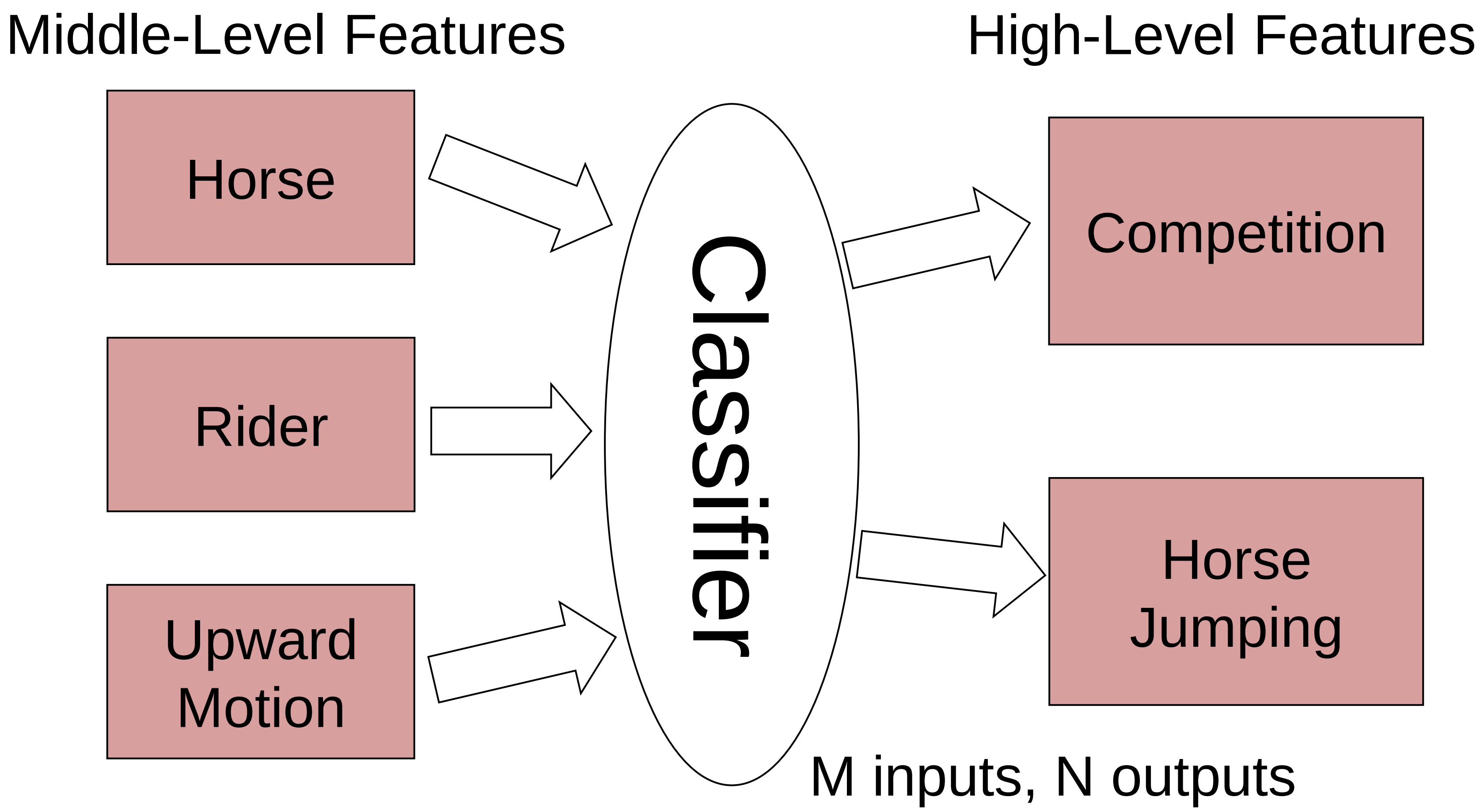
Now what??
How do i cross this bridge?

U. PORTO FC

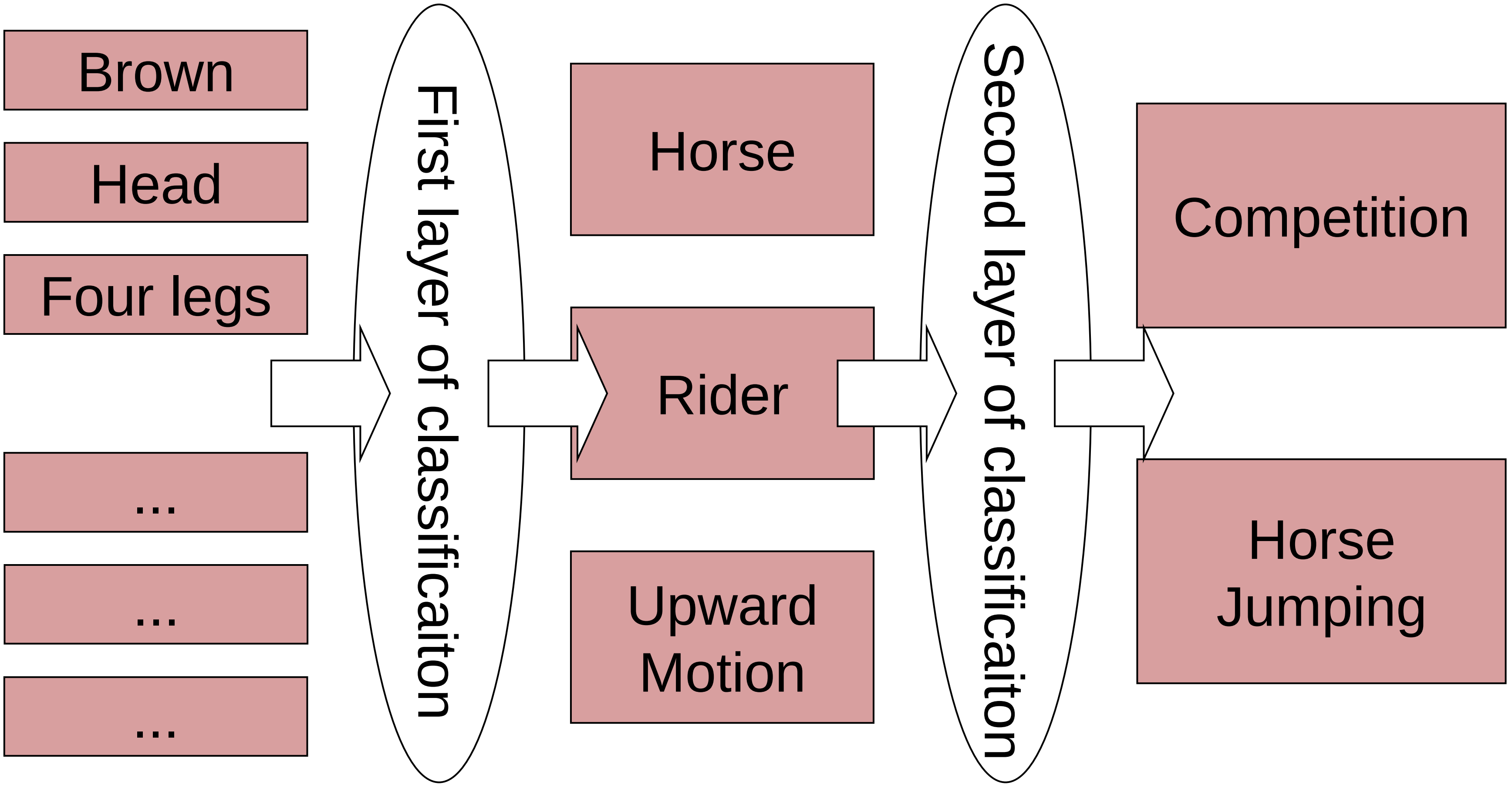
Features & Decisions



Classification



Layers of classification



Classifiers

- How do I map my M inputs to my N outputs?
- Mathematical tools:
  - Distance-based classifiers
  - Rule-based classifiers
  - Neural Networks
  - Support Vector Machines
  - ...

U. PORTO FC

Types of PR methods

- Statistical pattern recognition
  – based on statistical characterizations of patterns, assuming that the patterns are generated by a probabilistic system
- Syntactical (or structural) pattern recognition
  – based on the structural interrelationships of features

U.PORTO FC

Topic: Statistical Pattern Recognition

- Introduction to Pattern Recognition
- Statistical Pattern Recognition
- Visual Features

U. PORTO FC

Is Porto in Portugal?

Porto is in Portugal

- I want to make decisions
  - Is Porto in Portugal?
- I know certain things
  - A world map including cities and countries
- I can make this decision!
  - Porto **is** in Portugal
- I had enough ***a priori*** knowledge to make this decision

U. PORTO FC

What if I don't have a map?

- I still want to make this decision
- I observe:
  - Amarante has coordinates x_1,y_1 and is in Portugal
  - Viseu has coordinates x_2, y_2 and is in Portugal
  - Vigo has coordinates x_3, y_3 and is in Spain
- I classify:
  - Porto is close to Amarante and Viseu so **Porto <u>is</u> in Portugal**
- What if I try to classify *Valença*?

U.PORTO FC

Statistical PR

- I used **statistics** to make a decision
  - I can make **decisions** even when I don’t have full a priori knowledge of the whole process
  - I can make **mistakes**
- How did I **recognize** this pattern?
  - I **learned** from previous observations where I knew the classification result
  - I **classified** a new observation

What pattern?

U. PORTO FC

Back to the Features

- Feature F_i $F_i = [f_i]$
- Feature F_i with N values.

$$F_i = [f_{i1}, f_{i2}, \ldots, f_{iN}]$$

- Feature vector F with M features.

$$F = [F_1 \mid F_2 \mid \ldots \mid F_M]$$

- Naming conventions:
  - Elements of a **feature vector** are called **coefficients**
  - **Features** may have one or more **coefficients**
  - **Feature vectors** may have one or more **features**

U.PORTO FC

Back to our Porto example

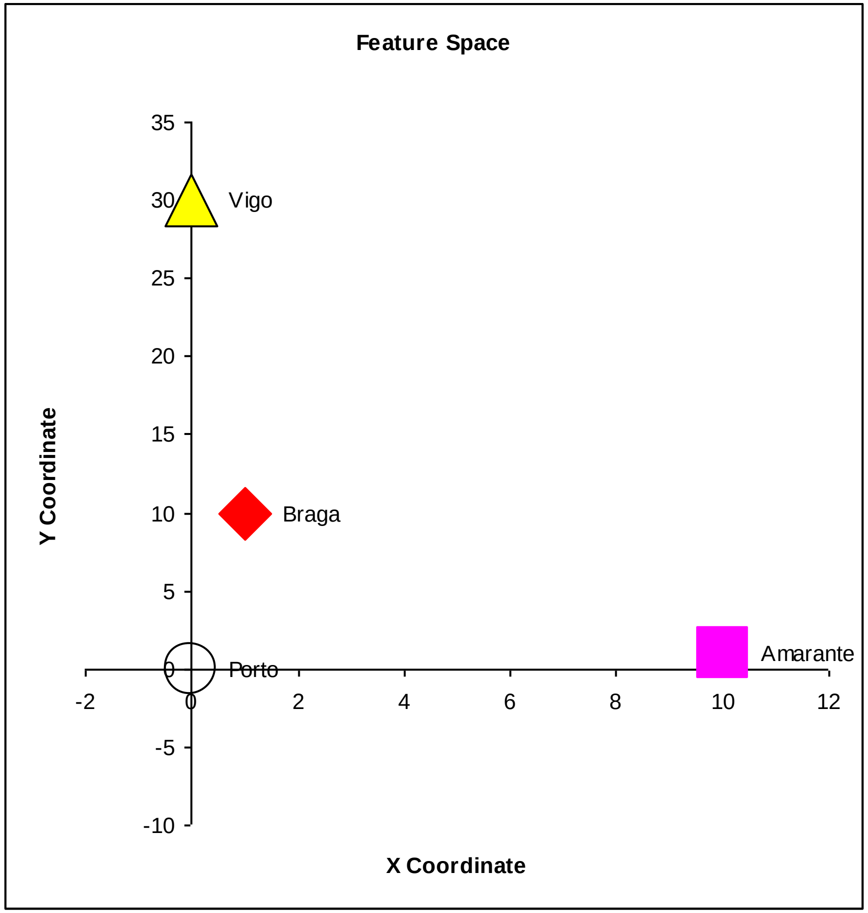
- I’ve classified that Porto is in Portugal
- What feature did I use?
  - Spatial location
- Let’s get more formal
  - I’ve defined a feature vector ***F*** with one feature $\boldsymbol{F_1}$, which has two coefficients $\boldsymbol{f_{1x}}$, $\boldsymbol{f_{1y}}$

$$F = [F_1] = [f_{1x}, f_{1y}]$$

U.PORTO FC

Feature Space

- Feature Vector
  - Two total coefficients
  - Can be seen as a feature 'space' with two orthogonal axis
- Feature Space
  - Hyper-space with N dimensions where N is the total number of coefficients of my feature vector



A Priori Knowledge

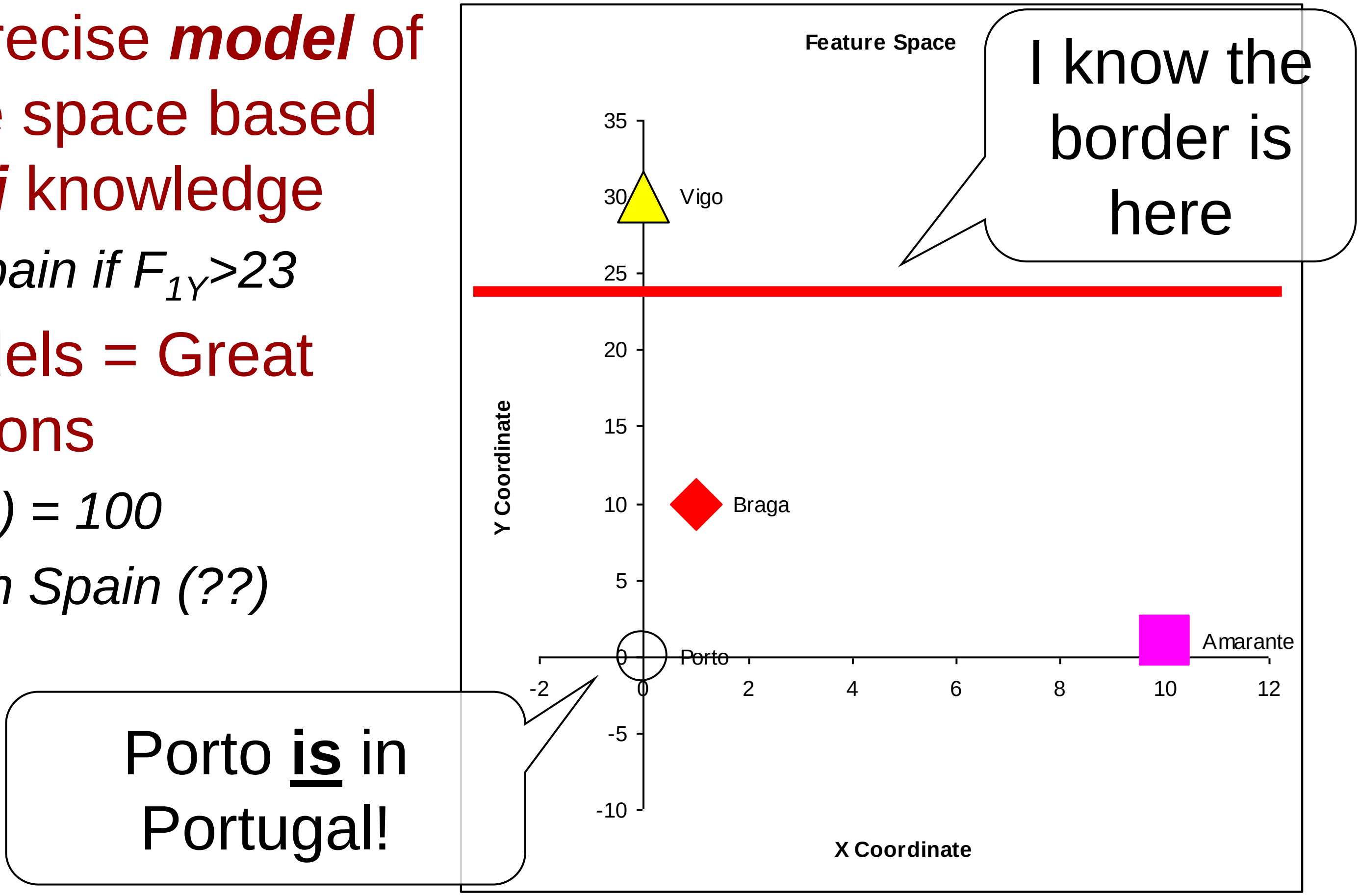
- I have a precise ***model*** of my feature space based on ***a priori*** knowledge

  City is in Spain if F_{1Y}>23

- Great models = Great classifications

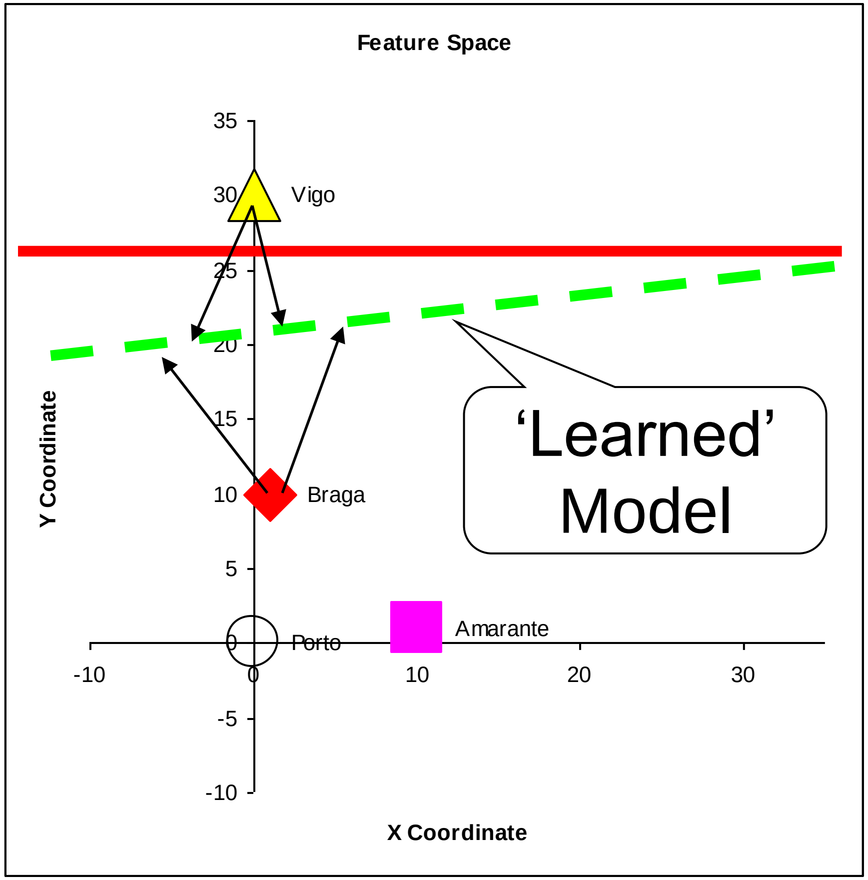
  F_{1Y}(London) = 100

  London is in Spain (??)



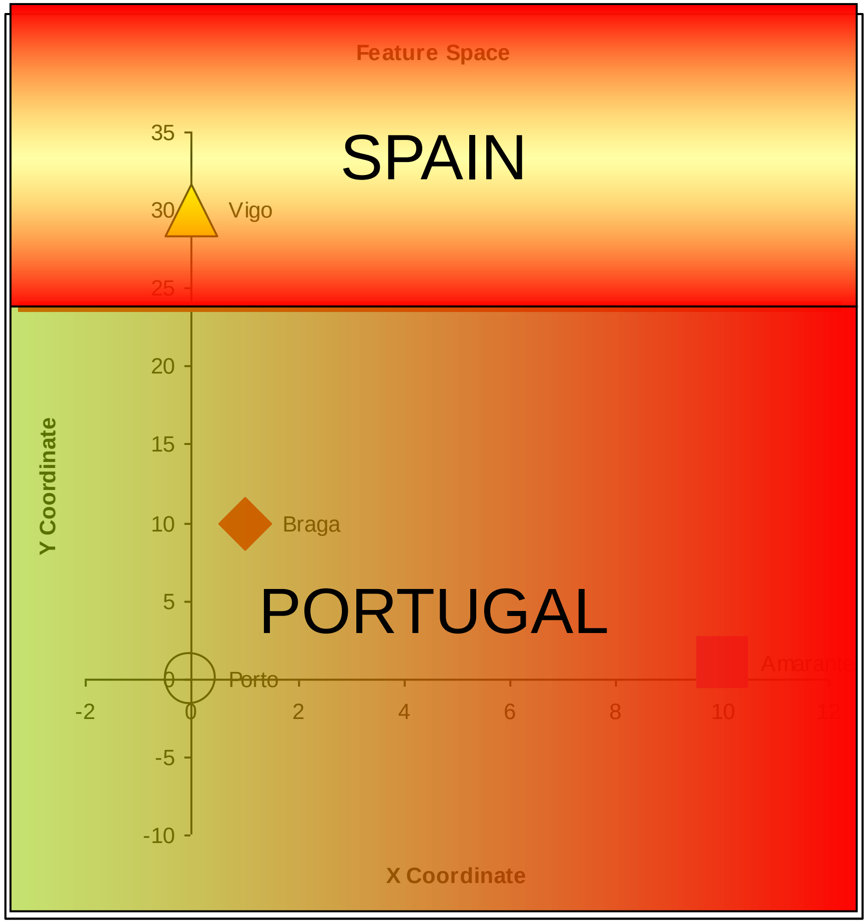
What if I don't have a model?

- I need to **learn** from observations.
  - Derive a model
  - Direct classification
- Training stage
  - Learn model parameters
- Classification



Classes

- In our example, cities can belong to:
  - Portugal
  - Spain
- I have two ***classes*** of cities
- A ***class*** represents a sub-space of my feature space



U. PORTO FC

Classifiers

- A **Classifier C** maps a class into the feature space

$$C_{\mathrm{Spain}}(x, y) = \begin{cases} true & , y > K \\ false & , otherwise \end{cases}$$

- Various types of classifiers
  - Nearest-Neighbours
  - Bayesian
  - Soft-computing machines
  - Etc...

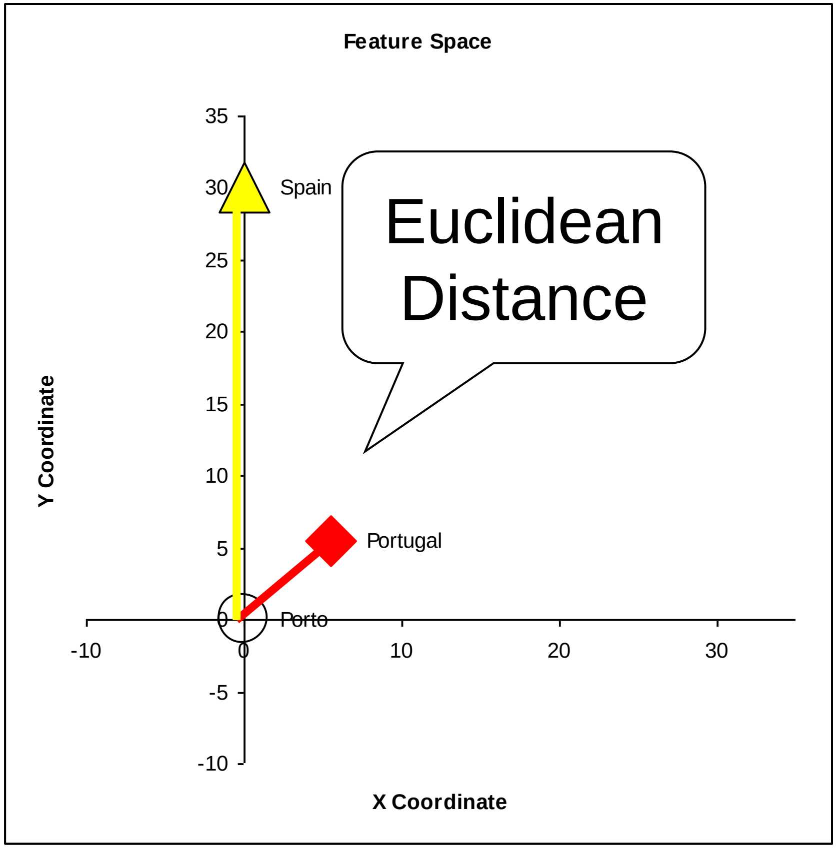
U.PORTO FC

Distance to Mean

- I can represent a class by its mean feature vector

$$C = \overline{F}$$

- To classify a new object, I choose the class with the closest mean feature vector
- Different distance measures!



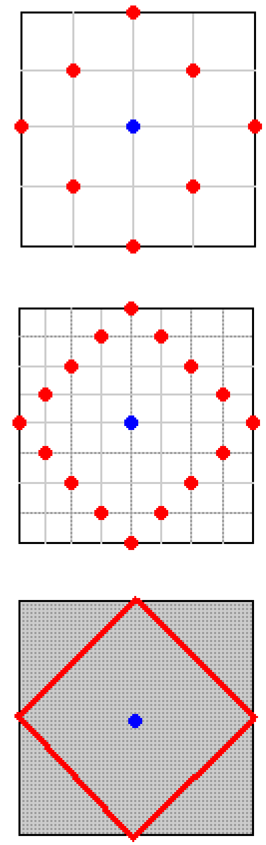
Possible Distance Measures

- L1 Distance

$$L1 = \frac{1}{N}\sum_{x=1}^{N}\left|S(x) - v(x)\right|$$

- Euclidean Distance (L2 Distance)

$$L2 = \frac{1}{N}\sum_{x=1}^{N}\left(S(x) - v(x)\right)^2$$

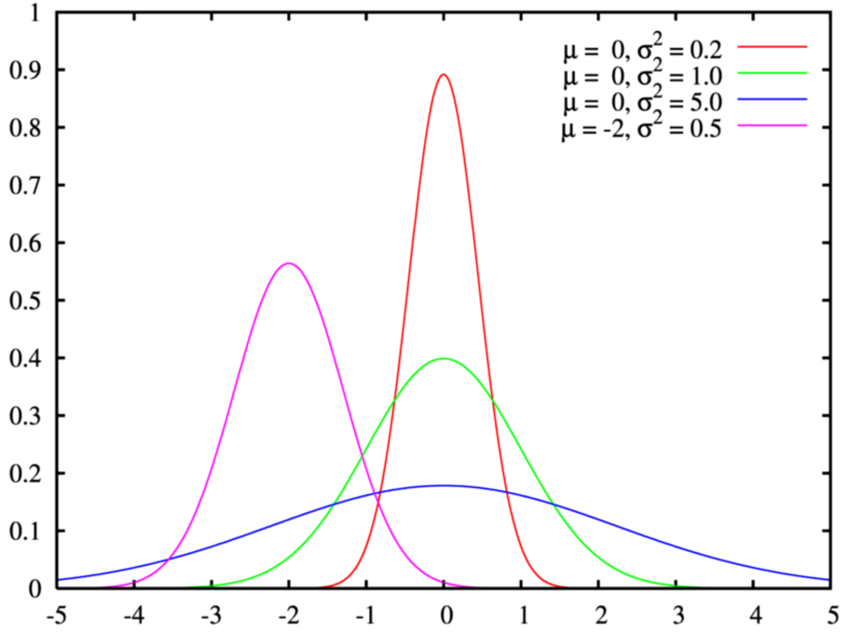


L1 or Taxicab Distance

Gaussian Distribution

- Defined by two parameters:
  - Mean: μ
  - Variance: σ^2
- Great approximation to the distribution of many phenomena.
  - *Central Limit Theorem*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right)$$



U. PORTO FC

Multivariate Distribution

- For N dimensions:

$$f_X(x_1, \ldots, x_N) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1} (x-\mu)\right)$$

- Mean feature vector:

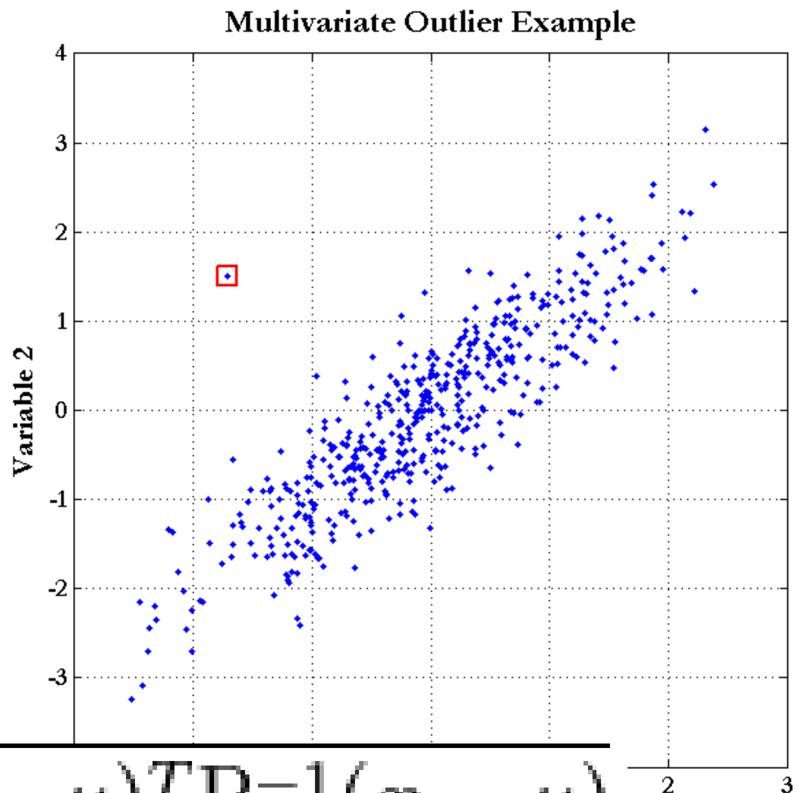
$$\mu = \overline{F}$$

- Covariance Matrix:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \qquad \mu_i = \mathrm{E}(X_i) \qquad \Sigma_{ij} = \mathrm{E}\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

Mahalanobis Distance

- Based on the covariance of coefficients
- Superior to the Euclidean distance

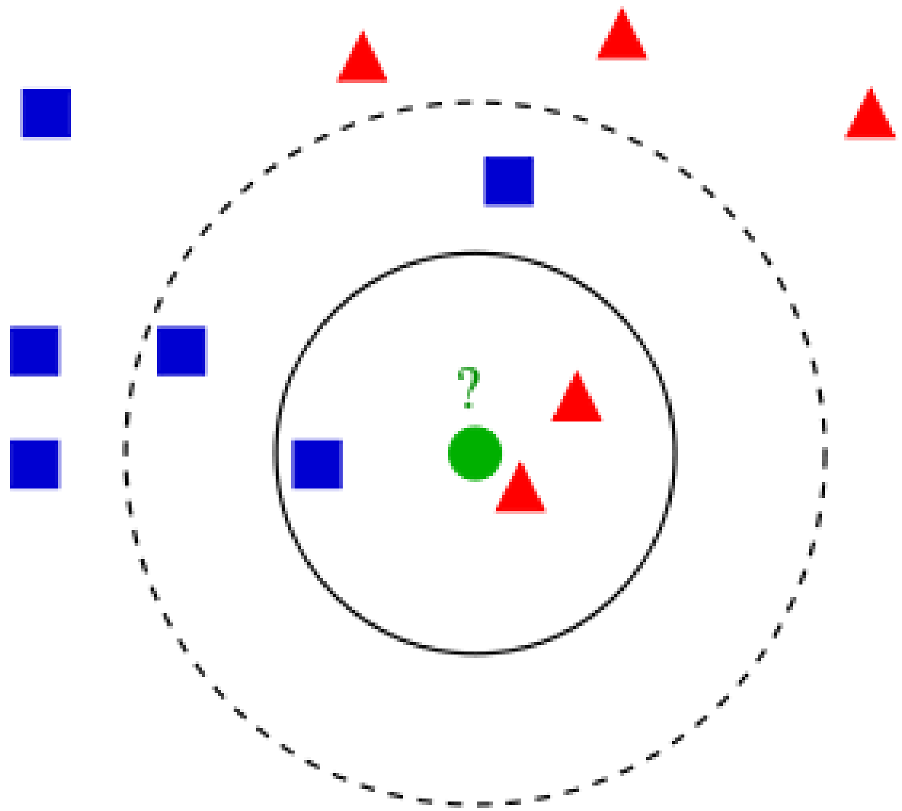


$$D_M(x) = \sqrt{(x-\mu)^T \mathrm{P}^{-1} (x-\mu)}.$$

U.PORTO FC

K-Nearest Neighbours

- Algorithm
  - Choose the closest K neighbours to a new observation
  - Classify the new object based on the **class** of these K objects
- Characteristics
  - Assumes no model
  - Does not scale very well...

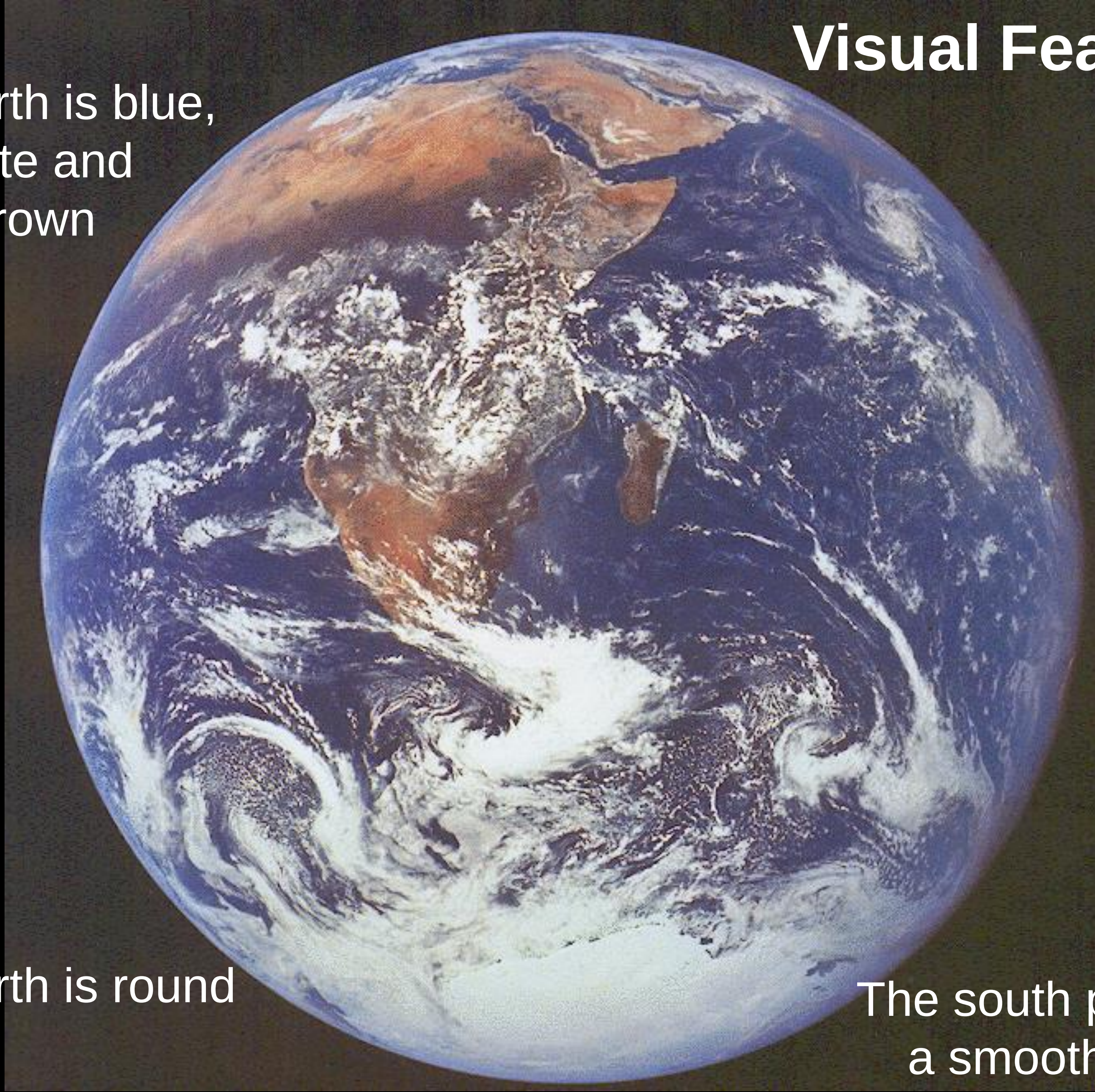


Topic: Visual Features

- Introduction to Pattern Recognition
- Statistical Pattern Recognition
- Visual Features

Visual Features

The earth is blue, white and brown



The earth is round

The south pole has a smooth texture

Visual Features

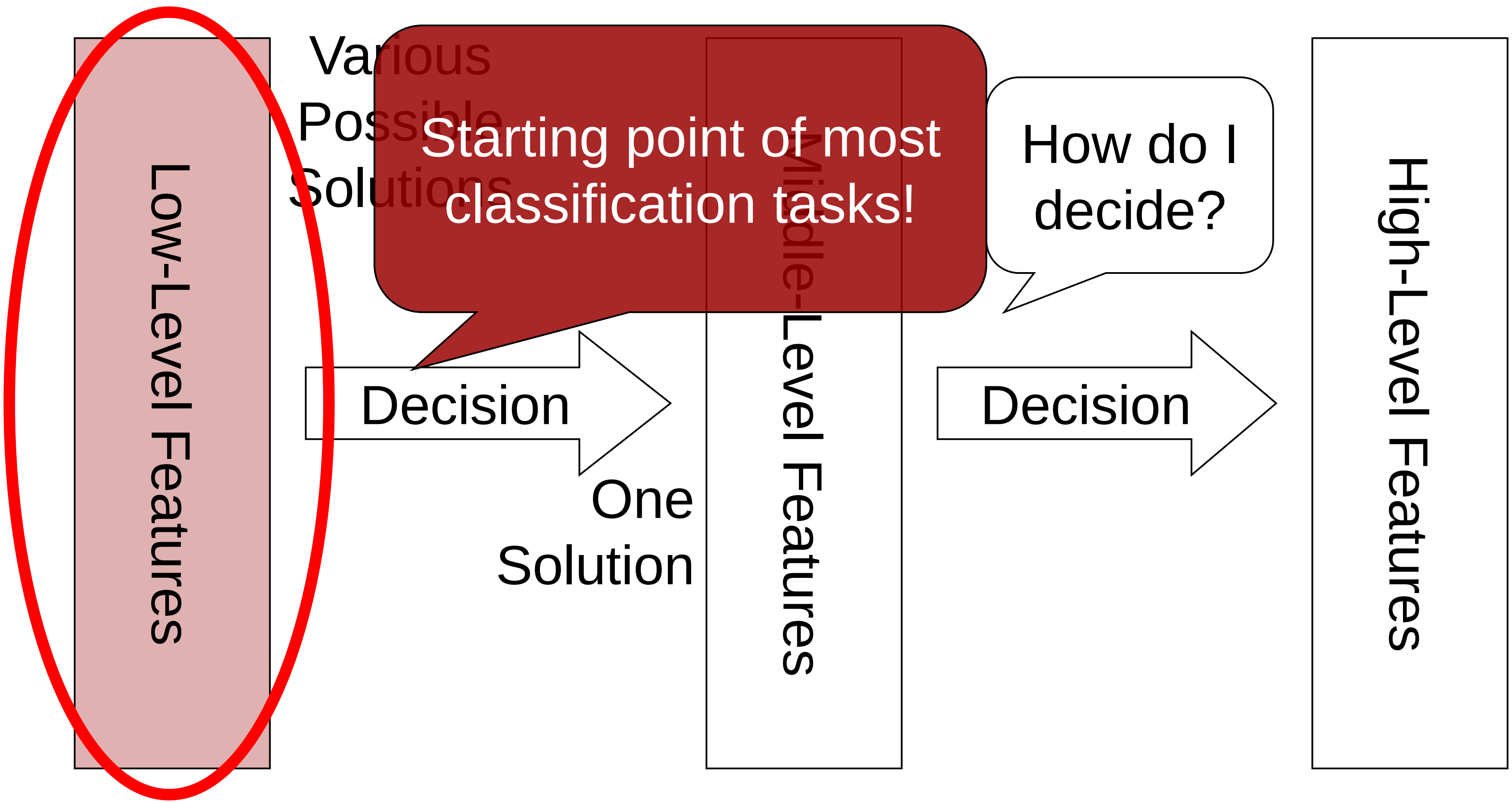
- Features
  - Measure specific characteristics
  - Numerical values
  - May have multiple values
- Visual Features
  - Quantify visual characteristics of an image
  - Popular features
    - Colour, Texture, Shape

Feature vector

- Feature $\boldsymbol{F_i}$ $$F_i = [f_i]$$
- Feature $\boldsymbol{F_i}$ with $\boldsymbol{N}$ values. $$F_i = [f_{i1}, f_{i2}, \ldots, f_{iN}]$$
- Feature vector F with M features. $$F = [F_1 \mid F_2 \mid \ldots \mid F_M]$$

- Naming conventions for this module:
  - Elements of a **feature vector** are called **coefficients**
  - **Features** may have one or more **coefficients**
  - **Feature vectors** may have one or more **features**

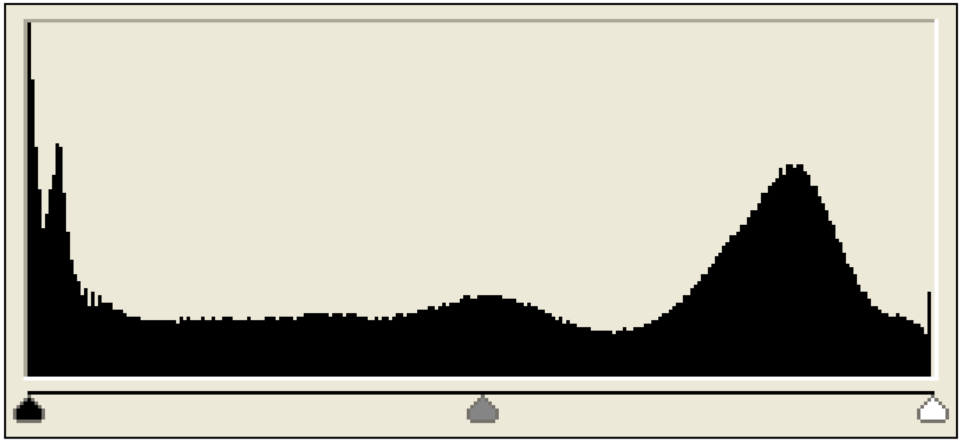
Features & Decisions



Gray-Level Histogram

- Intensity distribution (HSI)
- We can define the number of histogram bins
- Histogram bins = Feature coefficients

$$F = [f_0, \ldots, f_{255}]$$



Colour Histogram

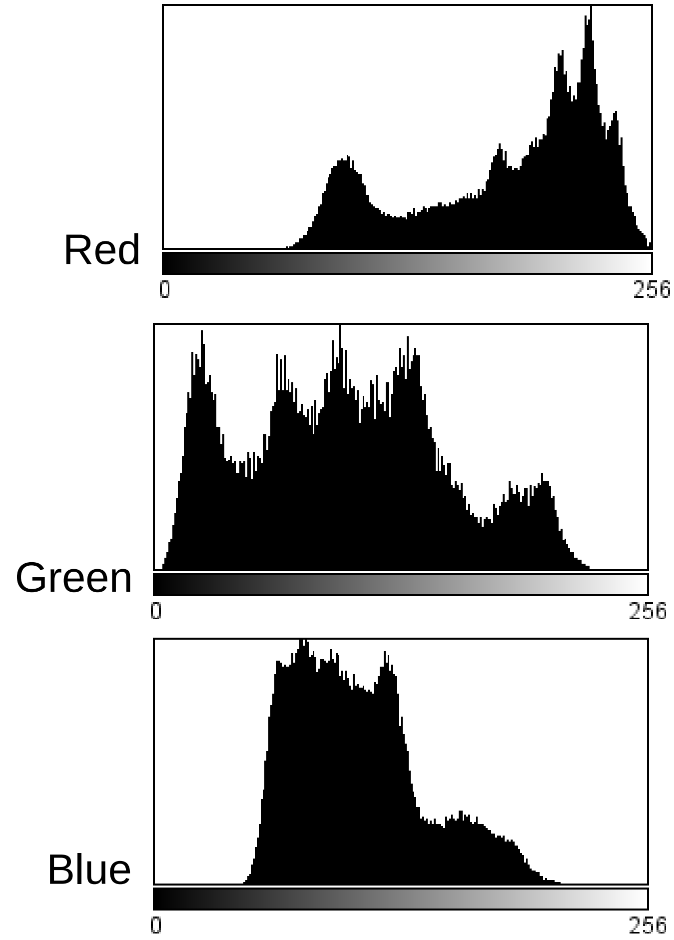
- We typically have three histograms

  Ex: RGB Colour space

  - Red Histogram

  - Green Histogram

  - Blue Histogram
- How do we build a feature vector?
  – Concatenate vectors
  – Multi-dimensional quantization of colour space



U. PORTO FC

RGB Histogram

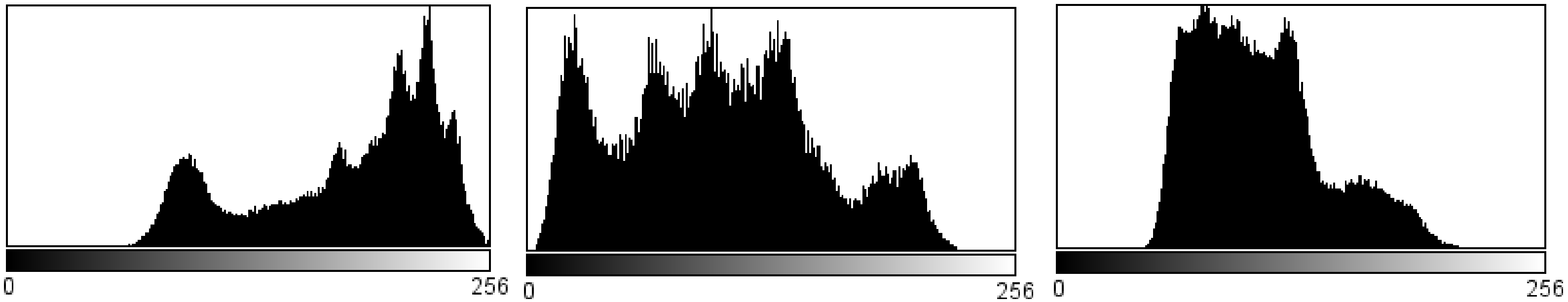
- Simply concatenate vectors
- Not very smart... (why?)

$$F_R = [f_{R0}, \dots, f_{R255}]$$

$$F_G = [f_{G0}, \dots, f_{G255}]$$

$$F_B = [f_{B0}, \dots, f_{B255}]$$

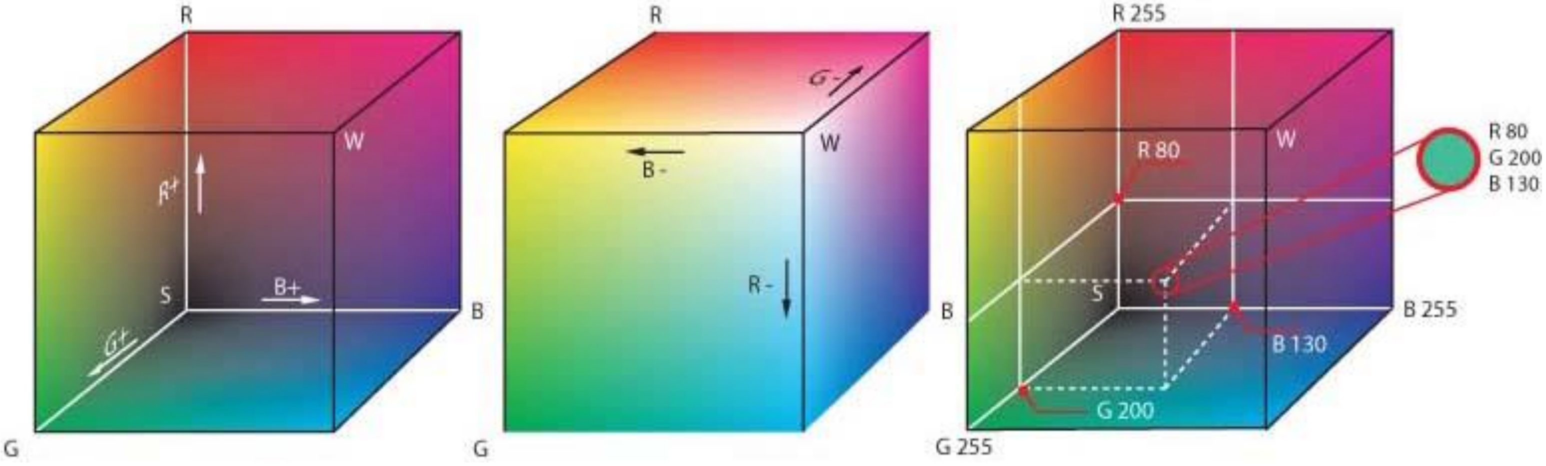
$$F_{RGB} = \left[F_R \mid F_G \mid F_B\right]$$



Combined Histogram

- Quantize multi-dimensional colour space
- RGB
  – Each coefficient is a small 'cube' inside the RGB cube

$$F = [f_0, \ldots, f_N]$$

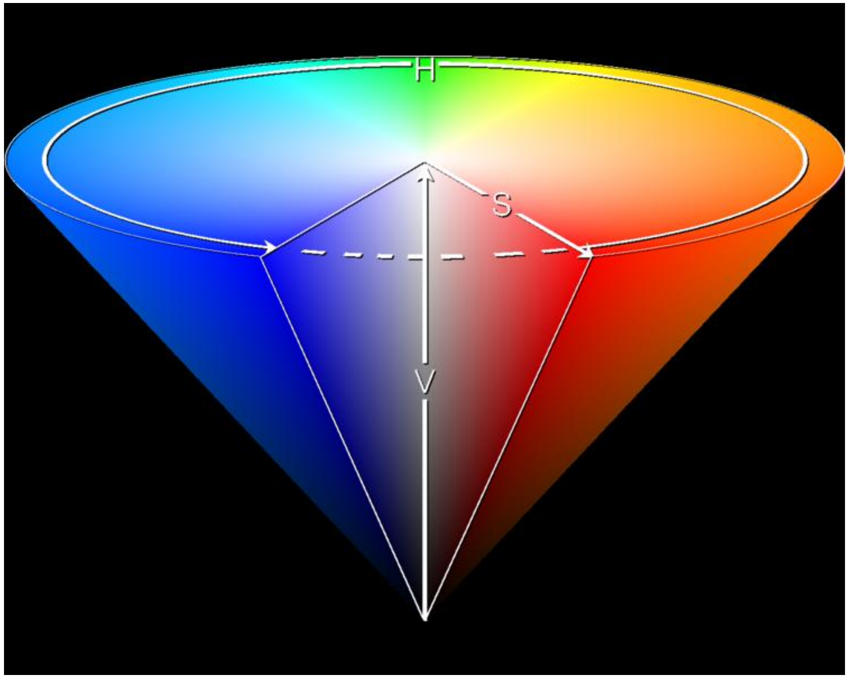


HSI Histogram

- Quantize HSI space
  - Define number of bins *N*.
  - Feature vector

$$F_{HSI} = [f_0, \ldots, f_N]$$

- Typically better for object description



Example: MPEG-7 Scalable Colour

- HSI Histogram
- Typical quantization: 256 bins.
  - 16 levels in H
  - 4 levels in S
  - 4 levels in I

$$F_{SC} = [f_0, \ldots, f_{255}]$$

- Very popular for CBIR (Content-Based Image Retrieval).

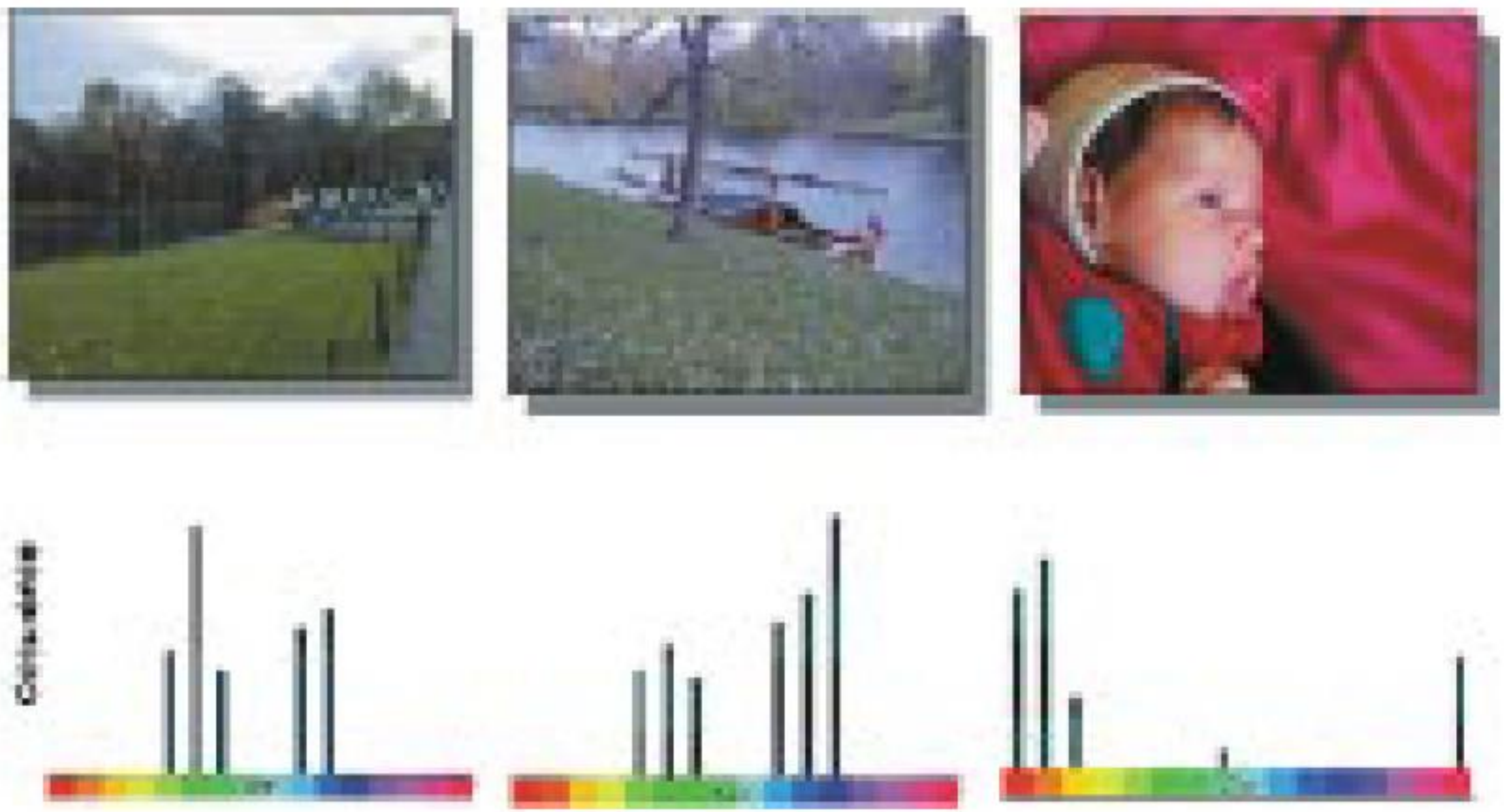


Fig. 2. Three color images and their MPEG-7 histogram color distribution, depicted using a simplified color histogram. Based on the color distribution, the two left images would be recognized as more similar compared to the one on the right.

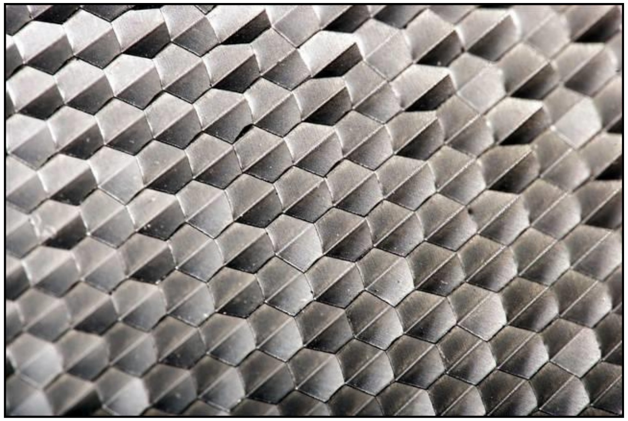
[Sikora 2001]

What is texture?

"Texture gives us information about the spatial arrangement of the colours or intensities in an image"

[L. Shapiro]





Two approaches to texture

- Structural approach
  - Texture is a set of primitive *texels* in some regular or repeated relationship
  - Good for regular, 'man-made' textures
- Statistical approach
  - Texture is a quantitative measure of the arrangement of intensities in a region
  - More general and easier to compute

Statistical approaches

- Grey level of central pixels
- Average of grey levels in window
- Median
- Standard deviation of grey levels
- Difference of maximum and minimum grey levels
- Difference between average grey level in small and large windows
- Sobel feature
- Kirsch feature
- Derivative in x window
- Derivative in y window
- Diagonal derivatives
- Combine features

How do I pick one??

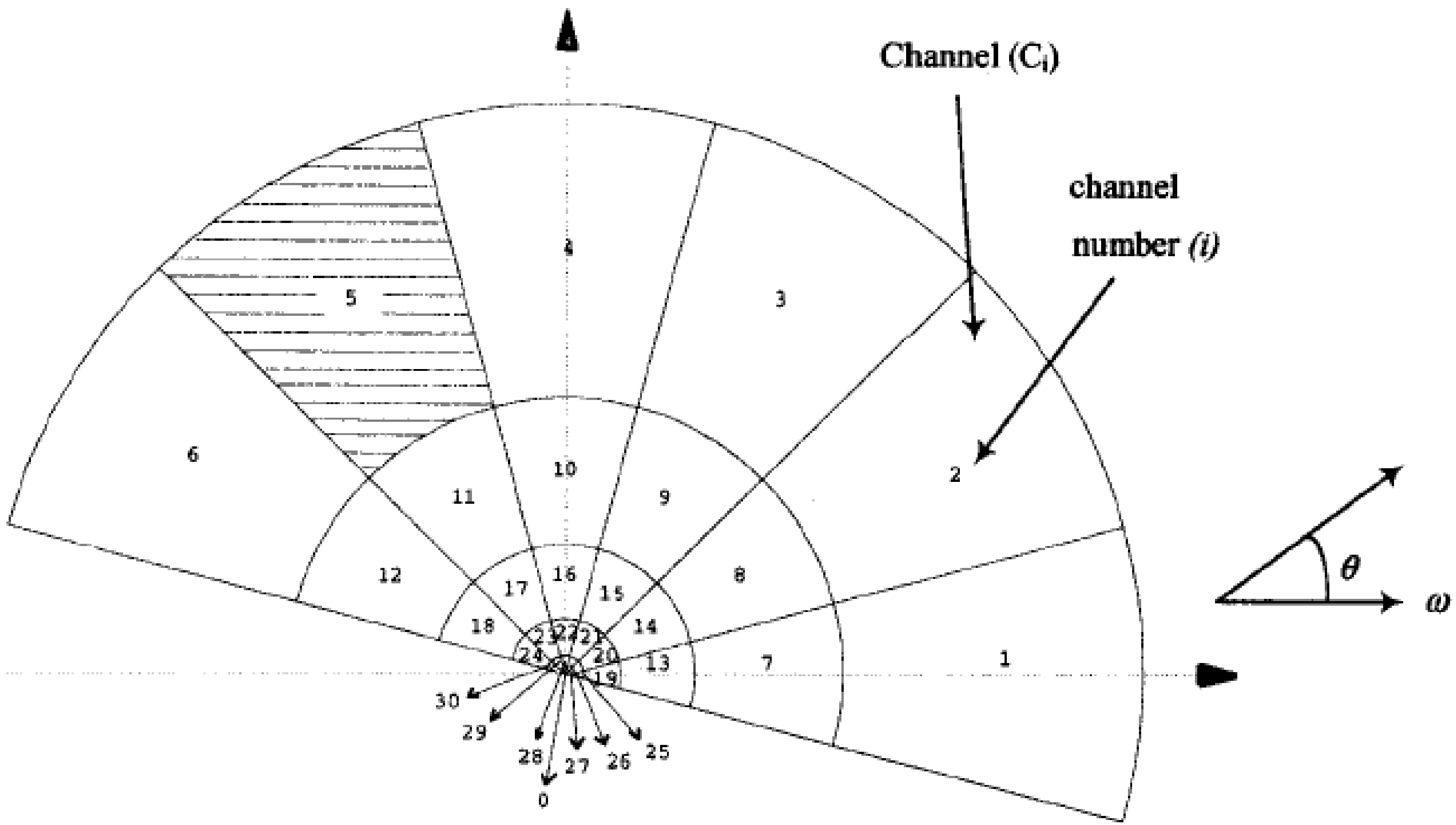
Gabor Filter Banks

- Filters the image with a set of orientation and scale sensitive filters
- Computes mean and standard deviation of response
- Example: 30 channels
  - 6 in angular direction, 5 in radial direction

$$F_{GFB} = [f_{DC}, f_{SC}, e_1, e_2, \ldots, e_{30}, d_1, d_2, \ldots, d_{30}]$$

fDC, *fSC* are the mean intensity and the standard deviation of image texture), where
ex and *dx* are the logarithmically scaled texture energy and texture energy deviation coefficients.

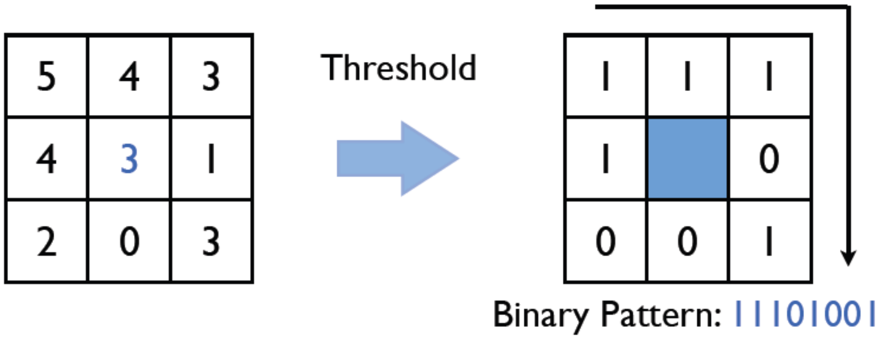
GFB Channels

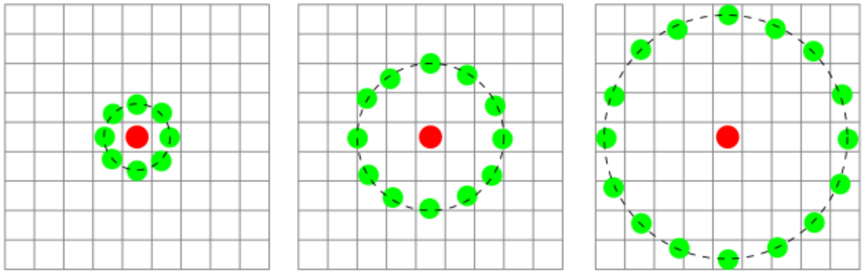


U.PORTO FC

Local Binary Patterns

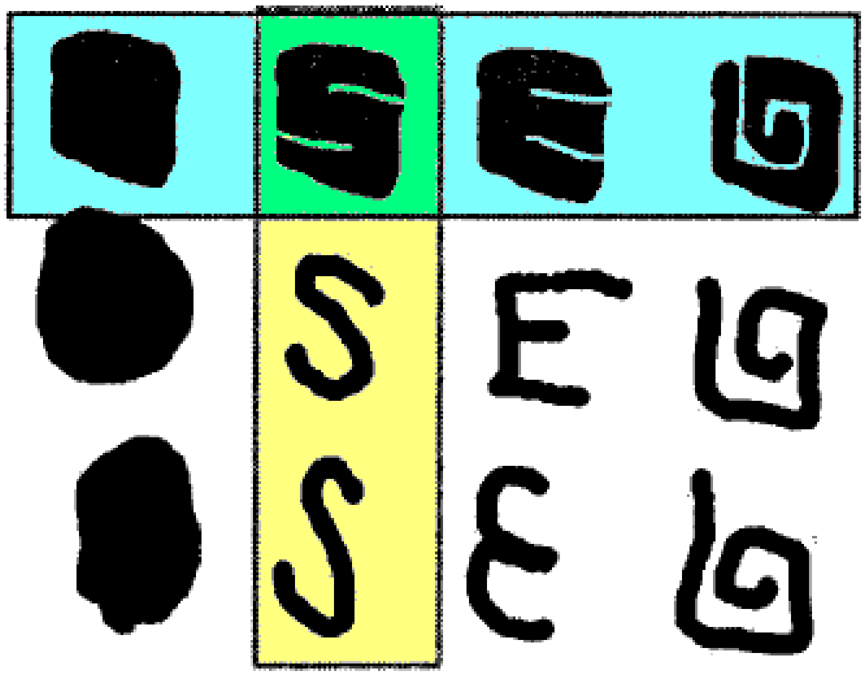
- Idea: Compare the intensity value of a pixel with its neighbors
  - 1 if neighbor is larger
  - 0 if not
- Combine results to generate a unique binary code
- Create a histogram of occurences of each binary pattern





U.PORTO FC

Shape Descriptors



- Blue: Similar shapes by Region-Based
- Yellow: Similar shapes by Contour-Based

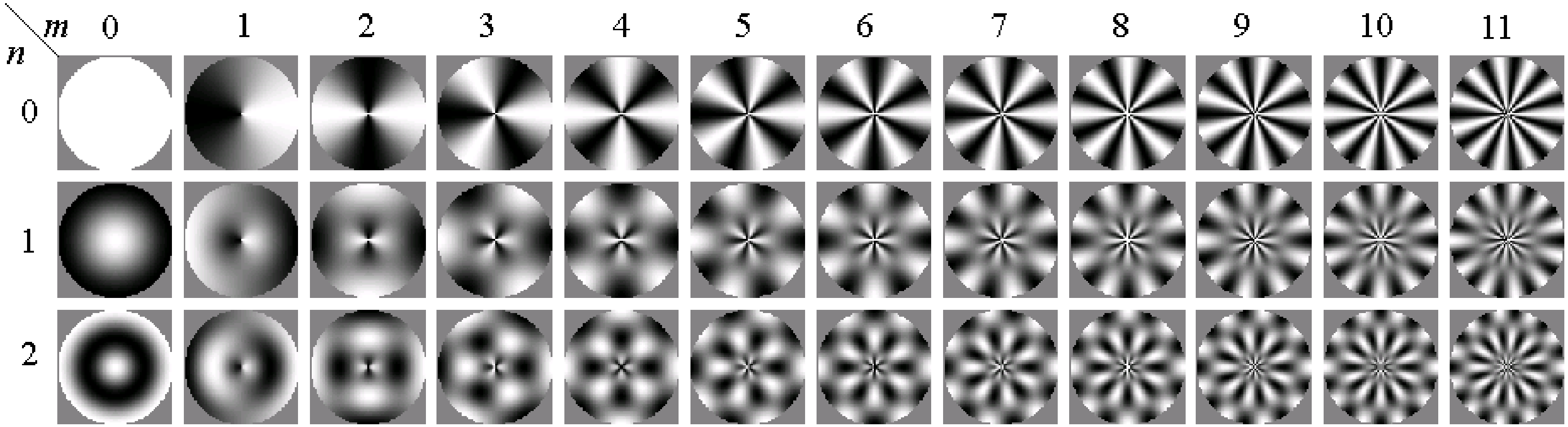
Example: Region-Based Shape Descriptor

- Use a set of separable ART (angular radial transformation) functions
- Classify shape along various angular and radial directions
- Totals 35 coefficients

$$F_{RBS} = [\,f_0,\ldots,f_{34}\,]$$

U. PORTO FC

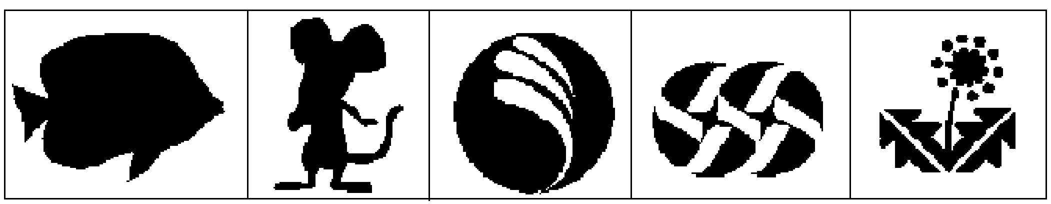
ART Basis Functions



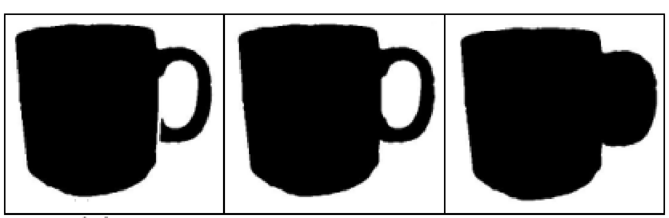
•Applicable to figures (a) – (e)

•Distinguishes (i) from (g) and (h)

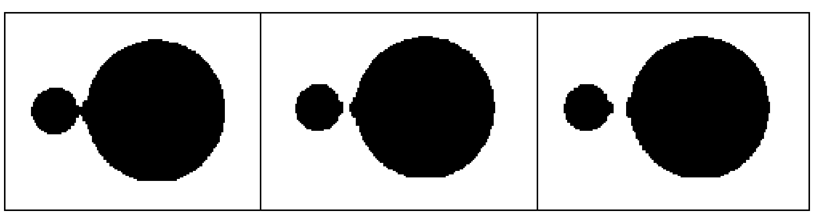
•(j), (k), and (l) are similar



(a) (b) (c) (d) (e)



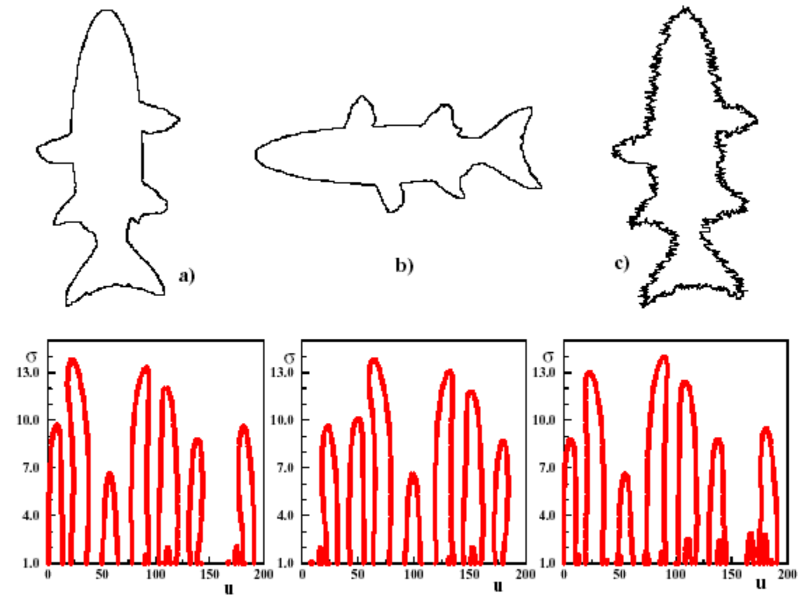
(g) (h) (i)



(j) (k) (l)

Example: Contour-Based Shape Descriptor

- Finds curvature zero crossing points of the shape's contour (key points)
- Reduces the number of key points step by step, by applying Gaussian smoothing
- The position of key points are expressed relative to the length of the contour curve



•Applicable to (a)

•Distinguishes differences in (b)

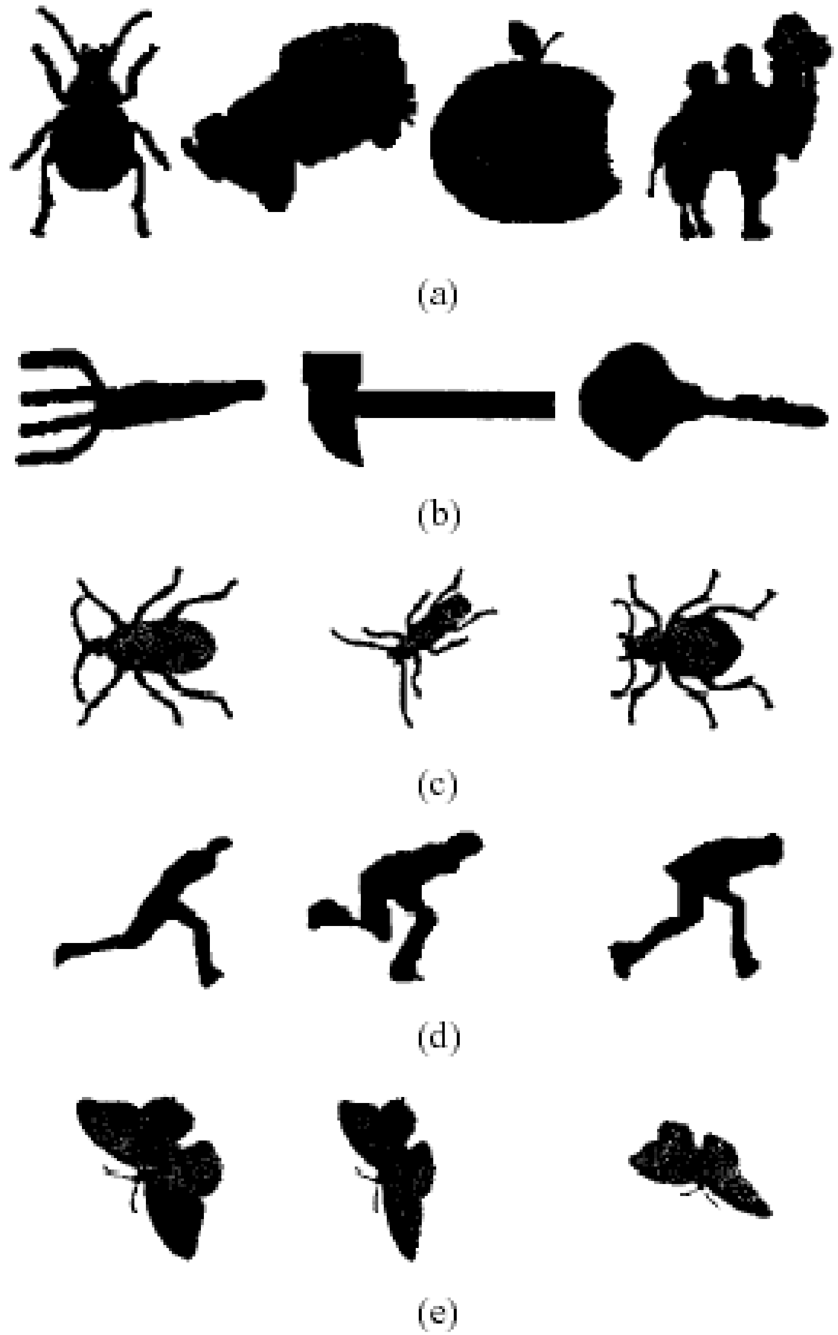
•Find similarities in (c) - (e)

Advantages:

• Captures the shape very well

• Robust to the noise, scale, and orientation

• It is fast and compact



Resources

- Szeliski, “Computer Vision: Algorithms and Applications”, Springer, 2011
  - Chapter 14 – “Recognition”