CHAPTER 1: INTRODUCTION TO PATTERN RECOGNITION

Grado en Ingeniería Informática

Curso 2014 / 15

Topics

1. Introduction
2. System examples
3. Characteristics
4. Problem types

Warning: similar nomenclature

- Respect to system:
  - ✓ Program design
  - ✓ Modeling techniques
  - ✓ Pattern Recognition
  - ✓ Simulation
  - ✓ Automatic learning

- Respect to system design:
  - ✓ Parameter estimation
  - ✓ Training

Introduction

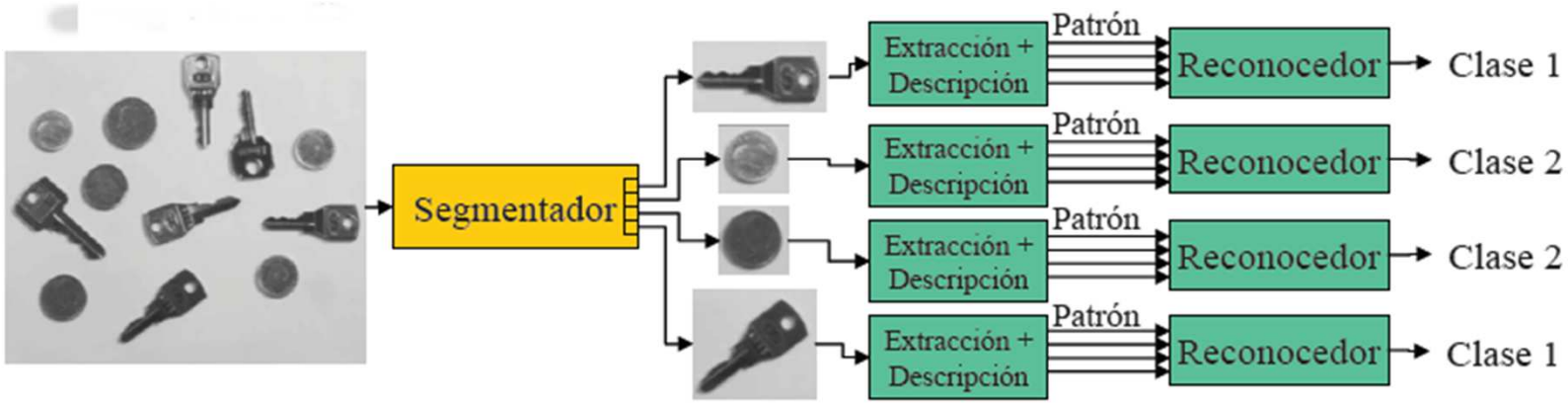
- Definitions:
  - **Object:** Representation of elements under study
  - **Pattern:** Set of rules which a object is described
  - **Characteristic:** Object property
  - **Class:** Characteristics objects set which distinguish it from others
  - **Recognition:** Labelling process to determine the class to which the object belongs

Pattern Recognition: Branch of knowledge of multidisciplinary approach whose subject of study are the identification, characterization, classification and reconstruction processes of sets of objects, as well as to develop theories, technologies and methodologies relating to these processes.

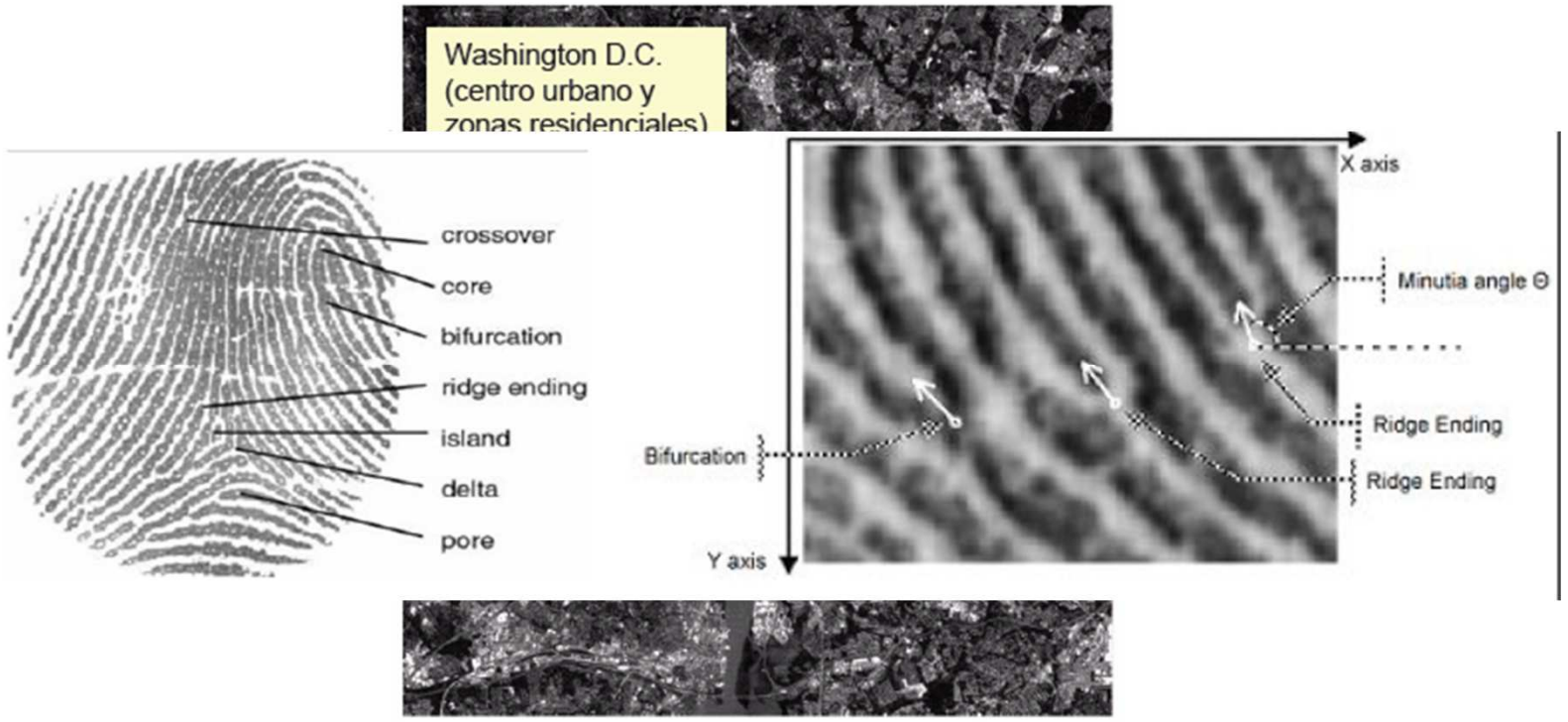
Introduction

- **Pattern Recognition:** basic feature in animal behavior
- What is a pattern? It is an object description
- Pattern types:
  - ✓ Specifics:
    - o Spatial: characters, images, etc.
    - o Temporary: waveforms (voice), series, etc.
  - ✓ Abstract: reasoning, solutions to problems, etc.
- **GOAL:** space – time pattern recognition

System examples



System examples

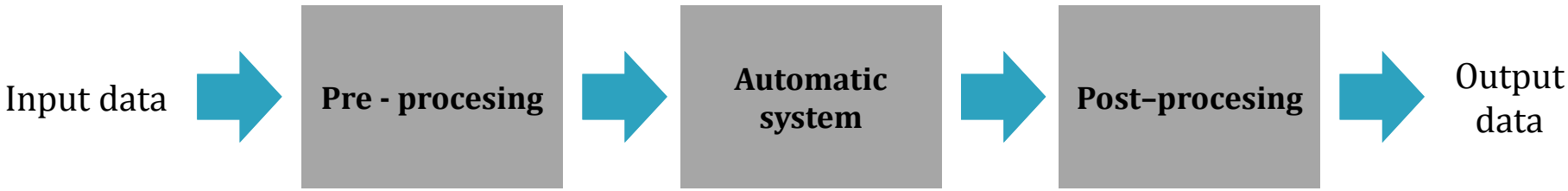


System examples

TASK	INPUT	OUTPUT
Character Recognition (OCR)	Written text	Recognize text
Voice processing	Speech signal	Sentence, speaker, etc.
Spectrometry	Spectral signal	Composition of matter
Weather	Meteosat images	Weather forecasting
Medical diagnosis	Symptoms	Disease

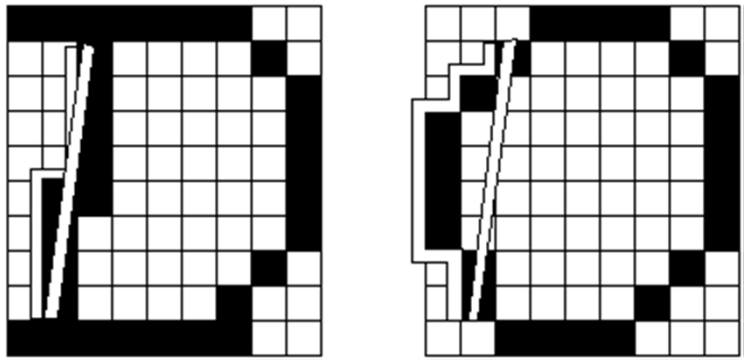
Automatic system

- It is required:
  - ✓ Similar characteristics between patterns which describes objects of a same class
  - ✓ Clearly differentiated characteristics between patterns which describes objects of different classes



Characteristics

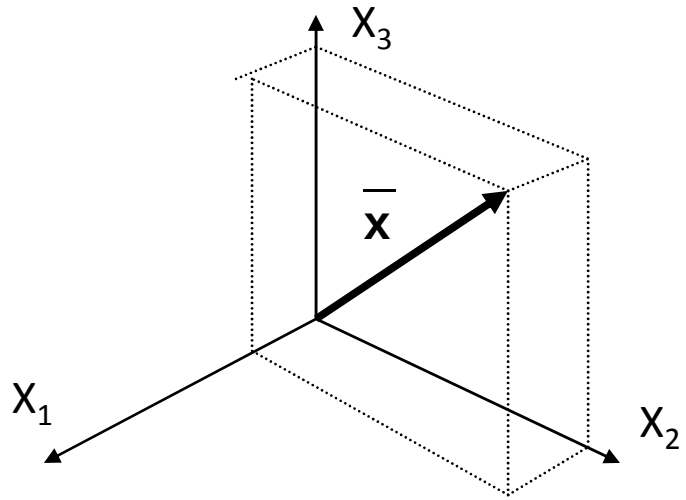
- **Characteristic:** important feature which allows to distinguish fully or partially an object (pattern, class, etc.) from another
  - ✓ **Simple:** based on pixels, width, height, etc.
  - ✓ **Complex:** perimeter, area, compactness, inclination...



Characteristics

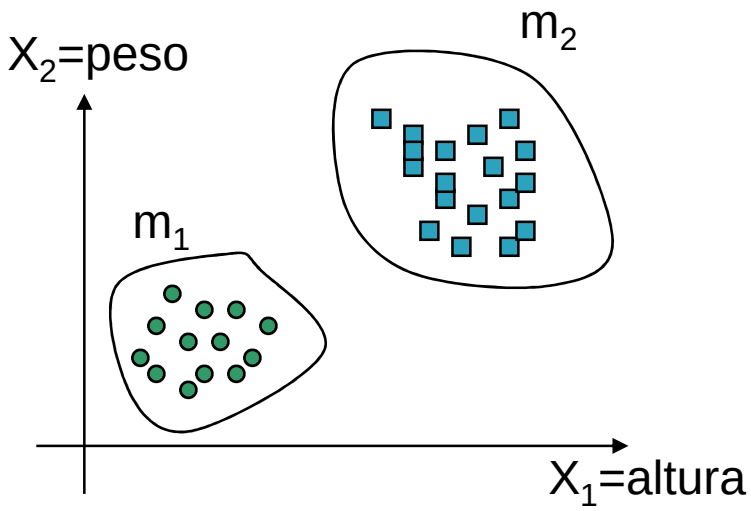
- **Vector of characteristics:** Composition of several characteristics in a vector
- A vector of characteristics defines a point in a n – dimensional space

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



Characteristics

- **Space of characteristics:** The set of patterns which belongs to the same class is grouped into any region of space
- Example:



$\mathbf{m_1}$ = children aged up to 10
$\mathbf{m_2}$ = basketball players

In this case, the separation is perfect, but this is not always the case, the classes are usually overlap

PROBLEM TYPES

Classification

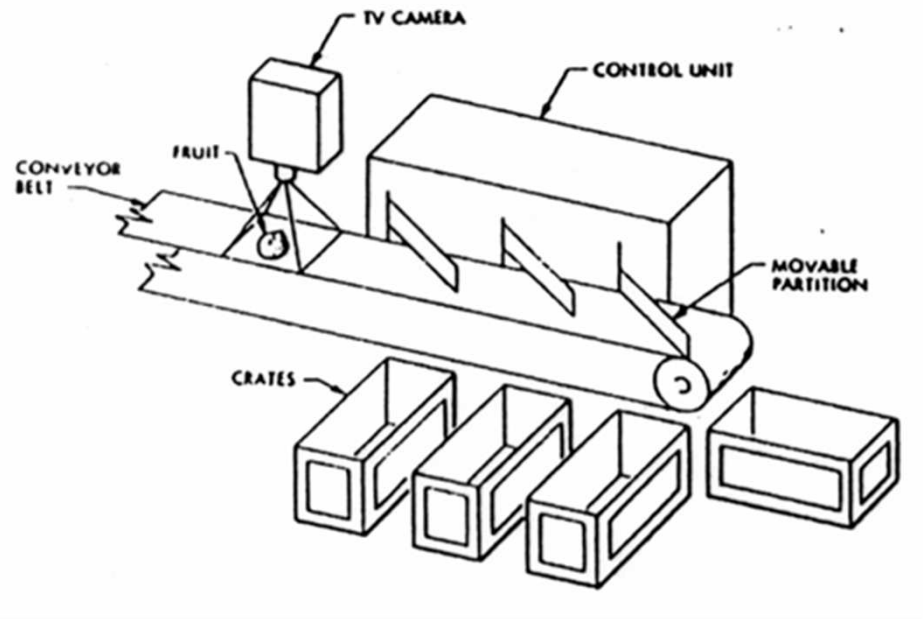
- It consists of allocating the input pattern to only one among N categories

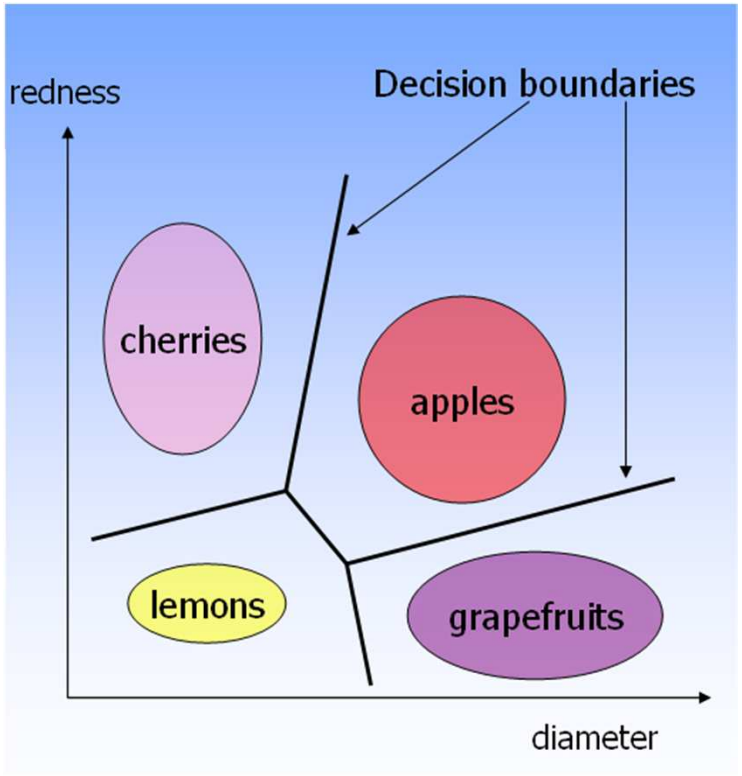
$$y = f : R^n \Rightarrow \{Y_0, Y_1, \ldots Y_N\}$$

- Examples:
  - ✓ Speech recognition
  - ✓OCR
  - ✓ Expert system

Classification

- Example: Fruits recognition

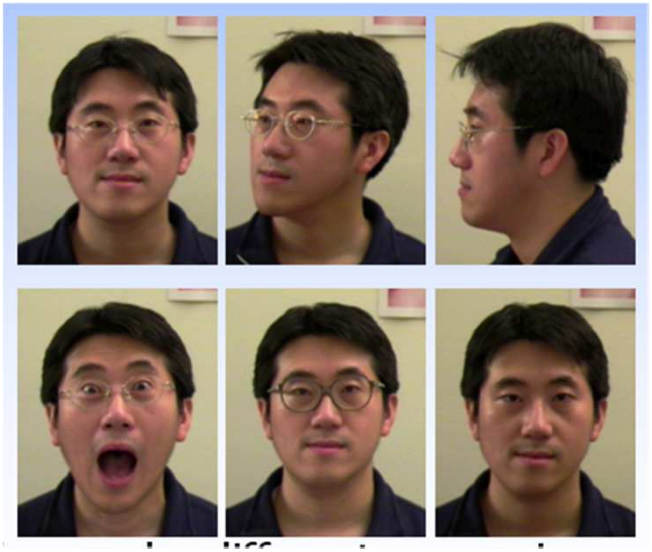




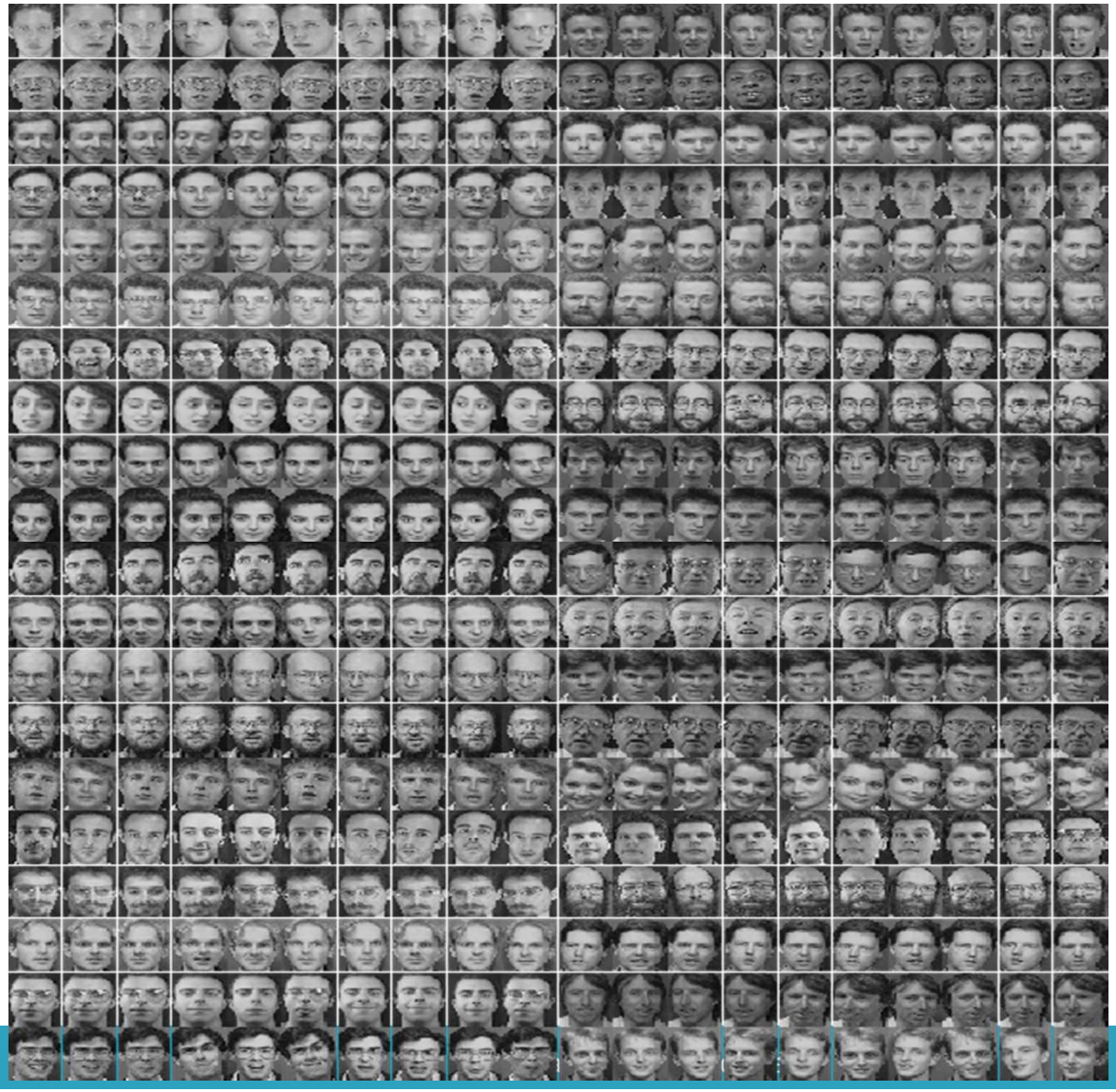
Classification

- Example: Character recognition

- Example: Face recognition



Classification



Classification

- Example: Dogs/cats recognition

Regression

- It assign the input pattern the corresponding output to a continuous variable

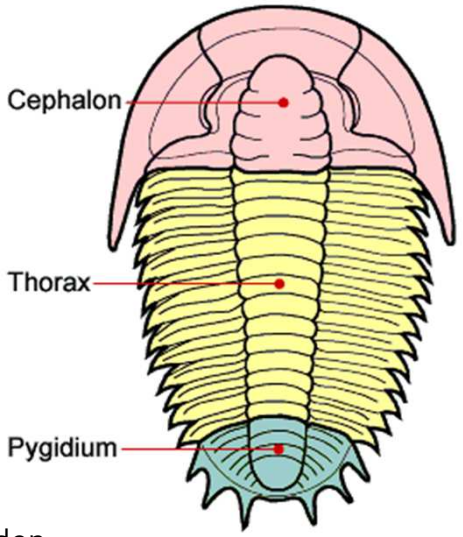
$$y = f : R^n \Rightarrow R$$

- General case:

$$y = f(x) \qquad x \in R^n, y \in R^m$$

Regression

- Example: Size estimation of Trilobites
  - ✓ Under most preservation conditions, it is difficult to find complete copies of Trilobites
  - ✓ The head (cephalon) is more common
  - ✓ Therefore, it is useful for being able to estimate the body from measure of the head, to establish whom prove the better determination of total size.

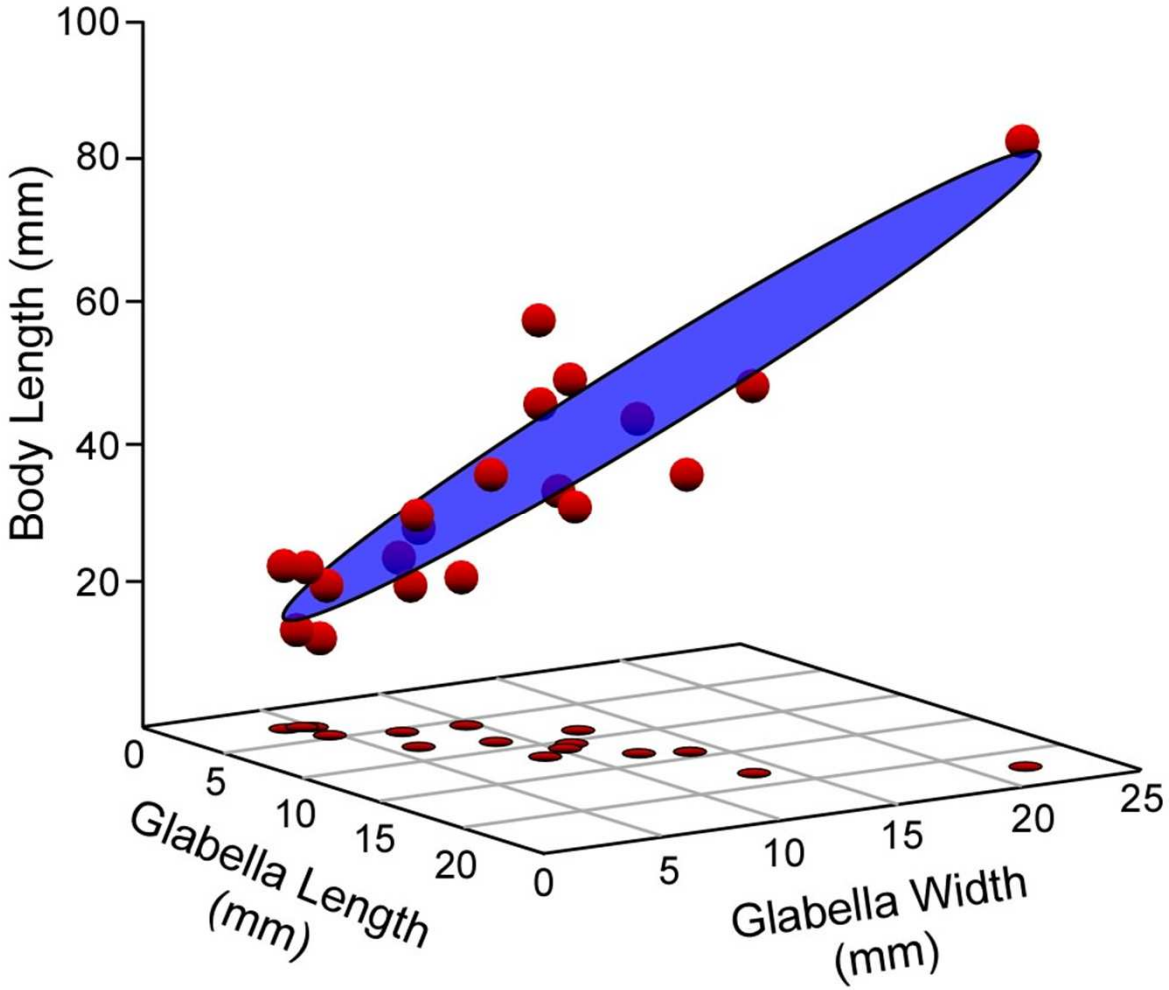


Norman MacLeod. Keeper of Palaeontology, The Natural History Museum, London

Table 1. Trilobite Data[1]

Genus	Body Length (mm)	Glabellar Length (mm)	Glabellar Width (mm)
Acaste	23.14	3.50	3.77
Balizoma	14.32	3.97	4.08
Calymene	51.69	10.91	10.72
Ceraurus	21.15	4.90	4.69
Cheirurus	31.74	9.33	12.11
Cybantyx	36.81	11.35	10.10
Cybeloides	25.13	6.39	6.81
Dalmanites	32.93	8.46	6.08
Delphion	21.81	6.92	9.01
Ormathops	13.88	5.03	4.34
Phacopdina	21.43	7.03	6.79
Phacops	27.23	5.30	8.19
Placopoaria	38.15	9.40	8.71
Pricyclopyge	40.11	14.98	12.98
Ptychoparia	62.17	12.25	8.71
Rhenops	55.94	19.00	13.10
Sphaerexochus	23.31	3.84	4.60
Toxochasmops	46.12	8.15	11.42
Trimerus	89.43	23.18	21.52
Zacanthoides	47.89	13.56	11.78
Mean	36.22	9.37	8.98
Std. Deviation	18.63	5.23	4.27

Regression


Body Length (mm)
100
80
60
40
20
0
5
10
15
20
Glabella Length (mm)
0
5
10
15
20
25
Glabella Width (mm)

Prediction

- It determine the pattern in the time *t + 1* from the pattern in the times *t, t – 1, t – 2, …*

$$y(t) = f(y(t-1), y(t-2), \ldots y(t-n))$$

- Examples:
  - ✓ Weather prediction
  - ✓ To invest in market
  - ✓ Forecast in companies

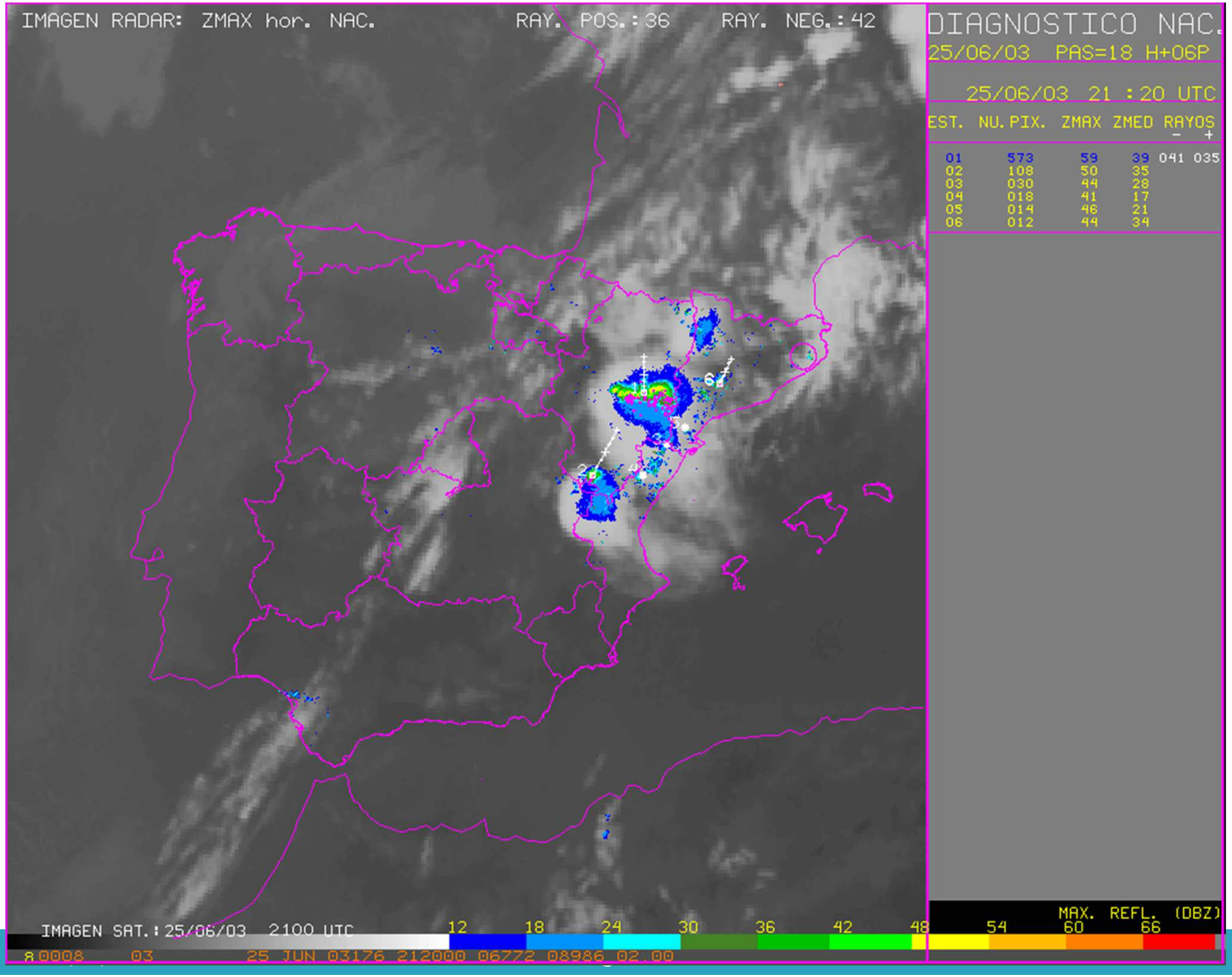


IMAGEN RADAR: ZMAX hor. NAC.
RAY. POS.: 36
RAY. NEG.: 42
DIAGNOSTICO NAC.
25/06/03 PAS=18 H+06P
25/06/03 21 :20 UTC
EST. NU.PIX. ZMAX ZMED RAYOS
- +
01 573 59 39 041 035
02 108 50 35
03 030 44 28
04 018 41 17
05 014 46 21
06 012 44 34
IMAGEN SAT.: 25/06/03 2100 UTC
12 18 24 30 36 42 48
MAX. REFL. (DBZ)
54 60 66
8 0008 03 25 JUN 03176 212000 06772 08986 02.00

Other issues

- Density estimation
- Clustering
- Adjusting functions
- Interpolation
- ...

Other issues

- **Density estimation:** It estimates the probability density function from which input data proceed

- It does not make any sense to talk about *y(t)* or *y(x)*

- Examples:
  - ✓ Classification
  - ✓ Statistics methods