8001652 Introduction to Pattern Recognition: Lecture 1

Jussi Tohka

jussi.tohka@tut.fi

Institute of Signal Processing

Tampere University of Technology

Lectures and exercises

- Lecturer: Jussi Tohka
  - e-mail: jussi.tohka@tut.fi
  - Office TE309
- Assistants: Heidi Koivistoinen and Margarita Magadan-Mendez.
- Lectures Mondays 14 - 16 at TB222.
- Exercises: Thurdays 12 - 13 TB216 (English) Wednesdays 9 - 10 TC133 (Finnish)
- First exam: Monday 15th December
- Homepage: http://www.cs.tut.fi/kurssit/8001652/

Generalities

- Lectures 28h (2h / week) Exercises 14h (1h /week)
- REQUIREMENTS: Final examination and active participation in exercises.
- LITERATURE: Duda, Hart, Stork: Pattern Classification, 2nd edition, Wiley, 2001.
- PREREQUISITES: Introduction to signal processing 2, some basic skills in mathematics.
- This course is a required prerequisite for advanced courses in pattern and speech recognition.

Goal and contents

The goal is to introduce basic methods and principles of pattern recognition. Basics of multivariate probability and statistics. Bayesian decision theory. Parameter estimation from training data. Non-parametric techniques for pattern classification. Algorithms for unsupervised classification.

Exercise bonus

- Exercise bonuses: 10 % of exercises must be done in order to pass the course.
- Thereafter every 20 % earns an extra point in the exam. Thus

10 %	pass the course
30 %	1 point to the exam
50 %	2 points to the exam
70 %	3 points to the exam
90 %	4 points to the exam

Course outline

- This lecture
- Basics on probability and statistics
- Basics on probability and statistics continued
- Bayesian decision theory
- Bayesian decision theory
- Parameter estimation, maximum-likelihood estimator
- Curse of dimensionality, computational complexity, and recap.

Course outline

- Parzen windows, k-Nearest neighbours rule, CART
- Linear discriminant functions
- Linear discriminant functions
- Unsupervised classification
- Unsupervised classification
- No free lunch-theorem and recap

What is pattern recognition?

- Briefly and broadly speaking, pattern recognition is a task of finding some conceptual and relevant information from raw data.
- The definition of 'relevant information' depends on the application as does that of raw data. In summary, 'pattern recognition' can mean a lot of things and finding an universal definition for it is hard if not impossible.
- However, that is not to say that good opinions about the meaning of 'pattern recognition' could not be given.

What is pattern recognition?

www.cs.uvm.edu/~snapp/teaching/CS295PR/whatispr.html

- It is generally a easy for a person to differentiate the sound of a human voice, from that of a violin; a handwritten numeral "3," from an "8"; and the aroma of a rose, from that of an onion. However, it is difficult for a programmable computer to solve these kinds of perceptual problems. These problems are difficult because each pattern usually contains a large amount of information, and the recognition problems typically have an inconspicuous, high-dimensional, structure.

What is pattern recognition?

www.cs.uvm.edu/~snapp/teaching/CS295PR/whatispr.html

- Pattern recognition is the science of making inferences from perceptual data, using tools from statistics, probability, computational geometry, machine learning, signal processing, and algorithm design. Thus, it is of central importance to artificial intelligence and computer vision, and has far-reaching applications in engineering, science, medicine, and business. In particular, advances made during the last half century, now allow computers to interact more effectively with humans and the natural world (e.g., speech recognition software). However, the most important problems in pattern recognition are yet to be solved.

What is pattern recognition?

www.um.ac.ir/~patternrec/about.htm

- Pattern recognition is the scientific discipline whose goal is the classification of objects into a number of categories or classes. Depending on the application, these objects can be images or signal waveforms or any type of measurements that need to be classified. We will refer to these objects using the generic term patterns.
- Pattern recognition has a long history, but before the 1960s it was mostly the output of theoretical research in the area of statistics. As with everything else, the advent of computers increased the demand for practical applications of pattern recognition, which in turn set new demands for further theoretical developments.

What is pattern recognition?

www.um.ac.ir/~patternrec/about.htm

- As our society evolves from the industrial to its postindustrial phase, automation in industrial production and the need for information handling and retrieval are becoming increasingly important. This trend has pushed pattern recognition to the high edge of today?s engineering applications and research. Pattern recognition is an integral part in most machine intelligence systems built for decision making.

What is pattern recognition?

From prlab.ee.memphis.edu/frigui/ELEC7901/INTRO/Intro.html

- Pattern: a description of an object.
- Recognition: classifying an object to a pattern class.

PR is the science that concerns the description or classification (recognition) of measurements.
PR techniques are an important component of intelligent systems and are used for

- Decision making
- Object and pattern classification
- Data preprocessing

What is pattern recognition?

The course book says:
The ease with which we recognize a face, understand spoken words, read handwritten characters, identify our car keys in our pocket by feel, and decide whether an apple is ripe by its smell belies astoundingly complex processes that underlie these acts of pattern recognition. Pattern recognition - the act of taking in raw data and making an action based on the category of the pattern - has been crucial for our survival, and over the past tens of millions of years we have evolved highly sophisticated neural and cognitive systems for such tasks.

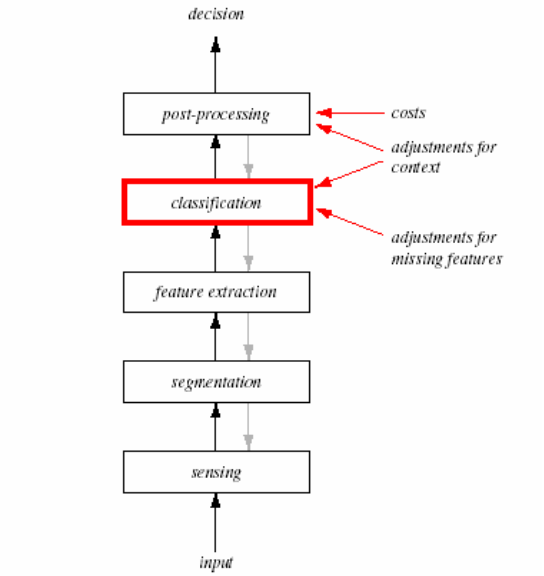
Machine perception

- The task: Seek to design and build machines that can recognize patterns.
- Applications: speech recognition, fingerprint identification, optical character recognition, DNA sequence identification.

Examples

- Differentiate between salmon and sea-bass
- Animal footprints

Pattern recognition systems



Feature extraction

- The distinction between feature extraction and classification is for practical reasons.
- The traditional goal of the feature extractor is to characterize an object by measurements. Measurement values should be similar for objects belonging to the same category and clearly distinct for objects in different categories.
- Invariant features: Features that remain same if something that is irrelevant is done to input.
- Feature extraction is very problem dependent: Good features for sorting fish are little use for recognizing fingerprints.

Classification

- The task of the classifier component is to use the feature vector provided by the feature extractor to assign the object to a category.
- Classification is the main topic of this course.
- The abstraction provided by the feature vector representation of the input data enables the development of a largely domain-independent theory of classification.

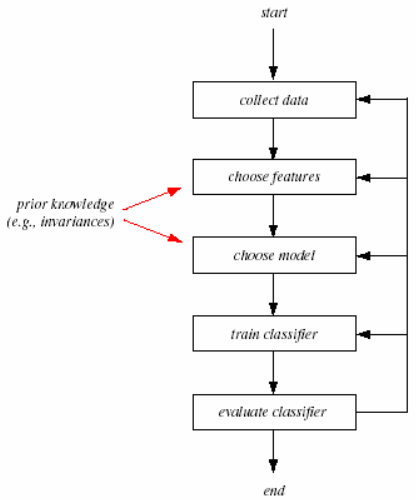
Classification

- The degree of difficulty of the classification problem depends on the variability in the feature values for objects in the same category relative to the feature value variation between the categories.
- Variability is natural or is due to noise.
- Variability can be described through statistics leading to statistical pattern recognition.
- Questions: How to design a classifier that can cope with the variability in feature values? What is the best possible performance?

Post processing

- The post-processor uses output of the classifier to decide on the recommended action.
- Simplest measure on a classifier performance is error rate. This leads to the minimum error rate classification. However, sometimes it is better to minimize the total expected cost of wrong classifications. This cost is called the risk.
- The post-processor could also exploit the context,
- or combine results of several classifiers.

The design cycle



Learning and adaptation

- In the broadest sense, any method that incorporates information from training samples in the design of a classifier employs learning.
- Due to complexity of classification problems, we cannot guess the best classification decision ahead of time, we need to learn it.
- Creating classifiers then involves positing some general form of model, or form of the classifier, and using examples to learn the complete classifier.

Unsupervised and supervised learning

- In supervised learning, a teacher provides a category label for each pattern in a training set. These are then used to train a classifier which can thereafter solve similar classification problems by itself.
- In unsupervised learning, or clustering, there is no explicit teacher or training data. The system forms natural clusters of input patterns and classifies them based on clusters they belong to.
- In reinforcement learning, a teacher only says to classifier whether it is right when suggesting a category for a pattern. The teacher does not tell what the correct category is.

Summary

- Pattern recognition systems aim to decide for an action based on the data provided.
- Classification is an important step in pattern recognition systems.
- A classifier takes feature values to assign an object into a category.
- Feature values contain variability which needs to be modeled $\Rightarrow$ statistics.
- Classifiers are challenging to construct. Nevertheless many problems can be solved.