Pattern Recognition Lecture
“Introduction and Outline”

Prof. Dr. Marcin Grzegorzek

Research Group for Pattern Recognition
www.pr.informatik.uni-siegen.de

Institute for Vision and Graphics
University of Siegen, Germany

Marcin's Short CV

Time Period	**University**	**Research Focus**
1996–2002	SUT Gliwice	Image Segmentation
2002–2006	FAU Erlangen	Object Recognition
2006–2008	QMUL London	Multimedia Retrieval
2008–2010	Univ. Koblenz	Semantic Multimedia
2010–	Univ. Siegen	Pattern Recognition

Our Place at the University of Siegen

University of Siegen

Faculty IV: Science and Engineering

Depart. for Electrical Eng. and Computer Science

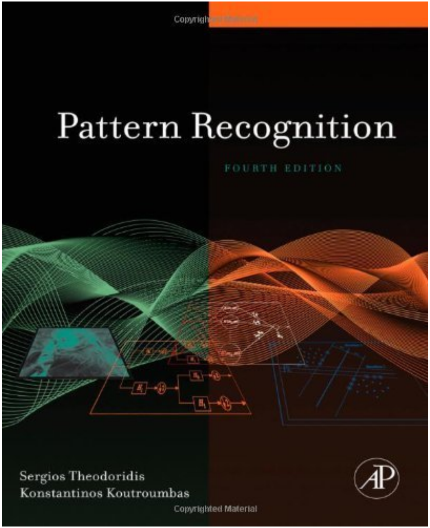
Institute for Vision and Graphics

- Computer Graphics and Multimedia Systems
- Realtime Learning Systems
- Media Systems
- Pattern Recognition

Topics

No	Topic
01	Introduction and Outline
02	Classifiers Based on Bayes Decision Theory
03	Linear Classifiers
04	Nonlinear Classifiers
05	Feature Selection
06	Feature Transformation
07	Feature Extraction
08	Template Matching
09	Context-Dependent Classification
10	Clustering: Basics Concepts
11	Clustering Algorithms I: Sequential Algorithms
12	Clustering Algorithms II: Hierarchical Algorithms
13	Clustering Algorithms III: Schemes Based on Function Optimisation
14	Summary, Applications, and Conclusions

Accompanying Book



The Term "Pattern Recognition"

Pattern Recognition

is a field whose objective is to assign an object or event to one of a number of categories, based on features derived to emphasise commonalities. In practice, features are often extracted from sensory signals, such as images or audio.

Pattern Recognition

is the act of taking in raw data and taking an action based on the category of the pattern.

Terminology

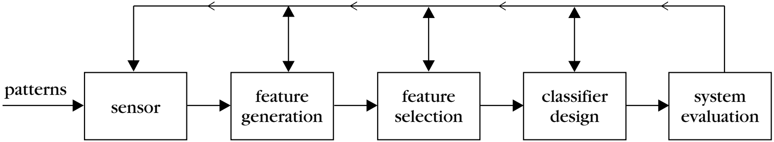
What is the difference between

- Image Processing,
- Image Recognition, and
- Pattern Recognition ?

PR Application Fields

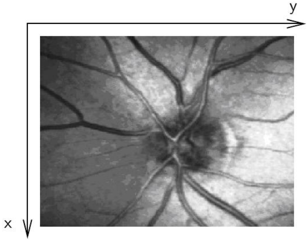
- Machine Vision
- Character Recognition
- Computed-Aided Diagnosis
- Speech Recognition
- Data Mining and Knowledge Discovery
- ...

Basic Stages of Pattern Analysis



- The stages are highly dependent on each other.
- In order to design an optimal pattern recognition system, they all have to be optimised at once.
- Patterns are analysed at different levels of abstraction.
- Integration of background knowledge into the process may be very useful.

Low-Level Interpretation of Patterns

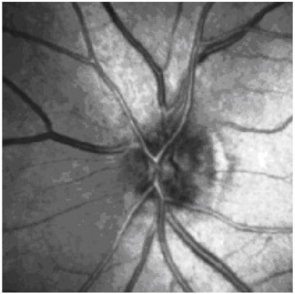


$$\begin{pmatrix} 039 & 041 & 041 & 043 & \dots & 057 \\ 041 & 043 & 043 & 043 & \dots & 066 \\ 041 & 051 & 053 & 055 & \dots & 066 \\ 041 & 051 & 053 & 056 & \dots & 066 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 054 & 052 & 042 & 042 & \dots & 095 \end{pmatrix}_{150\times 200}$$

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,199) \\ f(1,0) & f(1,1) & \dots & f(1,199) \\ \vdots & \vdots & & \vdots \\ f(149,0) & f(149,1) & \dots & f(149,199) \end{pmatrix} \quad ; \quad f(x,y) \in \{0,1,2,\dots,255\}$$

High-Level Interpretation of Patterns

Input Image



Gray Level Retina Image

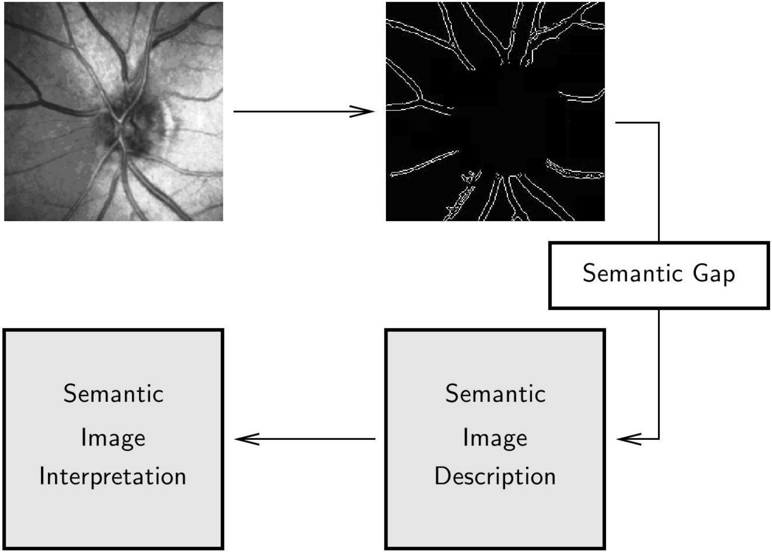
Papilla Shape - OK

Blood Vessel Width - OK

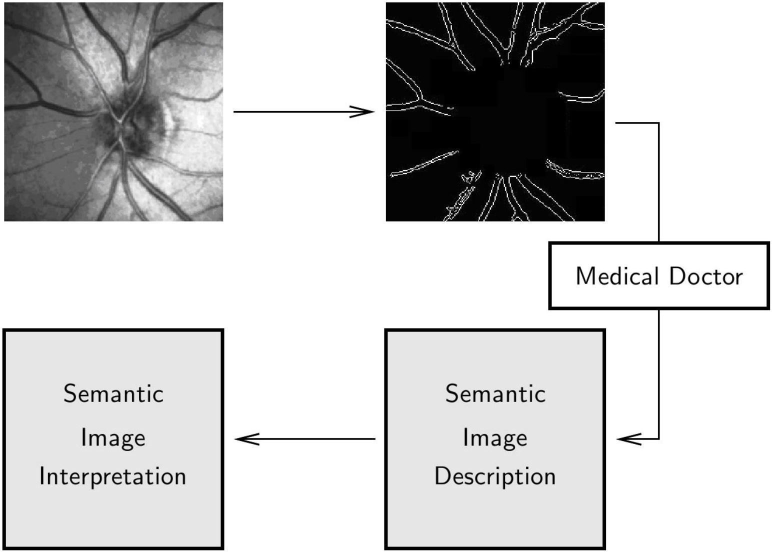
Research Challenges in Pattern Analysis

1. Optimisation of the Entire Processing Chain at Once
2. Combination of the Different Levels of Abstraction
3. Integration of Background Knowledge into the Process

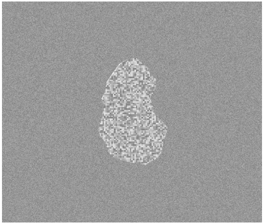
Semantic Gap in Image Understanding



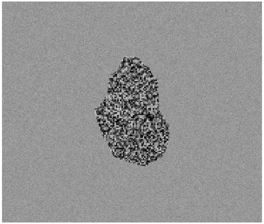
Semantic Gap in Image Understanding



Example for Medical Image Classification



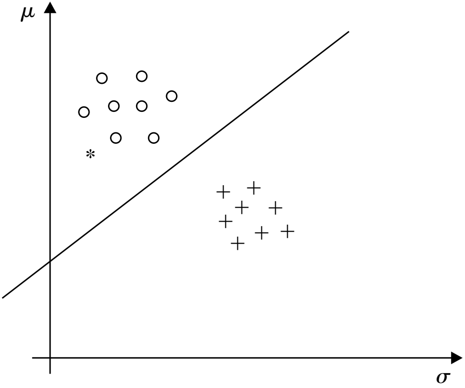
(a)



(b)

Examples of image regions corresponding to (a) class A and (b) class B.

Example Descriptors for the Image Regions



Plot of the mean value μ and standard deviation σ for a number of different images originating from class A ($\circ$) and class B ($+$).

Feature Vectors $\rightarrow$ Random Vectors

- Descriptors are called feature vectors

$$\mathbf{x} = [x_1, x_2, \ldots, x_l]^{\mathrm{T}}$$

- Each feature vector identifies a single pattern (object)

- Feature vectors are treated as random vectors

Signal Acquisition - Stochastic Process

$f(120, 180) = 219$

$f(120, 180) = 210$

$f(120, 180) = 208$

$f(120, 180) = 204$

$f(120, 180) = 198$

Bayes Decision Theory for a Two-Class Problem

Known

Classes:	$\{\omega_1, \omega_2\}$
A priori probabilities:	$P(\omega_1)$ and $P(\omega_2)$
Likelihood density functions:	$p(\mathbf{x}\|\omega_1)$ and $p(\mathbf{x}\|\omega_2)$
Pattern to be classified:	$\mathbf{x} = [x_1, x_2, \ldots, x_l]^{\mathrm{T}}$

Assumption

The feature vectors can take any value in the l-dimensional feature space: $\mathbf{x} = [x_1, x_2, \ldots, x_l]^{\mathrm{T}} \in \mathbb{R}^l$

Unknown

A posteriori probabilities: $P(\omega_1|\mathbf{x})$ and $P(\omega_2|\mathbf{x})$

Learning Strategies

Supervised Learning
: assumes that a set of labelled training data is available and the classifier is designed by exploiting this a-priori known information.

Unsupervised Learning
: clusters unlabelled training data described by feature vectors into similar groups.

Semi-Supervised Learning
: applies both the labelled and unlabelled training for designing a classification system.