Tutorial 3

Qu Xiaofeng, Teaching Asistant
quxiaofeng.at.polyu@gmail.com

COMP435p
Biometrics Authentication

February 6, 2012

The Hong Kong Polytechnic University
香港理工大學

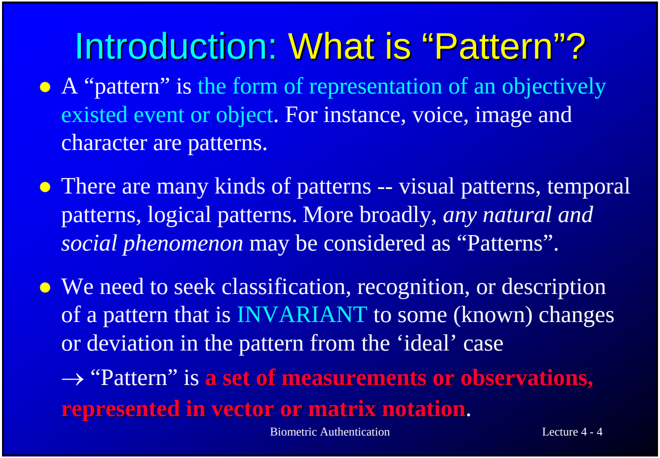
Table of Contents

Outline

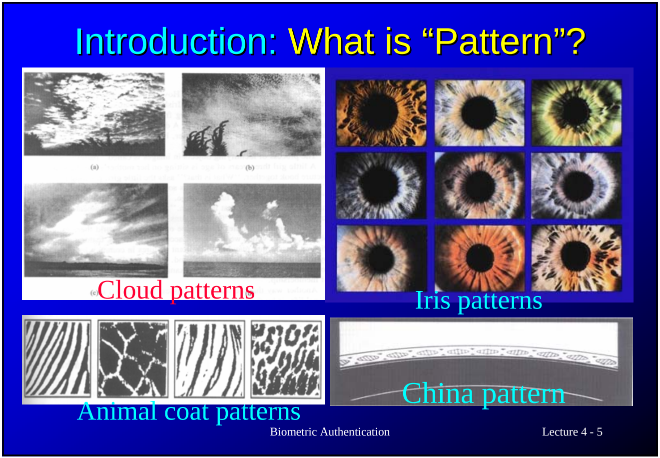
Problem 1.a

What is "Pattern"? According to its definition in (P4:4-7), could you list some patterns in your daily life?

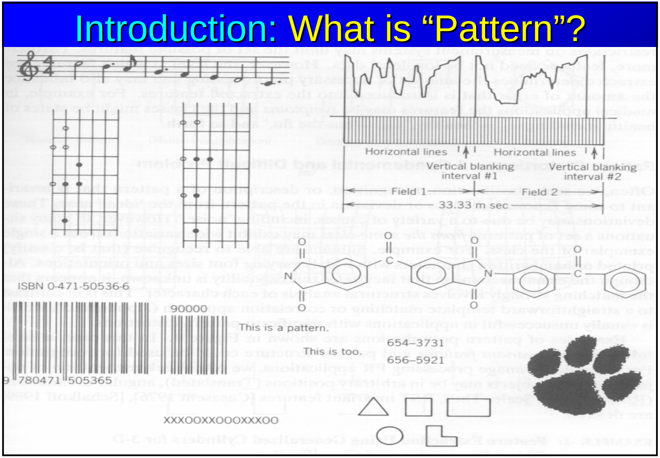
Problem 1.a


Introduction: What is "Pattern"?
A "pattern" is the form of representation of an objectively existed event or object. For instance, voice, image and character are patterns.
There are many kinds of patterns -- visual patterns, temporal patterns, logical patterns. More broadly, any natural and social phenomenon may be considered as "Patterns".
We need to seek classification, recognition, or description of a pattern that is INVARIANT to some (known) changes or deviation in the pattern from the 'ideal' case
→ "Pattern" is a set of measurements or observations, represented in vector or matrix notation.
Biometric Authentication
Lecture 4 - 4

Problem 1.a

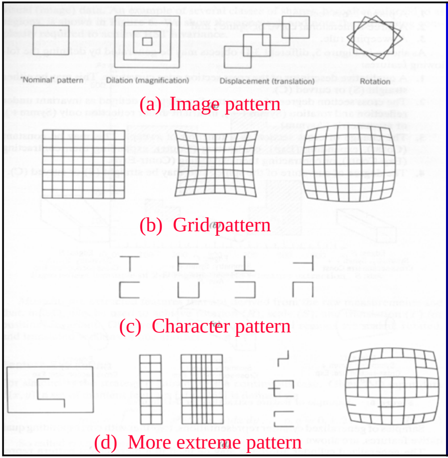

Introduction: What is "Pattern"?
Cloud patterns
Iris patterns
Animal coat patterns
China pattern
Biometric Authentication
Lecture 4 - 5

Problem 1.a



Introduction: What is "Pattern"?
Horizontal lines
Vertical blanking interval #1
Horizontal lines
Vertical blanking interval #2
Field 1
Field 2
33.33 m sec
ISBN 0-471-50536-6
90000
9 780471 505365
This is a pattern.
This is too.
654–3731
656–5921
XXXOOXXOOOXXOO

Problem 1.a



Pattern Distortion

- In many situations a set of patterns from the same class may exhibit wide variations from a single exemplar of the class.
- Invariant features: rotated, scaled, and translated invariant.

Lecture 4 - 7

Problem 1.b

"Pattern Recognition (PR)" is defined in P4:8. Could you give some PR samples and tell which features you use?

Problem 1.b

What is Pattern Recognition (PR)?

- A basic intelligent ability of human being or animal; Using a broad enough interpretation, we can find PR in every intelligent activity. For instance, you are able to recognize the number of classroom, which is the ability of *number recognition*, on the class you have to be able to understand what the teacher says and writes on the blackboard, this is the ability of *speech and character recognition*.
- From the system viewpoint, PR is an important component of intelligent systems.
- From the theoretical concept, PR is a mapping from **feature space** to **class space**.

Pattern Recognition Example

Stories of the Soap Box and the Fan

One very big cosmetics company in Japan received a complaint that a consumer had bought a soap box that was empty. Management of the company immediately traced the problem to the assembly line, that conveyed all the packaged boxes of soap to the delivery department. For some reason, one soap box went through the assembly line empty. Management asked its engineers to solve the problem. The engineers worked hard on the first solution that came to mind, using X-ray technology. They devised an X-ray machine with high-resolution monitors. Two workers were deployed full-time to monitor all the soap boxes that passed through the line to make sure they were not empty. They managed to solve the problem but it was an expensive solution.

Pattern Recognition Example

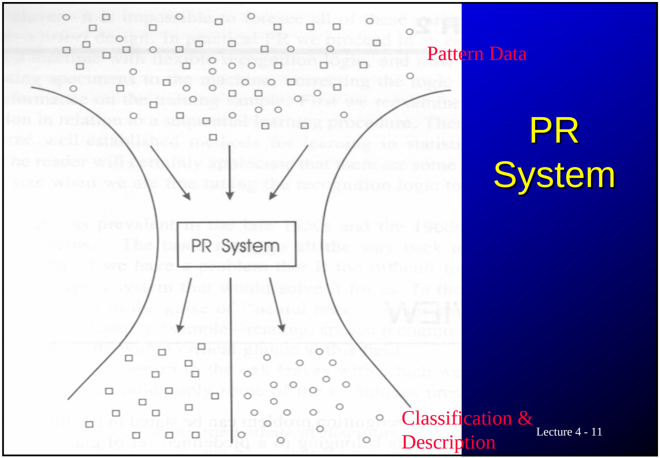
Stories of the Soap Box and the fan

Then one day, a new employee joined the company. He had previously worked in a small low-tech company. When posed with the same problem, he did not think of the X-ray machine. Instead, he bought a strong industrial electric fan and placed it facing the assembly line. He switched on the fan, and as soap boxes on the conveyor belt passed the fan, the empty ones were simply blown off the line. It was an effective and cheaper solution.

Problem 1.c

In P4:11-13, a PR system is given. Please understand each function in the system.

Problem 1.c


Pattern Data
PR System
PR
System
Classification &
Description
Lecture 4 - 11

Problem 1.c

PR System

- The design of a PR system essentially involves the following three aspects :

  1) data acquisition and preprocessing;

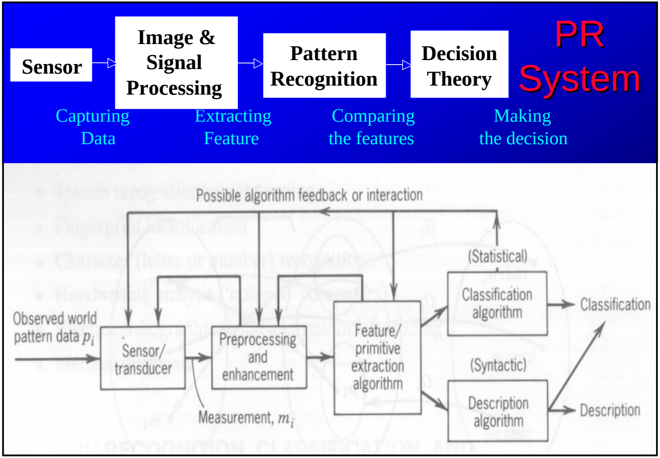
  2) data representation;

  3) decision making.

  The problem domain dictates the *choice of sensor*, *preprocessing technique*, *representation scheme*, and *the decision making model*

- It is generally agreed that a well-defined and sufficiently constrained recognition problem (small intraclass variations and large interclass variations) will lead to a compact pattern representation and a simple decision making strategy.

Problem 1.c


Sensor
Image & Signal Processing
Pattern Recognition
Decision Theory
PR System
Capturing Data
Extracting Feature
Comparing the features
Making the decision
Possible algorithm feedback or interaction
Observed world pattern data p_i
Sensor/ transducer
Preprocessing and enhancement
Measurement, m_i
Feature/ primitive extraction algorithm
(Statistical)
Classification algorithm
(Syntactic)
Description algorithm
Classification
Description

Problem 1.d

Understand the definitions related to PR: Classification, Recognition, Description, Pattern Class and Preprocessing (P4:14-15). Why classification and recognition require the different classes, c and c+1, respectively.

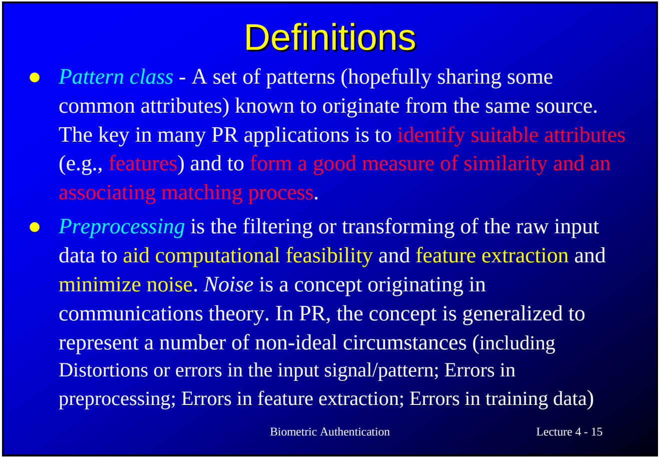
Problem 1.d

Definitions

- *Classification* - Assigns input data into one of *c* prespecified classes based on extraction of significant features or attributes and the processing or analysis of these attributes. It is common to resort to probabilistic or grammatical models in classification. A *Classifier* partitions feature space into class-labeled *decision regions*.
- *Recognition* is the ability to classify. Often we formulate PR with a *c + 1st* class, corresponding to the ‘unclassifiable’ or ‘can’t decide’ class.
- *Description* is an alternative to classification where a structural description of the input pattern is desired. It is common to resort to linguistic or structural models in description.

Biometric Authentication | Lecture 4 - 14

Problem 1.d


Definitions
Pattern class - A set of patterns (hopefully sharing some common attributes) known to originate from the same source. The key in many PR applications is to identify suitable attributes (e.g., features) and to form a good measure of similarity and an associating matching process.
Preprocessing is the filtering or transforming of the raw input data to aid computational feasibility and feature extraction and minimize noise. Noise is a concept originating in communications theory. In PR, the concept is generalized to represent a number of non-ideal circumstances (including Distortions or errors in the input signal/pattern; Errors in preprocessing; Errors in feature extraction; Errors in training data)
Biometric Authentication
Lecture 4 - 15

Problem 1.e

What is correlation between two vectors, x and y? Please use this definition in the function $||x - m_k||$ *(P4:34). How about orthogonal or uncorrelated between two vectors?*

Problem 1.e

Maximum when x and y point in the same direction

Inner Products (2)

Thus the **norm** of **x** (using the Euclidean metric) is given by

$$\| \mathbf{x} \| = \mathrm{sqrt}(\mathbf{x}' \mathbf{x}) .$$

Here are some important additional properties of inner products:

$$\mathbf{x}' \mathbf{y} = \mathbf{y}' \mathbf{x} = \| x \| \, \| y \| \cos(\text{angle between } \mathbf{x} \text{ and } \mathbf{y})$$

$$\mathbf{x}' (\mathbf{y} + \mathbf{z}) = \mathbf{x}' \mathbf{y} + \mathbf{x}' \mathbf{z} .$$

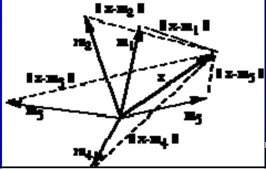
Thus, the inner product of **x** and **y** is maximum when the angle between them is zero, i.e., when one is just a positive multiple of the other. Sometimes we say (a little loosely) that **x' y** is the **correlation** between **x** and **y**, and that the correlation is maximum when **x** and **y** point in the same direction. If **x' y** = 0, the vectors **x** and **y** are said to be **orthogonal** or uncorrelated.

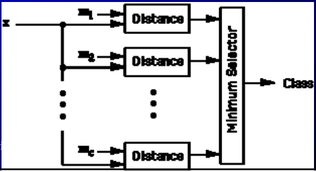
Problem 1.e

Minimum-Distance Classifiers

Template matching can easily be expressed mathematically. Let **x** be the feature vector for the unknown input, and let **m**1, **m**2, ..., **m**c be templates (i.e., perfect, noise-free feature vectors) for the c classes. Then the error in matching **x** against **m**k is given by || **x** - **m**k ||.

Here || **u** || is called the **norm** of the vector **u**. A minimum-error classifier computes || **x** - **m**k || for k = 1 to c and chooses the class for which this error is minimum. Since || **x** - **m**k || is also the distance from **x** to **m**k, we call this a **minimum-distance** classifier. Clearly, a template matching system is a minimum-distance classifier.





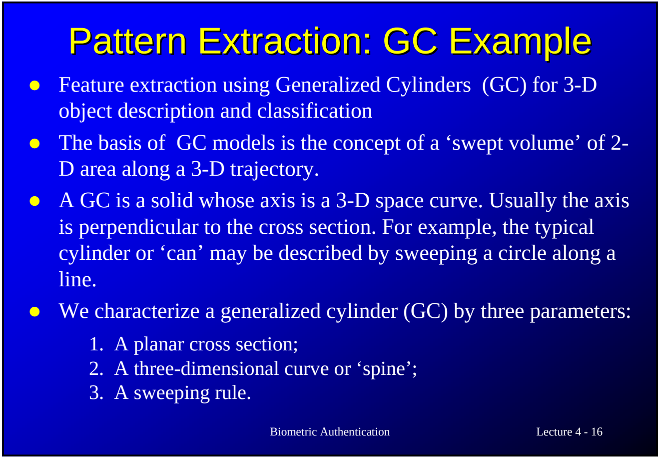
Outline

1. Problems
   - Problem 1: Answer the questions
   - Problem 2: Pattern extraction
   - Problem 3: Minimum-error classifier
   - Problem 4: Classify

Problem 2: Pattern extraction

A 3-D example for pattern extraction is defined in P4: 16-18. There are four kinds of features are extracted. Please select these features according to some guidelines (P4: 27-28) to classify the given six objects.

Problem 2: Pattern extractionterms



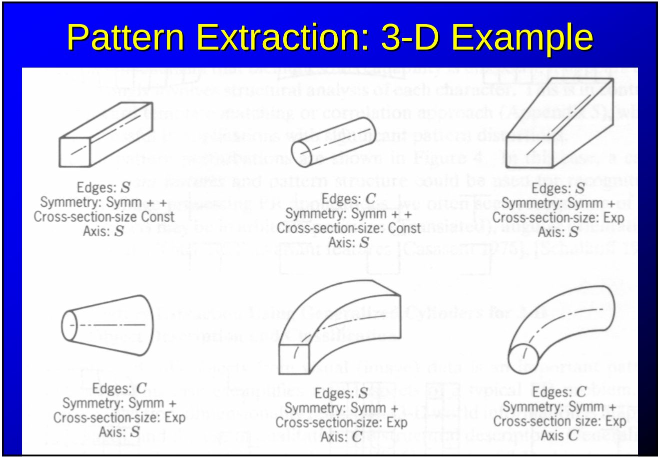
Problem 2: Pattern extractionterms

Pattern Extraction: 3-D Example

- 3-D objects, defining the following features

1. A qualitative descriptor of the cross section edge curvature. This may either straight (S) or curved (C).
2. The cross section degree of symmetry, defined as :
   -- Invariant under reflection and rotation (Symm ++)
   -- Invariant under reflection only (Symm +)
   -- Asymmetric (Asymm).
3. The change of cross section as a function of sweep, defined as
   -- Constant (Const)
   -- Expanding (Exp)
   -- Contracting (Contr)
   -- Contracting then expanding (Contr-Exp)
   -- Expanding then contracting (Exp-Contr)
4. The degree of curvature of the spine. This may be straight (S) or curved (C).

Problem 2: Pattern extractionterms



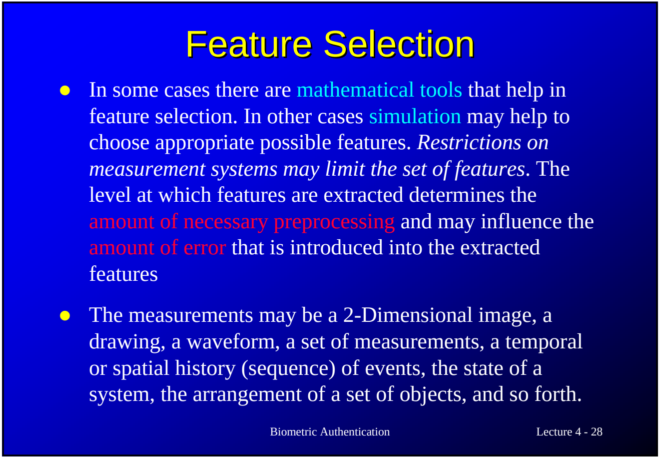
Problem 2: Pattern extractionterms

Feature Selection

- It is the process of choosing input to PR system and involves judgment.
- It is important that the extracted features be relevant to the PR task at hand.
- It has to be addressed at the outset of any PR system design.
- To choose and to extract features that :
  1. are computationally feasible;
  2. lead to ‘good’ PR system success;
  3. reduce the problem data (e.g., raw measurements) into a manageable amount of information without discarding valuable (or vital) information

Biometric Authentication Lecture 4 - 27

Problem 2: Pattern extractionterms



One possible selection: Edge+Symmetry+Axis

Outline

Problem 3: Minimum-error classifier

A minimum-error classifier based on $\|x - m_k\|$ *is defined in P4:34. How to get a linear discriminant function,* $g(x) = m'x - 0.5 \times \|m\|$ *in P4:39? Which difference between the following two classifiers using the approaches mentioned above.*

Problem 3: Minimum-error classifier

To find a k, which will get $Min(\|x - m_k\|)$ is equivalent to get the $Min(\|x - m_k\|^2)$

$$
\begin{aligned}
Min(\|x - m_k\|^2) =& (x - m_k)'(x - m_k) \\
=& x'x - x'm_k - m_k'x + m_k'm_k \\
=& x'x - 2m_k'x + \|m_k\|^2 \\
=& x'x - 2(m_k'x - 0.5\|m_k\|^2) \\
\textit{equivalent to } & Max(m_k'x - 0.5\|m_k\|^2) \\
\textit{so } g(x) =& m'x - 0.5 \times \|m\|^2
\end{aligned}
$$

Problem 3: Minimum-error classifier

Maximum when x and y point in the same direction

Inner Products (2)

Thus the **norm** of **x** (using the Euclidean metric) is given by

$$\| \mathbf{x} \| = \mathrm{sqrt}(\mathbf{x}' \mathbf{x}) .$$

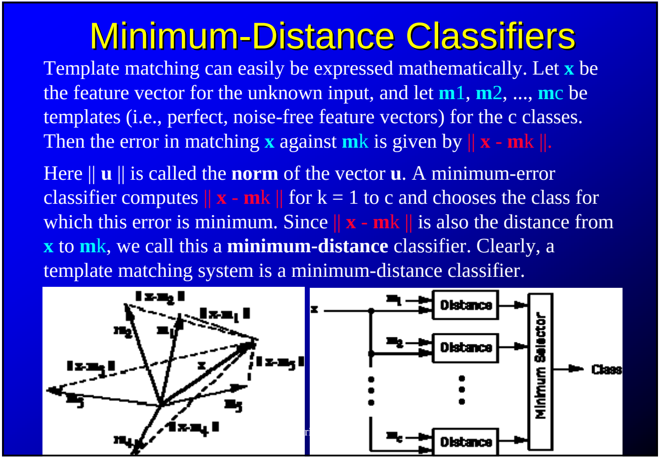
Here are some important additional properties of inner products:

$$\mathbf{x}' \mathbf{y} = \mathbf{y}' \mathbf{x} = \| x \| \, \| y \| \cos(\text{angle between } \mathbf{x} \text{ and } \mathbf{y})$$

$$\mathbf{x}' (\mathbf{y} + \mathbf{z}) = \mathbf{x}' \mathbf{y} + \mathbf{x}' \mathbf{z} .$$

Thus, the inner product of **x** and **y** is maximum when the angle between them is zero, i.e., when one is just a positive multiple of the other. Sometimes we say (a little loosely) that **x' y** is the **correlation** between **x** and **y**, and that the correlation is maximum when **x** and **y** point in the same direction. If **x' y** = 0, the vectors **x** and **y** are said to be **orthogonal** or uncorrelated.

Problem 3: Minimum-error classifier

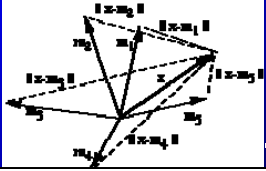


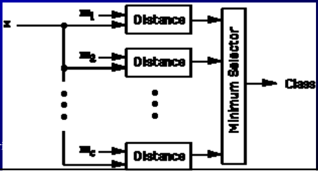
Problem 3: Minimum-error classifier

Minimum-Distance Classifiers

Template matching can easily be expressed mathematically. Let **x** be the feature vector for the unknown input, and let **m**1, **m**2, ..., **m**c be templates (i.e., perfect, noise-free feature vectors) for the c classes. Then the error in matching **x** against **m**k is given by || **x** - **m**k ||.

Here || **u** || is called the **norm** of the vector **u**. A minimum-error classifier computes || **x** - **m**k || for k = 1 to c and chooses the class for which this error is minimum. Since || **x** - **m**k || is also the distance from **x** to **m**k, we call this a **minimum-distance** classifier. Clearly, a template matching system is a minimum-distance classifier.





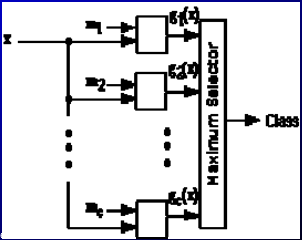
Problem 3: Minimum-error classifier

Linear Discriminants (2)

Let us define the **linear discriminant function** $g(\mathbf{x})$ by

$$g(x) = \mathbf{m}'\,\mathbf{x} - 0.5\,\|\,\mathbf{m}\,\|_2 .$$

Then we can say that a minimum-Euclidean-distance classifier classifies an input feature vector **x** by computing *c* linear discriminate functions $g1(\mathbf{x})$, $g2(\mathbf{x})$, ... , $gc(\mathbf{x})$ and assigning **x** to the class corresponding to the maximum discriminant function. We can also think of the linear discriminant functions as measuring the correlation between **x** and **m**k, with the addition of a correction for the "template energy" represented by $\|\,\mathbf{m}k\,\|_2$. With this correction included, a minimum-Euclidean-distance classifier is equivalent to a maximum-correlation classifier.



Biometric Auth

Outline

Problem 4: Classify

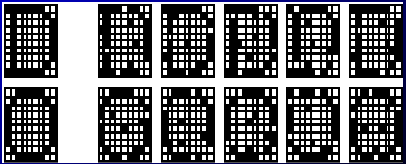
In P4:33 there are two templates ("D" and "O") and ten input samples, each five. Please classify the third input example in the "D" row by using the maximum correlation approach and minimum error approach, respectively.

Problem 4: Classify

Template Matching

Template matching is a natural approach to pattern classification. For example, consider the noisy "D"'s and "O"'s. The noise-free versions can be used as **templates**. To classify one of the noisy examples, simply compare it to the two templates. Two ways can be done:

1. Count the number of agreements (black matching black and white matching white). Pick the class that has the maximum number of agreements. This is a **maximum correlation** approach.
2. Count the number of disagreements (black where white should be or white where black should be). Pick the class with the minimum number of disagreements. This is a **minimum error** approach.



Problem 4: Classify

For Maximum Correlation Approach: "D"-Input: 82/90=91% and "O"-Input: 66/90=73%) □

Problems

Any questions?