9. Pattern Recognition

1. Basic concepts of pattern recognition

A pattern is the description of an object. We can spot a friend in a crowd and recognize what he says, we can recognize the voice of a known individual, we can read handwriting and analyze fingerprints, we can distinguish smiles from gestures of anger.

We can divide our acts of recognition into two major types: the recognition of ***concrete*** items and the recognition of ***abstract*** items.

- Concrete items : we recognize characters, pictures, music, and object around us. This may be referred to as sensory recognition, which includes visual and aural pattern recognition. This recognition process involves the identification and classification of spatial and temporal patterns.
- Abstract items : we can recognize an old argument, or a solution to a problem, with our eyes and ears closed. This process involves the recognition of abstract items and can be termed ***conceptual recognition.***

We will concentrate on the first type, which include ***spatial patterns*** : characters, fingerprints, weather maps, physical objects, and pictures, and ***temporal patterns*** : speech, speaker, target signature.

Pattern recognition can be defined as the categorization of input data into identifiable classes via the extraction of significant features or attributes of the data from a background of irrelevant detail.

Table (1) describes several classification tasks, together with the corresponding input data and output responses.

Task of Classification	Input Data	output Response
Character recognition	Optical signal	Name of character
Speech recognition	Acoustic waveform	Name of word
Speaker recognition	Voice	Name of speaker
Weather prediction	Weather maps	Weather forecast
Medical diagnosis	Symptoms	Disease

Table (1)

2. Fundamental problems in pattern recognition system design

The design of an automatic pattern recognition system generally involves several major problem areas.

First : is concerned with the representation of the input data which can be measured from the objects to be recognized. This is the sensing problem. Each measured quantity describes a characteristic of the pattern or object. Suppose, for example, that the pattern in question are alphanumeric characters.

The ***pattern vectors*** contain all the measured information available about the patterns. The measurements performed on the objects of a pattern class may be regarded as a coding process which consist of assigning to each pattern characteristic a symbol from the alphabet set$\{x_i\}$. When the measurements yield information in the form of real numbers, it is often

useful to think of a pattern vector as a point in the n-dimensional Euclidean space.

Second: this problem concerns the ***extraction*** of character ***features*** or attributes from the received input data and the ***reduction*** of ***dimensionality*** of pattern vectors. This is often referred to as the preprocessing and feature extraction problem.

Correct recognition will depend on :

1. The amount of ***discrimination information*** contained in the measurement .
2. The ***effective utilization*** of this information.

Third : this problem in pattern recognition system design involves the determination of ***optimum decision procedures***, which are needed in the identification and classification process. After the observed data from patterns to be recognized have been expressed in the form of pattern points or measurement vector in the pattern space, ***we want the machine to decide to which pattern class these data belong.***

Characteristic of adaptive pattern recognition

1. Resistance مقاوم to distortions.
2. Flexible under large pattern deviationsانحرافات.
3. Capable of self-adjustment تعديل.

3. Design Concepts and Methodologies

Three basic design concepts are discussed in the following paragraphs.

1. Membership-roster concept

Characterization of pattern class by roster of its members suggest automatic pattern recognition by *template matching*. The set of patterns belonging to

the same pattern class is stored in the pattern recognition system. When an unknown pattern is shown to the system, it is compared with the stored patterns one by one. The pattern recognition system classifies this input pattern as a member of a pattern class if it matches one of the stored patterns belonging to that pattern class.

2. **Common-property concept**

   Characterization of pattern class by common properties shared by all of its members suggests automatic pattern recognition via the detection and processing of **similar features**. The basic assumption in this method is that ***patterns belonging to the same class posses certain common properties or attributes which reflect similarities among these patterns.***

3. **Clustering concept**

   - When the patterns of a class are vectors whose components are ***real numbers***, a pattern class can be characterized by its clustering properties in the pattern space. The design of pattern recognition system based on this general concept is guided by the relative ***geometrical arrangement*** of the various pattern clusters.
   - If the class are characterized by clusters which are ***far apart***, simple recognition schemes such as the ***minimum-distance classifiers*** may be successfully employed.
   - When the ***clusters overlap***, however, it becomes necessary to utilize more ***sophisticated techniques*** for ***partitioning the pattern space.*** Overlapping clusters are the result of a ***less*** in observed information and the presence of measurement ***noise***.

The ***basic design concepts*** for automatic pattern recognition described above may be ***implemented by three principal*** categories of ***methodology*** :

1. **Heuristic methods**

   The heuristic approach is based on ***human intuition and experience***, making use of the common-property concepts. A system designed using this principle generally consists of ***a set of ad hoc*** procedures developed for specialized recognition tasks. The structure and performance of a heuristic system will ***depend*** to a large degree on the ***cleverness and experience of the system designers .***

2. **Mathematical methods**

   The mathematical approach is based on classification rules which are formulated and derived in a mathematical framework , making use of the common-property and clustering concepts. This is in contrast with the heuristic approach , in which decisions are based on ad hoc rules.

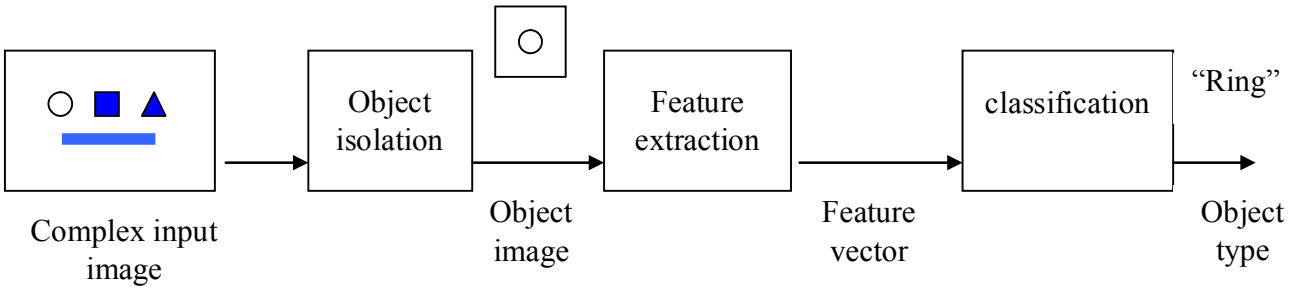
3. **Linguistic (syntactic) methods**

Characterization of patterns by primitive elements (sub patterens) and their relationships suggest automatic pattern recognition by the linguistic or syntactic approach, making use of the common-property concept. This permits application of formal languages theory to the pattern recognition problem. A pattern grammar is considered as consisting of finite sets of elements called ***variables, primitives, and production. The rules of production determine the type of grammar***.

Among the most studied grammars are regular grammars, context-free grammars, and context-sensitive grammar.

Optical Pattern Recognision

Given a digitized image containing several objects, the pattern recognition process consists of three major phases:



The three phases of pattern recognition

The ***first*** phase is object isolation, in which each object must be found and its image isolated from the rest of the scene. The ***second*** phase is called feature extraction. The feature is a set of measurable properties. The feature extraction phase measures these properties, producing a set of measurements called the feature vector. This drastically reduced amount of information represents all the knowledge upon which the subsequent classification must be based. The third phase is object classification and its output is merely a decision regarding the class to which the object belongs. The object is thus recognized as being one particular type of objects, and the recognition is implemented as a classification process. Each object is assigned to one of several pre-established groups (classes) that represent the possible type of objects expected to be encountered. The classification is based solely on the feature vector.

1. Pattern Recognition System Design

The design of a pattern recognition systems is usually done in the five steps listed in table below. These are object locator design, feature selection, classifier design, classifier training, and performance evaluation.

The object locator is the algorithm that isolates the image of the individual objects in the complex scene, this is scene segmentation feature selection involve deciding which properties of the object (size, shape , etc) best distinguish among the various object type and thus should be measured. Classification design consist of establishing a mathematical basis for the classification procedure. The various adjustable parameters of the classifier itself (decision threshold , etc) are pinned down in the training stage. Finally , it is usually desirable to estimate the error rates that can be expected with the system. This constitutes the performance evaluation step.

Pattern Recognition System Design

Step	**Function**
1. Object location design	**Select the scene segmentation algorithm that will isolate the individual objects in the image.**
2. Feature selection	**Decide Which properties of the objects best distinguish the object type and how to measure these.**
3. classification design	**Establish the mathematical basis of the classification algorithm and select the type of classifier structure to be used.**
4. Classifier training	**Fix the various adjustable parameters (decision boundaries ,etc) in the classifier to**

	suit the object being classified.
5. performance evaluation	**Estimate the expected rates of the various possible misclassification errors.**

2. *Patterns and pattern classes*

A pattern is a quantative or structural description of an object or some other entity of interest in an image. A pattern class is a family of patterns that share some common properties. Pattern recognition by machine involve techniques for assigning patterns to their respective classes automatically and with as human intervention as possible.

The two principal pattern arrangements used practice are vectors (for quantitative descriptions) and trees (for structural descriptions). Pattern vectors are represented by lower case letters , such as x,y and z and take the form :

$$X= \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ xn \end{pmatrix}$$

Where each component xi represent the i th descriptor (feature) and n is the number of such descriptors. The nature of the components of a pattern vector x depends on the measurement technique used to describe the physical pattern itself. For example, suppose that we want to describe three type of iris flowers (Iris setosa, virglnica, and versicolor) by measuring the width

and length of their petals. In this case we would be dealing with pattern vector of the form

$$X= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Where x1, x2 correspond to petal length and width, respectively. Because the petals of all flowers vary in width and length to some degree, the pattern vectors describing these flowers also will vary, not only between different classes, but also within a class.

3. ***Feature Selection***

The goal in image analysis is to extract information useful for solving application based problems. This is done by intelligently reducing the amount of image data with the tools we have explored, including edge detection and segmentation. After we have performed these operations, we have modified the image from the lowest level of pixels data into higher-level representations. Now, we can consider extraction of features that can be useful for solving computer imaging problem.

If we desire a system to distinguish objects of different type, we must first decide which parameters, descriptive of the objects, will be measured. The particular parameters that are measured are called the features.