Introduction to Pattern Recognition[1]

Anil K. Jain, Robert P.W. Duin

Informally, a pattern is defined by the common denominator among the multiple instances of an entity. For example, commonality in all fingerprint images defines the fingerprint pattern; the commonality in fingerprint images of John Doe's left index finger defines the John-Doe-left-index-fingerprint pattern (see Figure 1 - showing a bunch of fingerprints of the same finger; and a bunch of impressions of arbitrary fingers in Figure 2). Thus, a pattern could be a fingerprint image, a handwritten cursive word, a human face, a speech signal, a bar code, or a web page on the Internet (see Figure 3). Often, individual patterns may be grouped into a category based on their common properties; the resultant group is also a pattern and is often called a pattern class. Pattern recognition is the science for observing (sensing) the environment, learning to distinguish patterns of interest (e.g., animals) from their background (e.g., sky, trees, ground), and making sound decisions about the patterns (e.g., Fido) or pattern classes (e.g., a dog, a mammal, an animal).

Figure 1: Examples of patterns: six fingerprints from the same finger of the same person.

Figure 2: Examples of patterns: six fingerprints from different fingers of different persons.

1. Introduction.

Since our early childhood, we have been observing patterns in the objects around us (e.g., toys, flowers, pets, and faces). Learning patterns also reinforces, and is reinforced by, the acquisition of language. By the time children are five years old, most can recognize digits and letters. Small and large characters, handwritten and machine printed characters, characters of different colors and orientations and partially occluded letters - all are easily recognized by the young. We take this ability for granted until we face the task of teaching a machine how to recognize the characters. In spite of almost 50 years of research, design of general-purpose machines for pattern recognition remains an elusive goal.

Humans are the best pattern recognizers in most scenarios, yet we do not fully understand how we recognize patterns. Ross (1998) emphasizes the work of Nobel Laureate Herbert Simon whose central finding is that pattern recognition is critical in most human decision making tasks: "The more relevant patterns at your disposal, the better your decisions will be. This is hopeful news to proponents of artificial intelligence, since computers can surely be taught to recognize patterns. Indeed, successful computer programs that help banks score credit applicants, help doctors diagnose diseases and help pilots land airplanes depend in some way on pattern recognition."

We will first describe the area of pattern recognition in detail and relate it to the more restricted problem of pattern classification. This is followed by systems for automatic pattern recognition. In particular, we describe some methods for generalization, i.e., how can the derived decision rules be applied to new observations. Next, some aspects of pattern learning are discussed that may play a role in human learning as well. Finally, some applications are described that are already in use in various sectors of our society.

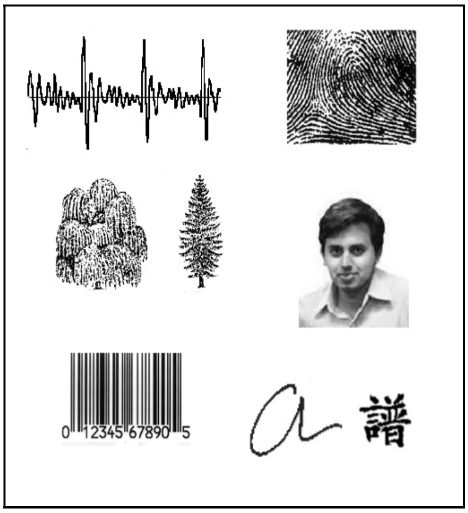

Figure 3: Examples of patterns: Sound wave, fingerprint, trees, face, bar code and character images

1. This text is taken from a contribution of the authors in R.L. Gregory (eds.), *The Oxford Companion to the Mind, Second Edition*, Oxford University Press, Oxford, UK, 2004, 698-703.

2. Pattern recognition and classification.

Pattern recognition aims to make the process of learning and detection of patterns explicit, such that it can partially or entirely be implemented on computers. Automatic (machine) recognition, description, classification (grouping of patterns into pattern classes) have become important problems in a variety of engineering and scientific disciplines such as biology, psychology, medicine, marketing, computer vision, artificial intelligence, and remote sensing. In almost any area of science in which observations are studied but the underlying mathematical or statistical models are not available, pattern recognition can be used to support human concept acquisition or decision making. Given a group of objects, there are two ways to build a classification or recognition system (Watanabe 1985), supervised, i.e., with a teacher, or unsupervised, without the help of a teacher, see Figure 4.

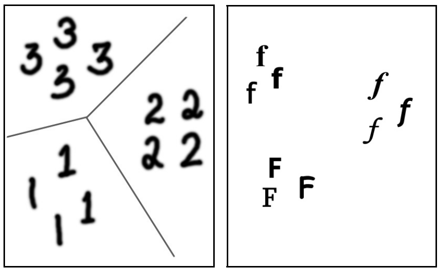

Figure 4: 4(a). Supervised pattern recognition deals with classifying objects with (known) different labels.

4(b). In unsupervised pattern recognition, classes or subclasses have to be derived from the data.

Interest in pattern recognition has been renewed recently due to emerging applications which are not only challenging but also computationally more demanding, such as data mining, document classification, organization and retrieval of multimedia databases, and biometric authentication (i.e., face recognition and fingerprint matching).

3. Systems for automatic pattern recognition.

Rapid advances in computing technology not only enable us to process huge amounts of data, but also facilitate the use of elaborate and diverse methods for data analysis and classification. At the same time, demands on automatic pattern recognition systems are rising enormously due to the availability of large databases and stringent performance requirements (faster recognition speed and higher accuracy at a lower cost). In many emerging applications, it is clear that no single approach for classification is "optimal" and multiple methods and approaches have to be used. Consequently, combining several sensing modalities and classifiers is now a common practice in pattern recognition.

The design of a pattern recognition system essentially involves the following four aspects:

(i) data acquisition and preprocessing, e.g., taking a picture of an object and removing the irrelevant background,

(ii) data representation, e.g. deriving relevant object properties (like its size, shape and color) which efficiently offer pertinent information needed for pattern recognition,

(iii) training, e.g., imparting pattern class definition into the system, often, by showing a few typical examples of the pattern, and (iv) decision-making that involves finding the pattern class or pattern description of new, unseen objects based on a training set of examples. The application domain dictates the choice of sensor(s), preprocessing technique, representation scheme, and the decision making model. It is generally agreed that a well-defined and sufficiently constrained classification problem will lead to a compact pattern representation and a simple decision making strategy. Learning from a set of examples (training set) is an important and desired characteristic of most pattern recognition systems, in contrast with systems consisting of handcrafted decision rules only.

The five major approaches for pattern recognition are (Jain, Duin, Mao 2000):

- *Template matching*. Objects are directly compared with a few stored examples or prototypes that are representative of the underlying classes. Because of the large variations often encountered in these examples, the template matching is not the most effective approach to pattern recognition.
- *Geometrical classification*. Classes are represented by regions in the representation space (e.g. a feature space as in Figure 5) defined by simple functions such that the training examples are

classified as correctly as possible. Suppose, the average value of (height, weight) of women is (5’5”, 125 lb) and that of men is (5’11”, 157lb). A simple geometric woman vs. man classifier using (height, weight) as a two-dimensional representation may simplistically divide the representation space into two triangular regions (similar to Figure 4(a)). So, a person with (height, weight) = (5’2”, 122lb) will classified by this classifier as a woman.

- *Statistical classification*. Continuing with the foregoing example, a statistical classifier may estimate the statistical distribution of the two features, namely, height and weight of the two classes of interest (women and men) from known samples. At any coordinate or point in the representation space, one could estimate the likelihoods of it being a man or a woman; depending upon which likelihood is higher, one could determine the class of an entity. This method differs from the geometrical method in that the classes are not (pre)defined in terms of any regular shapes in the representation space.

- *Syntactic or structural matching*. The height and weight representation space is too simplistic and it is conceivable that a person’s body shape is a better representation for determining his/her gender. One could decompose the shape of a person into component parts and describe the shape in terms of component parts and their relationships (e.g., how they are attached to each other). Now, the determination of gender could be performed either based on the shapes of the individual body parts, their relationships or both. In syntactic or structural approach, a complex pattern (e.g., animal) is described in terms of component patterns (e.g., hair and head, or, torso and limbs) and their relationship (e.g., articulated joints). Strategies for learning such a language (defining the structure) from examples are problematic, as it is essentially difficult to compensate for noise (see Fu (1983), also Perlovsky (1998)).

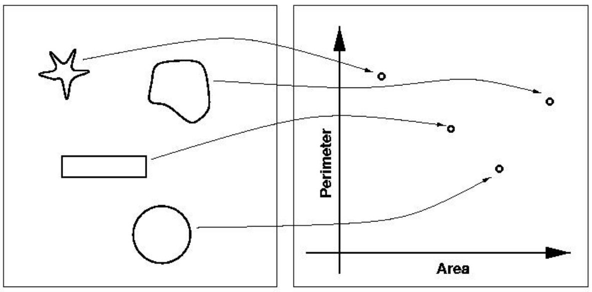

Figure 5: Example of four objects represented by two features (area and perimeter) in a two-dimensional feature space.

- *Artificial neural networks*. These networks attempt to apply the models of biological neural systems to solve practical pattern recognition problems. This approach has become so popular that the use of neural networks for solving pattern recognition problems has become an area on its own, and is often studied outside the biological context, e.g., see the books by Bishop (1995) and Ripley (1996).

It is interesting to compare these approaches for automatic pattern recognition with the various ways the human learning process may be modeled: simulation of the neural system itself, or simulation of the processes in that system, either based on direct information from the senses (like in the statistical and the geometrical approaches) or on higher level symbolic information (like in the structural approach). The template matching procedure can be compared with learning by storing all facts without understanding them.

4. Some challenges in pattern and class learning

(i) *Selection of training sets*. If we want to learn from examples, care should be paid to the way the examples are selected. For instance, a system for the recognition of electrocardiograms (say, into normal heart vs. diseased heart) can be based on examples collected in hospitals, examples collected in a general screening test, on typical cardiograms that are clear examples of particular classes of heart problems or on selected cardiograms that are the border cases between these classes. The choice of such a strategy is strongly related to the learning approach to be used and to the way the recognition system can be used.

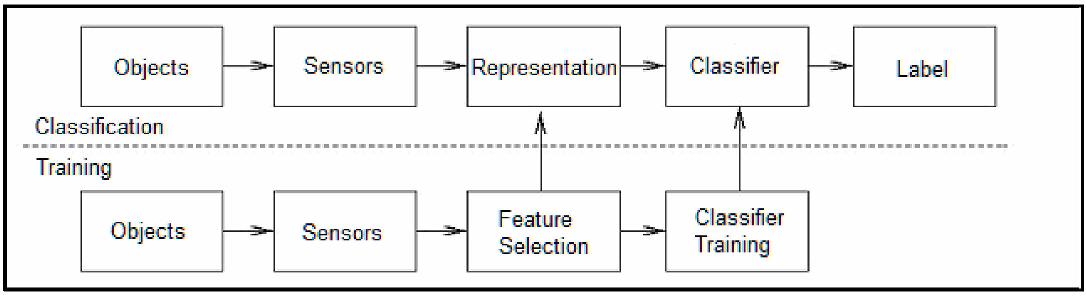

Figure 6: Design of a pattern recognition system.

(ii) *Representation of objects*. There are various ways to represent objects: raw data measurements (e.g., overall height, overall weight), derived measurements or features (e.g., ratio of height to weight), a structural description (e.g., height to weight ratios of parts of bodies and spatial relationship of the body parts), etc. In the statistical approach, the feature representation is the most common. For the recognition of simple real world objects, the features can be their sizes, shapes, colors, etc. More features do not necessarily imply a better classification performance. Given a representation scheme, an objective measure (e.g., “distance” or “score”) needs to be defined to quantify the (dis)similarity between any two representations.

(iii) *Inter- and intra-class distances*. A direct and intuitive way to see whether a feature representation is good for a classification problem is to compare the inter-class distances (e.g., between the two sets of pictures of two different persons) with the intra-class distances (e.g., between all pictures of a single person). If the inter-class distances are much larger than the intra-class distances, the classification problem is easy. If they are of similar orders, either the classes overlap, or a more advanced procedure is needed to separate the classes. Obviously, a representation with large inter-class variability and small intra-class variability is desirable. See Figure 7 for an illustration on inter- and intra-class distances.

(iv) *Invariance of representation*. Some object variations may not be important for the classification task, e.g., the size of a character, the angle (pose) at which a face is observed, the speed by which a word is spoken. These variations may influence the representation so that the position of the object in feature space is changed. An important problem is how to identify and extract these so-called invariants. We can collect objects under all possible variations, which is expensive. A preferred approach is to use invariant features.

(v) *The problem of overtraining*. An overly complex pattern recognition system may learn unnecessary details of training samples of a pattern and consequently, will be unable to recognize the essential commonality defining the pattern. It is necessary to adapt the complexity of the recognition system to the complexity and size of the data set under consideration.

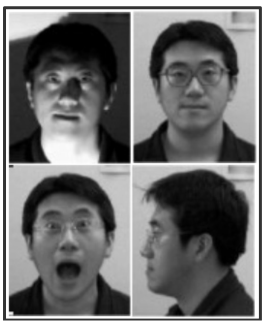

Figure 7: Different faces of the same person, or different persons?
(a) faces belonging to two different people.
(b) multiple faces of the same person.

http://www.visi.com/~charlesr/twinsdays_98/

5. Pattern recognition applications.

Pattern recognition is used in any area of science and engineering that studies the structure of observations. It is now frequently used in many applications in manufacturing industry, health care and military. Examples include:

- Optical character recognition (OCR) is becoming an integral part of document scanners, and is also used frequently in banking and postal applications. Printed characters can now be accurately recognized, and the improving performance of automatic recognition of handwritten cursive characters has diminished significantly the need of human interaction for OCR tasks.
- Automatic speech recognition is very important for user interaction with machines. Commercial systems for automatic response to flight queries, telephone directory assistance and telebanking are available. Often the systems are tuned to a specific speaker for better recognition accuracy.
- Computer vision deals with the recognition of objects as well as the identification and localization of their three-dimensional environments. This capability is required, for example, by robots to operate in dynamic or unknown environments. This can be useful from applications ranging from manufacturing to household cleaning, and even for rescue missions.
- Personal identification systems that use biometrics are very important for security applications in airports, ATMs, shops, hotels, and secure computer access. Recognition can be based on face, fingerprint, iris or voice, and can be combined with the automatic verification of signatures and PIN codes.
- Recognition of objects on earth from the sky (by satellites) or from the air (by airplanes and cruise missiles), is called remote sensing. It is important for cartography, agricultural inspection, detection of minerals and pollution, and target recognition.
- Many tests for medical diagnosis utilize pattern recognition systems, from counting blood cells and recognition of cell tissues through microscopes to the detection of tumors in magnetic resonance scans and the inspection of bones and joints in X-ray images.
- Many large databases are stored on the repositories accessible via Internet or otherwise in local computers. They may have a clear structure such as bank accounts, a weak structure such as consumer behavior, or no obvious structure such as a collection of images. Procedures for finding desired items (database retrieval) as well as to learn or discover structures in databases (data mining) are becoming more and more important. Web search engines and recommender systems are two example applications.

References

Bishop C.M. (1995), *Neural Networks for Pattern Recognition*. Clarendon Press.

Duda, R.O., Hart, P.E. and Stork, D.G. (2001), *Pattern Classification and Scene Analysis*, 2nd ed. Wiley.

Fu, K.S. (1983). ‘A step towards unification of syntactic and statistical pattern recognition’. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 5, no. 2.

Jain, A.K., Duin, R.P.W. and Mao, J. (2000), ‘Statistical Pattern Recognition: A Review’. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1.

Perlovsky, L.I. (1998), ‘Conundrum of combinatorial complexity’. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6.

Picard, R. (1997). *Affective Computing*. MIT Press.

Ripley, B. (1996), *Pattern Recognition and Neural Networks*. Cambridge University Press.

Ross, P.E. (1998). ‘*Flash of Genius*’. Forbes.

Watanabe, S. (1985). *Pattern Recognition: Human and Mechanical*. Wiley.