UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org

Distr.
LIMITED

UNIDO/IS.609
14 February 1986

ENGLISH

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

15437

APPLICATIONS OF PATTERN RECOGNITION AND IMAGE PROCESSING TO INDUSTRIAL PROBLEMS IN DEVELOPING COUNTRIES*

prepared by

TATA Research Development and Design Centre**

---

* The views expressed in this paper are those of the author and do not necessarily reflect the views of the secretariat of UNIDO. Mention of firm names and commercial products does not imply the endorsement of UNIDO. This document has been reproduced without formal editing.

** 1, Mangaldas Road, Pune 411 001, India.

V.86-51868

Summary

Outlines the problems connected with pattern recognition and image processing especially in the context of the developing country problematic. The basic ideas of pattern recognition and image processing are presented, including a review of image processing and pattern recognition techniques. Several examples of recognition systems are given. The relevance of pattern recognition and image processing in the industrial context of developing countries is discussed and specific applications selected for this environment are outlined The document also looks at the elements supporting introduction of pattern recognition and image processing techniques to the industrial practice of developing countries such as education and training, hardware and software. There is also a glossary of terms specific to the field.

TABLE OF CONTENTS

I. INTRODUCTION

Computers are becoming an integral part of modern society in developed countries. Computers allow better management and use of information, better planning, and they are essential for industrial automation which permits tremendous flexibility in manufacturing. All these uses of computers generally lead to greater productivity and competitive advantage. Some even suggest that the dominance in the information processing field will lead to world leadership in the 21st century. As envisioned by the Japanese, by the early 1990 s almost everybody in developed countries will be users of computers, and new applications of computers will touch almost every aspect of life.

Computer applications range from industrial and office automation to education and health care. In industrial automation, computers are useful for inventory management, product-cost estimation, and designing automated factories. Computers permit integration, storage, and retrieval of all types of office information (images, graphics, and text). Different offices in a large organization can be connected via local-area networks to support electronic mail. In the field of education, computer-aided instruction permits students to learn new material at their own pace. Current systems are beginning to make use of natural language and voice commands to control the flow of information. The health care industry is using computers for automatic analysis of laboratory tests, for monitoring and management of hospital patients, and for developing sensory prostheses.

Most developing countries do realize that computers are here to stay. They are beginning to develop an infrastructure to spread the use of computers. However, the adaptation of computers in developing countries has been slow, and is currently limited to business data processing (billing, inventory control, etc.) and scientific computing. It is safe to say that a typical developing country is 10-20 years behind a developed country in the utilization of computers. This can be attributed to several reasons: shortage of capital, too few trained computer scientists and, lack of appreciation for the benefits and costs of computing.

Recent concerns over productivity, cost, precision, and reliability in the industrial environment of the developed countries have led to an upsurge in the activities associated with the design of automated factories [1]. The black box representation of an automated factory has raw materials entering at the input port and the finished products exiting at the output port with the least amount of human intervention. To design such factories, we need capabilities in computer-aided design (CAD), computer-aided manufacturing (CAM), robotics, pattern recognition, and image processing (often called machine vision). While the automated factories are still mostly at the design stage, the use of above techniques has resulted in a better quality of finished products at reduced prices in a number of manufacturing environments.

In an industrial environment, pattern recognition and image processing techniques are useful to sense the position, orientation, and identity of parts. This is useful for both inspection and assembly tasks. Examples include automated visual inspection of transistor ignition assemblies and printed circuit boards, and measurement of hole diameter in metal castings connectors for the telephone system. The traditional inspection methods in a factory usually examine only a small percentage of the parts or components in a lot due to high labor cost and the large inspection time. Automated inspection techniques permit examinations of all the parts at a reasonable cost. Consider the problem of inspecting glass bottles for glass and metal fragments in a food industry where the production lines operate at a speed of 12 bottles per second. Manual inspection cannot match this speed. Further, because of the repetitive nature of what they do, human inspectors cannot maintain the 100 percent efficiency level which an automated system can. This is why automated vision inspection is being adopted in microelectronics manufacturing. Vision systems are also being integrated into computer-aided design manufacturing systems.

Pattern recognition and image processing have numerous other applications besides automatic inspection and assembly in manufacturing environments, which will benefit scientists and engineers in developing countries. For example, x-ray diffraction and electron microscopy used in chemistry, physics, metallurgy, microbiology, biochemistry, geology, and pesticide research generate a rich source of photographic images which are currently analyzed manually. Manual optical resolution procedures have counterparts in automated image processing. Automatic recognition and counting of particulate matter describable by size and shape parameters—such as cells, bacteria, viruses, and powders—will assist other investigators. Scientists working in agriculture, forestry, hydrology, mineral exploration, and land use management share a common goal—they all need information about the resources of the earth. Image processing and pattern recognition techniques along with remote sensing have revolutionized the acquisition and processing of satellite data to provide information about soil conditions, types of vegetation, thermal characteristics, weather patterns, and overall topography of a terrain. Even the food industry, which is very sensitive to the problem of contaminants in its products, is adopting automated visual inspection systems.

It is important that the developing countries understand the implication of these advanced technologies with respect to their own industries and production problems. They must at the least develop capabilities to continuously monitor, identify the relevance, and assess the utilization of these new technologies. In its absence, the scientific and technological gap between the developed and developing countries will further widen, and the developed countries will strengthen their competitive edge in the open market due to the lower cost of the manufactured goods using the new technologies. A number of recent UNIDO reports highlight this point in connection with computer-aided design and software production [2,3,4].

This report summarizes the pattern recognition and image processing techniques and its current use and significance in different scientific, engineering, agricultural, and manufacturing disciplines. Developing countries can derive substantial benefit by adopting pattern recognition and image processing technology. Information about hardware, software, and technical literature pertinent to developing a recognition system is provided along with initiatives to be taken by a developing country.

References

1. J. Hollingum, Machine Vision: The Eyes of Automation, Springer Verlag, 1984.

2. R. Narasimhan, 'Guidelines for Software Development in Developing Countries', UNIDO report UNIDO/IS.439, Feb. 10, 1984.

3. H. Kopetz, 'Guidelines for Software Production in Developing Countries', UNIDO report UNIDO/IS.440, Feb. 10, 1984.

4. 'The Impact of Electronics on the International Economic Setting - The Case of Computer-Aided Design', UNIDO/IS.297, March 19, 1982.

II. WHAT IS PATTERN RECOGNITION AND IMAGE PROCESSING?

The main objective of pattern recognition and image processing is to either aid a human decision maker or to automate the decision making process. Both these techniques require data acquisition via some sensors, data representation, and data compression. The data is usually in the form of pictures (or images) or signals. While the images or signals can be interpreted, analyzed or classified by trained human operators, computer processing often leads to more reliable and faster analysis. This can be illustrated by the following example:

Continuous casting of steel is a common technology existing in virtually all developing countries. The steel slabs need to be examined for surface imperfections (longitudinal face and corner cracks, transverse face cracks, scum patch, etc.) before any further processing. The current inspection technique requires the hot steel slab to be cooled to the ambient temperature and then examined by an inspector. If there are no imperfections, the steel slab has to be heated again for sending it to, say, a hot strip rolling mill. The Honeywell Systems and Research Center, U.S.A. has developed an automated steel slab inspection system which eliminates this intermediate step of cooling and reheating, resulting in significant energy savings. This real-time inspection system utilizes various pattern recognition and image processing algorithms as demonstrated in the following block diagram [1]:

Moving Hot slab → Slab image by CCD camera → Object Segmentation Labeling, and Feature Extraction → Component classifier → Multi component combiner → Identified Imperfections

There are numerous other industrial inspection systems [2-5] whose performance is superior to that of human inspector, and more importantly, these systems can eliminate the need of human presence in potentially hazardous invironments.

2.1 Image Processing

Image processing techniques augment human visual analysis of images. Modern scientific equipment in many disciplines has automated the image data gathering process; photography, television, electron scanning microscopes, x-rays, multi-spectral scanners, bubble chambers, and ultrasound devices all produce images which can be subjected to computerized or digital image processing. However, the time saved in gathering the data may frequently be lost in routine interpretation. The collection of computer techniques known as 'image processing' allows automation of much of the routine analysis of pictorial data, thus freeing individual scientists for more challenging and productive aspects of their work. In some cases computer programs can recognize features or objects in the pictures (this is called pattern recognition), or enhance a poor image. The best known example of computer image enhancement lies in the satellite photographs of the moon and Mars; in those cases the computer-enhanced images are a significant improvement over the originally transmitted data. Long-range goals of image processing include enabling computers to do routine evaluation of complex pictorial data and improving the information content of pictures which must be examined by people.

2.2 Pattern Recognition

Pattern recognition techniques assign a physical object or an event to one of several prespecified categories. Thus, a pattern recognition system can be viewed as an automatic decision rule; it transforms measurements on a pattern into class assignments. The patterns themselves could range from agricultural fields in a remotely sensed image from Landsat satellites to a speech waveform or utterance from an individual; the associated recognition or classification problems are to label an agricultural field as wheat or non-wheat and to identify the spoken word. Thus, the patterns can be 1-dimensional (amplitude versus time waveform), 2-dimensional (images where light intensity at each point (x,y) is available), or 3-dimensional (objects in robotics application). Patterns are recognized based on some features or measurements made on the pattern. In remote sensing applications, the features are usually the reflected energies in several wavelength bands of the electromagnetic spectrum, whereas in word identification or speech recognition problems, one common set of features are the LPC's (Linear Predictive Coefficients) computed from the speech waveform.

The recognition problem is difficult because various sources of noise distort the patterns and there exists a substantial amount of variability among the patterns belonging to the same category. The end result of this is that the features or measurements on patterns belonging to the same category are different. Hopefully, this difference will be small compared to the differences between patterns belonging to different categories. For example, the speech waveforms associated with the utterance of the word 'Computer' by different speakers will be different. Even the same speaker uttering the same word at different times will generate slightly different speech waveforms. Similarly, different wheat fields will generate different spectral signatures depending on the soil conditions, atmospheric conditions, and the stage of plant growth. A successful recognition system must be able to handle these pattern variabilities.

2.3 Applications

Recent concerns over productivity and quality control in the developed countries have led to an upsurge in pattern recognition and image processing activity. Virtually every segment of the industry in developed countries has either employed or investigated the use of pattern recognition and image processing for automated assembly and inspection. Most robots currently used in industrial plants are preprogrammed and do not have any 'intelligence'. The new generation of robots have a vision system so that they can 'see' things and recognize them. The subject of pattern recognition and image processing deals with this process of 'seeing' and 'recognition'. This ability allows a system to inspect a part and determine if it is normal or defective. An example of such a system is CONSIGHT [6], which has been developed for locating and inspecting integrated circuit chips for their automatic alignment during various manufacturing processes at the General Motors Delco Electronics Division. This system is currently in production use and achieves the required production accuracy within acceptable costs. Other industrial applications utilizing the pattern recognition and image processing methodology can be labeled as:

(i) Parts Verification

(ii) Process Control

(iii) Materials Handling

(iv) Sorting

(v) Proper location/Orientation of Parts

(vi) Assembly Automation

The applications of pattern recognition and image processing are not just restricted to manufacturing processes or inspection techniques. In fact, these applications are very recent compared to remote sensing and medical applications. Satellite images have been used for determining crop yield, monitoring environmental factors, land use planning, oil and mineral exploration, water pollution, etc. These applications are as important to a developing country as the industrial applications. Various applications of pattern recognition and image processing can be categorized under the following major areas:

1. Document Processing	Recognition of printed or written characters. E.g. reading machine for the blind, sorting objects using bar codes, automatic input of printed text to word processing systems.
2. Industrial Automation	inspection and assembly of complex objects. E.g. printed-circuit board inspection, high-speed inspection of small machined parts, inspection of fuel gauges and speedometers
3. Remote Sensing	Earth observation by sensors aboard satellite or aircrafts. E.g. Forecasting crop yield, land use planning, environmental monitoring, meteorology, mineral exploration, topographic mapping.
4. Medicine & Biology	Processing various medical images and signals

	E.g. blood cell counting, presence of tumors in chest x-ray, tissue characterization using ultrasound, chromosome image analysis, interpretation of electrocardiogram.
5. Nondestructive Testing	Detection of flaws and defects in materials and components before the failure occurs. E.g. eddy current inspection for cracks in heat exchanger tubes, ultrasound for cracks in pressure vessels, radiography for voids and inclusions in pipeline welds.
6. Personal Identification	Restricted entry in secure installations. E.g. speech recognition, fingerprint identification, face recognitions.
7. Scientific Applications	E.g. interpretation of seismic waves for earthquake prediction, identification of tracks in bubble chamber photographs, analysis of molecular composition from electron microscope images.
8. Agriculture Application	E.g. Guidance of equipment, inspections of products, sorting and packing of products.

The recent book by Fu [7] gives a detailed account of many of these applications. While the images or signals associated with any of the above applications can be examined and interpreted by a trained human operator, automating the interpretation process using pattern recognition and image processing algorithms implemented on a digital computer has almost always led to better accuracy in reduced time and lower cost.

References

1. B. R. Suresh, et.al., 'Real-time Automated Visual Inspection System for Hot Steel Slabs', IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 5, Nov. 1983.

2. Special Issue on Image Processing, Proceedings of IEEE, May 1981.

3. Special Issue on Industrial Applications of Machine Vision, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 5, Nov. 1983.

4. D. L. Hudson and J. E. Tromblay, 'Industrial Uses for Machine Vision', Computers in Mechanical Engineering, April 1983, pp. 18-24.

5. Special Issue on Machine Perception for Industrial Applications, Computer, Vol. 13, May 1980.

6. M. R. Ward, et. al., 'CONSIGHT: A Practical Vision-Based Robot Guidance System', Proc. 9th Int'l Symp. Industrial Robots', Society of Manufacturing Engineers and Robot Institute of America, Washington, D.C., Mar. 1979, pp. 195-212.

7. K. S. Fu (Ed.), Applications of Pattern Recognition, CRC Press, 1982.

III. REVIEW OF IMAGE PROCESSING TECHNIQUES

Image processing turns images or pictures from a camera into useable data using digital computers. These images could be obtained from different sensors in a variety of environments. For example, multispectral scanners aboard Landsat satellites provide images of the earth's surface, x-ray sources give images of human bodies, and laser scanners provide range images in industrial environments [1]. A typical image processing application consists of three fundamental operations:

(i) Image acquisition,

(ii) processing and quantitatively assessing features in an image,

(iii) making decisions based on the available features.

Therefore, the field of image processing must deal with the acquisition, storage, enhancement, restoration, and interpretation of images or pictures. The phrases picture processing, machine vision, computer vision, scene analysis, and image understanding also refer to this activity of processing and interpreting images using digital computers.

There are two main reasons for digital processing of images: (i) To improve image quality to facilitate human interpretation and (ii) To automatically determine the presence/absence of some objects in an image. Consider the field of remote sensing in which images are gathered by satellites and then transmitted to earth [2,3]. These images are studied by trained photo-interpreters to derive information regarding land-use categories, mineral deposits, wet lands, crop disease, forest inventory, etc. These images need to be corrected for various distortions (e.g. earth's curvature) and brought into geographic alignment. Various kinds of noise present in these images (e.g. shading effects, scattered light from the atmosphere) need to be removed. Digital image processing allows efficient processing of these images to get rid of various forms of degradation and noise. This makes the job of a human photo interpreter much easier. However, even a thoroughly trained photo interpreter cannot process large numbers of images in a consistent and reliable manner. Therefore the remotely sensed images are also classified automatically into various categories depending on the particular application. Automatic decision-making techniques can also easily incorporate ancillary information (e.g., topographic information, images taken at different times).

Figure 3.1 shows the block diagram for a typical image processing application. The image of the scene which is to be interpreted is acquired by sensors. The analog or continuous image (signal) is digitized so that the image can be entered into a digital computer and the intensities in the image are quantized to a finite number of gray levels. This image is enhanced to remove or reduce noise and degradation. At the third stage, image points or pixels with similar properties are grouped to form regions or segments. Regions are described by properties like shape, area, texture, color, etc. Relationships (adjacency, left, right, surround) among the various scene components or regions are then used to recognize the objects and interpret the scene. At each stage of processing, certain assumptions are made which do not arise from the sensed data itself. For example, in the enhancement stage, apriori assumptions about the imaging noise are used. At the segmentation stage, assumptions are made about the smooth variation of object boundaries and uniform intensity of object interiors. These various stages in a typical image processing system are described below [4-14]. A collection of key papers on various aspects of image processing can be found in [15,16].

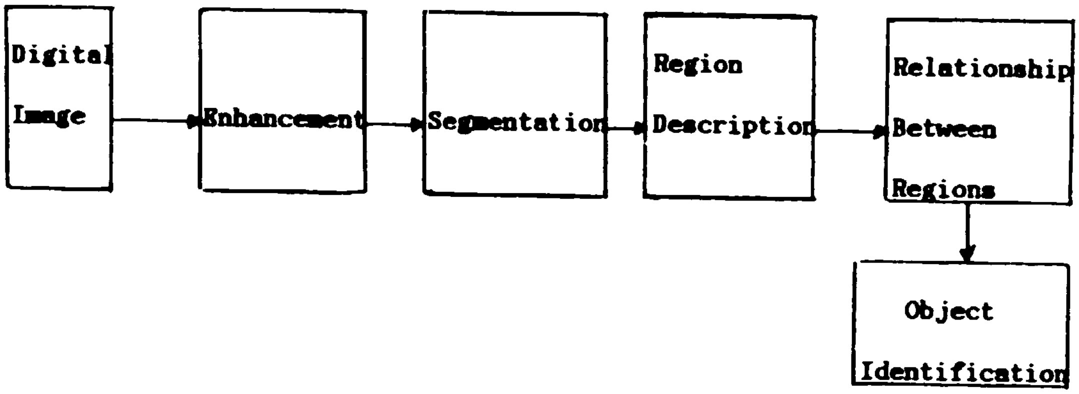


Figure 3.1 Block Diagram of a Typical Image Interpretation System

3.1 Image Acquisition

We are all familiar with photographic cameras which give a picture of the environment by recording intensity variations by differential exposure of light-sensitive chemicals. In many image processing applications, we need the image only in the form of signals rather than a hard-copy of the scene. Three common methods of electronically recording an image are briefly described below. Most of our discussion here is limited to the acquisition of two-dimensional images.

(i) Vidicon tube: Their operation is similar to a television camera. The image of the scene is focused on a light sensitive plate which is scanned by an electron beam. At any point in time, the light intensity at a particular point in the image is proportional to the current or voltage output from the tube.

(ii) Solid-state camera: The image is focused on an array of discrete sensing elements, each one of which builds up a charge proportional to the intensity of light falling on it.

(iii) Multispectral scanner: These scanners measure the components of the incoming light in several bands of the electromagnetic spectrum. For example, in Landsat 2 satellite, four bands (2 in the visible region and 2 in the infrared region) were used.

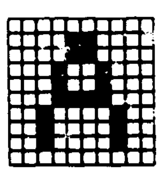
Most image sensors are passive, i.e. they record emitted energy. However, in many applications active sensing is necessary. Examples are laser scanners for depth measurement, and x-ray and ultrasonic sensors for biomedical applications. For industrial applications, particularly in robotics, solid state cameras are more appropriate: They are light weight, rugged, and use little power. With advances in hardware technology, solid state cameras are available with higher resolution and lower cost. A good survey on image sensors can be found in [17].

3.2 Image Digitization

Image Digitization refers to the process of converting an image or picture into a set of numbers so that it can be entered into a digital computer. There are two important considerations in converting an image into digital form. The first deals with spatial resolution which is measured to be the number of pixels (picture elements) which can be sensed in the image. With higher resolution, we can detect smaller objects in the image. Thus resolution refers to the spatial detail captured during the sampling process. The process of image digitization then refers to converting an image into an NxN array or grid of pixels. Typical values of

N are 256 and 512, although in some applications 1024x1024 arrays are used. The physical size of the pixel depends on the camera optics and the distance of the camera from the object (standoff). The Landsat 2 satellite gives images with a pixel size of 80 meters by 80 meters. Some newer satellites (e.g. SPOT) will have pixels of size 10 meters by 10 meters. In biomedical applications, pixel size is measured in micrometers.

For digital processing, the intensity at a point in the image also needs to be quantized. The gray level resolution refers to the range of densities or reflectances to which the image scanner responds and how finely it divides the range. Earlier industrial vision systems used only binary images, for which intensity values are either 0 or 1 ('black' or 'white'). Figure 3.2 (a) shows the binary digitized image of character A on a 10x10 grid. This binary image can be viewed as a matrix of 0's and 1's where a 0 represents the background and a 1 represents the object (character A). Binary images are easy to process and take little storage space. However, to process complex images we need to utilize subtle changes in the reflectance in a scene. Most digitizers and scanners will code each intensity value into 256 possible intensity encodings. For industrial tasks where illumination can be controlled, 16 levels of intensity information are enough.



(a)

```
0000000000
0000110000
0000110000
0001001000
0001001000
0001111000
0010000100
0010000100
0000000000
```

(b)

Figure 3.2 (a) Binary Image of character A on a 10 x 10 grid

(b) Representation of binary image as a matrix of 0's and 1's

Once an image has been digitized and quantized, it is represented as a matrix of integer numbers which can be entered into a computer. These numbers can then be manipulated or processed to extract relevant information about the image.

3.3 Image Enhancement

A digital image usually contains several kinds of noise and degradation which make it difficult to interpret the image. Some sources of distortion are: sensor noise, motion blur, lens aberrations, noise introduced during the transmission of the image, non-uniform illumination, and specular reflection. Image enhancement and restoration refers to a collection of techniques to improve the quality of a noisy and degraded image, either for human inspection or subsequent automatic analysis.

Digital image restoration typically involves modelling the image formation process and the degradation process, and eliminating the degradation through a convolution or filtering operation [7]. For example, if we knew that during the imaging process the camera or the object moved, then the blurred image can be brought into focus by digital restoration techniques. The parameters of the degradation process may or may not be known apriori. Image enhancement utilizes various heuristics to improve the quality of the image. There is no 'optimal' enhancement technique and the choice of a particular technique is data dependent. Enhancement operations include histogram modification for increasing contrast, smoothing or low pass filtering, sharpening or high pass filtering, and geometric correction. The smoothing operation performs a local averaging of image intensity values to remove noise whereas the sharpening operation highlights the local discontinuities in image intensities to facilitate extracting the boundaries of the objects present in the image.

3.4 Image Segmentation

Once an image has been acquired and processed to improve its quality, one usually wants to automatically isolate regions of interest in the image. For example, given a digital image obtained from a Landsat satellite, the desired output is a map of the region showing specific types of terrain features (e.g., forest, urban areas, water bodies etc.). The output requires the location and identification of the desired terrain types. Image segmentation refers to a partitioning of the image into regions such that pixels belonging to a region are more similar to each other than pixels belonging to different regions [18]. What criterion of similarity should be used in grouping the pixels? In the context of remote sensing, similarity between two pixels is measured in terms of the gray values in several spectral bands.

In the simplest case for image segmentation we have a dark object on a light background or vice versa. The pixels belonging to the object can be isolated from the background pixels by a simple thresholding operation: pixels with intensity or gray levels $\geq$ t belong to the object whereas the remaining pixels belong to the background. However, in realistic situations, more powerful segmentation techniques based either on edge detection or region extraction are needed. The edge detection techniques essentially try to find the boundary between the object and the background and then determine the interior of the object. Region extraction techniques begin with some pixels which are known to belong to the object and then start merging these pixels with other neighboring pixels with similar properties until the entire object is extracted. Another phrase for region extraction technique is clustering which has been successfully used in several remote sensing applications [19].

3.5 Region Description

Regions or segments in an image can be distinguished and classified based on their color, shape, and texture. A useful description of an object should be independent of its size (viewing distance), orientation, and position. Thus color, shape, and texture properties of a region are used to derive features needed by a classifier or a pattern recognition system (see section 4). The shape of a region can be described simply by its boundary. The boundary can be encoded by looking at local properties. Freeman chain codes and Fourier descriptors are two well-known boundary encoding approaches. A well-known internal representation of a region is the set of invariant moments, which have been used extensively in target classification. Details of various shape representation methods can be found in [20].

Texture is a property related to the spatial distribution of gray levels or the overall pattern of gray level changes in a region. Quite often, a region in an image can be distinguished from the others based on the differences in the texture. For example, in remote sensing applications different land use categories (orchard, residential, water, forest, etc.) have different textures. Two textured regions may not be separated by an edge or a sharp transition in the gray levels. Thus texture is useful in both image segmentation and classification. Several statistical, structural, and modelling approaches to texture have been proposed in the literature [21].

3.6 Image Interpretation

Image interpretation or understanding refers to a symbolic description of the image. This requires not only the identification of individual objects present in the image but also the spatial relationship among the objects. In order to achieve this, the recognition system must have access to the same domain-specific knowledge which the human interpreter uses. The difference between image processing and image understanding can be illustrated by the following example. Given an aerial image of a scene, an image processing system would label or classify each pixel into one of several prespecified categories. An image understanding system, on the other hand, is capable of generating a narrative from the image which reads something like:

'The image contains two major highways, a river in the southeast corner and there are two bridges on this river.'

In order to generate this narrative, the system must know the concept of a highway, a bridge and a river. The difficulty lies in representing this knowledge so that it can be used efficiently in the analysis of an image.

3.7 Recent Developments in Image Processing

Image processing has been successfully used in recognizing 2-dimensional objects in a variety of applications. However, to interpret complex images and to recognize 3-dimensional objects (necessary in robotics applications), more powerful techniques and additional information about the domain need to be incorporated. Therefore, current research efforts are concentrated in the following areas.

3.7.1 Intrinsic Images

Attention has shifted from the application domain of a vision system to restrictions on visual abilities. Substantial progress has been made in the interpretation of surface contours, surface orientation from texture, and the computation of motion [22].

3.7.2 Recognition of 3-dimensional Objects

Much recent work has been done to extract 3-dimensional information about a scene either directly by laser scanner or indirectly by, say, binocular stereo, and to utilize it to duplicate human performance in navigating through space and manipulating 3-dimensional objects [23].

3.7.3 Expert Vision Systems

It is now generally agreed that to interpret a complex image, the image processing system must utilize some knowledge about the image domain which a human expert uses. Several such systems are currently under development [24,25].

3.8 Summary

Image processing is useful for improving the quality of an image to aid human interpretation, and to make automatic decisions about the objects present in the image. One of the first applications of digital image processing was to remove the noise and degradation present in the images of Mars and the moon transmitted to earth by satellite. Since then, image processing has been routinely used to process and interpret images in a variety of fields such as remote sensing, astronomy, medicine, chemistry, physics, and industrial inspection. Current vitality in this area is due to its important role in robotics where image processing is needed to analyze 3-dimensional indoor and outdoor scenes. It is already revolutionizing the manufacturing and inspection processes in most of the developed countries.

References

1. A. Rosenfeld (ed.), Digital Picture Analysis, Springer Verlag, 1976.

2. J. Lintz and D. S. Simonett, Remote Sensing of Environment, Addison-Wesley, 1976.

3. R. Hord, Digital Image Processing for Remotely Sensed Data, Academic Press, 1982.

4. R. Gonzalez and P. Wintz, Digital Image Processing, Addison-Wesley, 1977.

5. A. Rosenfeld and A. Kak, Digital Picture Processing, Vol. I and II, Academic Press, 1976 and 1980.

6. W. Pratt, Digital Image Processing, Wiley, 1978.

7. H. Andrews and B. Hunt, Digital Image Restoration, Prentice Hall, 1977.

8. D. Ballard and C. Brown, Computer Vision, Prentice Hall, 1982.

9. R. Nevatia, Machine Pereception, Prentice Hall, 1982.

10. T. Pavlidis, Algorithms for Graphics and Image Processing, Computer Science Press, 1982.

11. E. Hall, Computer Image Processing and Recognition, Academic Press, 1979.

12. O. D. Faugeras (ed.), Fundamentals in Computer Vision, Cambridge University Press, 1983.

13. S. Tanimoto and A. Klinger (eds.), Structured Computer Vision, Academic Press, 1980.

14. A. R. Hanson and E. M. Riseman (eds.), Computer Vision Systems, Academic Press, 1978.

15. J. K. Aggarwal et. al. (eds.), Computer Methods in Image Analysis, IEEE Press, 1977.

16. H. C. Andrews (eds.), Selected Papers in Digital Image Processing, IEEE Press, 1978.

17. G. Nagy, 'Optical Scanning Digitizers', Computer, May 1983, pp. 13-23.

18. K. S. Fu and J. K. Mui, 'A Survey on Image Segmentation', Pattern Recognition, Vol. 13, 1981, pp. 3-16.

19. P. M. Narendra and M. Goldberg, 'A Non-parametric Clustering Scheme for Landsat', Pattern Recognition, Vol. 9, 1977, pp. 207-215.

20. T. Pavlidis, Structural Pattern Recognition, Springer Verlag, 1977.

21. R. M. Haralick,'Statistical and Structural Approaches to Texture', Proc. IEEE, Vol. 67, 1979, pp. 786-804.

22. D. Marr, Vision, Freeman, 1982.

23. E. Hall et. al., 'Measuring Curved Surfaces for Robot Vision', Computer, Vol. 15, Dec. 1982, pp. 42-54.

IV. REVIEW OF PATTERN RECOGNITION TECHNIQUES

Pattern recognition techniques are used to describe and classify measurements or objects. It is also referred to as automatic decision making. The basic idea is that patterns with some similar properties or attributes should be grouped into a class or category. There are essentially two basic paradigms to classify a pattern into one of C different categories or pattern classes.

4.1 Geometric or statistical Approach

A pattern is represented in terms of N features or properties. Thus a pattern can be viewed as a point in this N-dimensional space. If the choice of features is good then pattern vectors belonging to different categories will occupy different regions of this feature space. Given certain sample patterns from each pattern class (training samples), the objective is to establish decision boundaries in the feature space to separate patterns belonging to different classes. A large number of books [1-7] cover this approach.

For example, a time-varying waveform can be represented by peak amplitude, number of zero crossings, rise time, and moments. One could also obtain some additional features by transforming the time-varying waveform to the frequency domain. These features include maximum frequency response, power spectrum parameters, and energy in selected bandwidths. These features collectively form a signature for the waveform. One expects that the signature for, say, an ultrasound waveform corresponding to a "no-defect" situation would be different from that for a "defect" situation.

The objective of an automated recognition system is to assign a given signature to either the defective or non-defective category.

The choice of features is data dependent but is crucial to the performance of the recognition system. Defining appropriate features requires interaction with a person who is an expert in the application area. Features needed to recognize a speaker from a speech waveform are different from those needed to determine the presence of a crack in a pressure vessel from ultrasound return waveforms. Different approaches to the well-known problem of character recognition essentially differ in the kinds of features used [34]. However, most commercial systems for character recognition use a set of binary masks for each character.

Consider the problem of designing a fruit-sorter system [8]. Cherries, apples, lemons, and grapefruits are to be packaged in different boxes as they come on a conveyor belt. These four types of fruits or pattern classes can be easily separated in the redness-diameter feature space as shown in Figure 4.1. The decision boundaries are also displayed in Figure 4.1. These boundaries are estimated from the training samples.

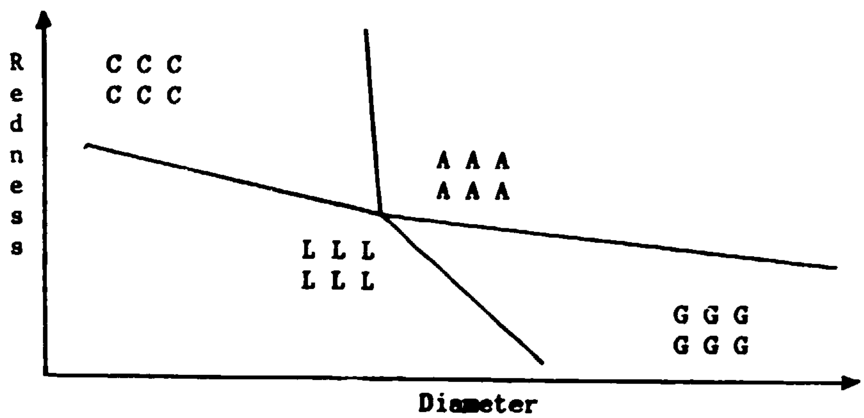


Figure 4.1 Representation of fruits in redness-diameter ; L: Lemon, C: Cherries, A: Apples, G: Grapefruits

The block diagram of a general pattern recognition system is shown in Figure 4.2. The input to the feature extractor is the data (usually a 1-dimensional or a 2-dimensional digitized signal measured by sensors: x-ray, TV camera, ultrasound).

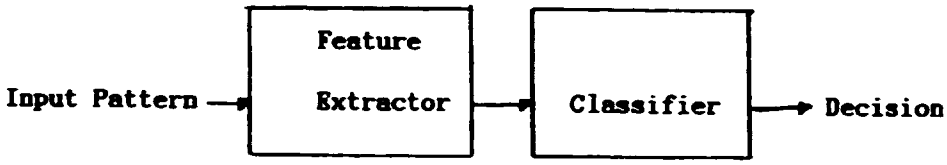


Figure 4.2: A Pattern Classification System

4.2 Structural or Syntactic Approach

In many recognition problems involving complex patterns, the number of features required to establish a reasonable decision boundary is very large. In such situations it is more appropriate to view a pattern as being composed of simple subpatterns [9]. A subpattern itself could be built from yet simpler sub-patterns. The simplest subpatterns to be recognized are called primitives and the given complex pattern is represented in terms of the interrelationships among these primitives. The main advantage of the structural approach over the geometric approach is that, in addition to classification, it also provides a description of the pattern. The description takes the form of how the given pattern is constructed from the primitives. This paradigm has been used in situations where the patterns have a definite structure which can be captured in terms of a set of rules (e.g., EKG waveforms, textured images, shape analysis of contours).

In syntactic pattern recognition [10-12], an analogy is drawn between the structure of patterns and the syntax of a language. The patterns belonging to a category w_i are viewed as sentences belonging to a language L_i. The sentences are generated according to a grammar G_i and the primitives are viewed as the alphabet of the language. For example, in the classification of EKG waveforms, one can define a grammar for abnormal waveforms and one for the normal waveforms in terms of the P-Q-R-S complex. A given waveform is classified as normal or abnormal based on which of the two grammars can correctly parse the pattern. Thus a large set of complex patterns can be described by using a small number of primitives and grammatical rules. The grammar for each pattern class must be inferred from the available training samples.

The block diagram of a syntactic recognition system is shown in Figure 4.3.

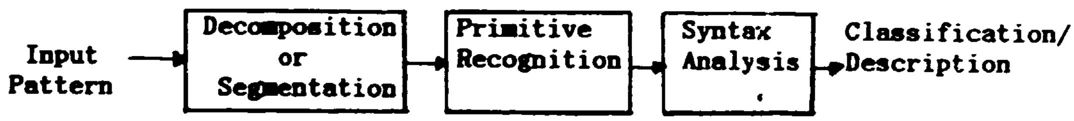


Figure 4.3: Block Diagram of Syntactic Approach

The challenging problems associated with the syntactic approach are: (i) Find a grammatical structure that is an efficient and effective way to represent the pattern classes (ii) Infer a grammar from a finite sample of patterns, and (iii) Parse strings of primitives. These concepts can be illustrated by the following example. Four types of chromosomes shown in Figure 4.4 need to be distinguished based on their images. The primitives which are used to represent the boundaries of the chromosomes constitute the terminals for the chromosome grammars [12]. These terminals are:

$$V_T = \{ a , b , c , d , e \}.$$

The intermediate subpatterns (more complex than primitives but less complex than the chromosome image) are called the non-terminals. The set of nonterminals used and their interpretation are given below:

$$V_N = \{A, B, C, D, E, F\}.$$

A:	arm pair	D:	right pair
B:	bottom	E:	left pair
C:	side	F:	arm

The production rules for the submedian chromosome grammar are given below.

$S \longrightarrow AA$

$A \longrightarrow cA/Ac/FD/EF$

$E \longrightarrow Fc$

$D \longrightarrow cF$

$B \longrightarrow bB/Bb/e/bC/cb/b$

$C \longrightarrow d$

$F \longrightarrow BF/FB/\varepsilon$

These rules can be used to determine whether a given chromosome image belongs to the submedian category. Similar production rules exist for the other three types of chromosomes.

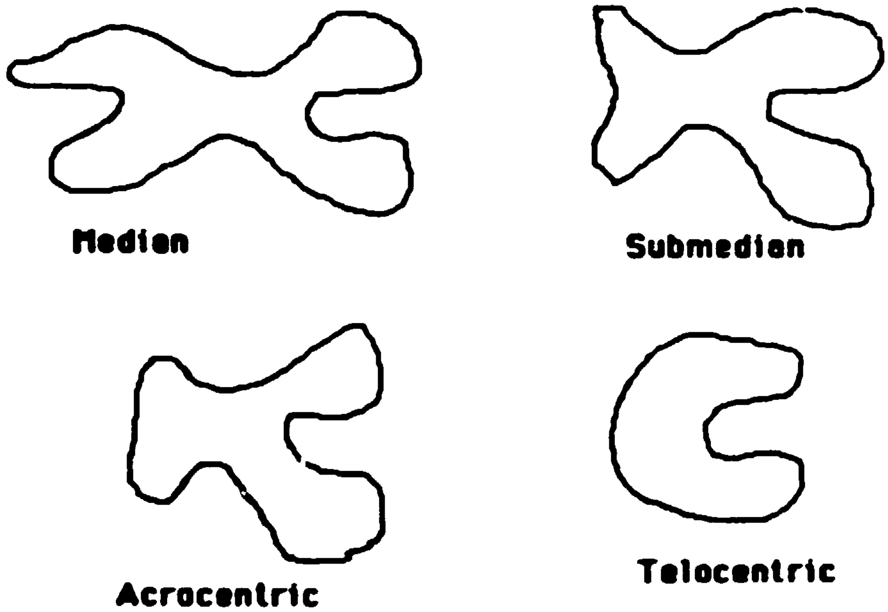


Figure 4.4 Four Types of Chromosomes

Syntactic pattern recognition is intuitively appealing. However, implementation of this approach leads to many difficulties which primarily have to do with the segmentation of noisy patterns to detect the primitives and inference of the grammar. While there are some problems with the statistical approach also, it is better understood and relies on more firmly established elements of statistical decision theory. Perhaps this is why most commercial recognition systems utilize decision - theoretic approaches [13]. Therefore, we will now concentrate on the statistical approach.

4.3 Statistical Pattern Recognition

The paradigm of statistical pattern recognition can be summarized as follows. A given pattern is to be assigned to one of C categories $w_1, w_2, \ldots, w_C$ based on its feature values $(x_1, x_2, \ldots, x_N)$. The features are assumed to have a density function conditioned on the pattern class. Thus a pattern vector $\underline{x}$ belonging to class w is viewed as an observation drawn randomly from the class-conditional density $p(\underline{x}|w_i)$. Well known concepts from statistical decision theory are utilized to establish decision boundaries between the pattern classes. The decision rule involves the class-conditional densities $p(\underline{x}|w_i)$ and the apriori probabilities $p(w_i)$, $i = 1,2,\ldots, C$. The class-conditional densities are never known in practice and must be "learned" from the training samples. Various strategies utilized to design a classifier in statistical pattern recognition depend on what kind of information is available about the class-conditional densities. If it is assumed that the form of the class-conditional densities is known then we have a parametric decision problem. Otherwise, we must either estimate the density function or use some nonparametric decision rule. The estimation of parameters or the density function requires the availability of learning or training samples. Another dichotomy in statistical pattern recognition is that of supervised learning (labeled training samples) versus unsupervised learning (unlabeled training samples). The label on each training pattern represents the category to which that pattern belongs.

These various dichotomies which appear in statistical pattern recognition are shown in the tree structure of Figure 4.5. The classification problems get more difficult as one traverses the tree from top to bottom and left to right. The field of cluster analysis essentially deals with decision making problems in the non-parametric and unsupervised learning mode [14-17].

Many factors need to be considered in the design of a recognition system. The performance of the recognition system critically depends on the choices made.

4.3.1 <u>One-shot Versus Hierarchical (tree) Classifier</u>

In the one-shot classification scheme, distinction between all classes is made in one stage. This may not be the most appropriate scheme, especially if the number of pattern classes is large. An alternative classification structure is a binary tree where the most obvious discriminations are done first, postponing the more subtle distinctions until a later stage [18]. This results in a higher processing speed since the average number of features per node is substantially reduced.

4.3.2 <u>Parametric versus Nonparametric Approach</u>

Most of the popular parametric techniques are optimal if the class-conditional densities happen to be Gaussian. Even though statistics based on Gaussian distributions are robust, it is wise to consider non-parametric techniques, e.g. nearest - neighbor classifier, if the data substantially departs from normality [1].

4.3.3 <u>Dimensionality and Sample Size Relationship</u>

The well-known phenomenon of the curse of dimensionality cautions a system designer to use a limited number of features for a given sample size

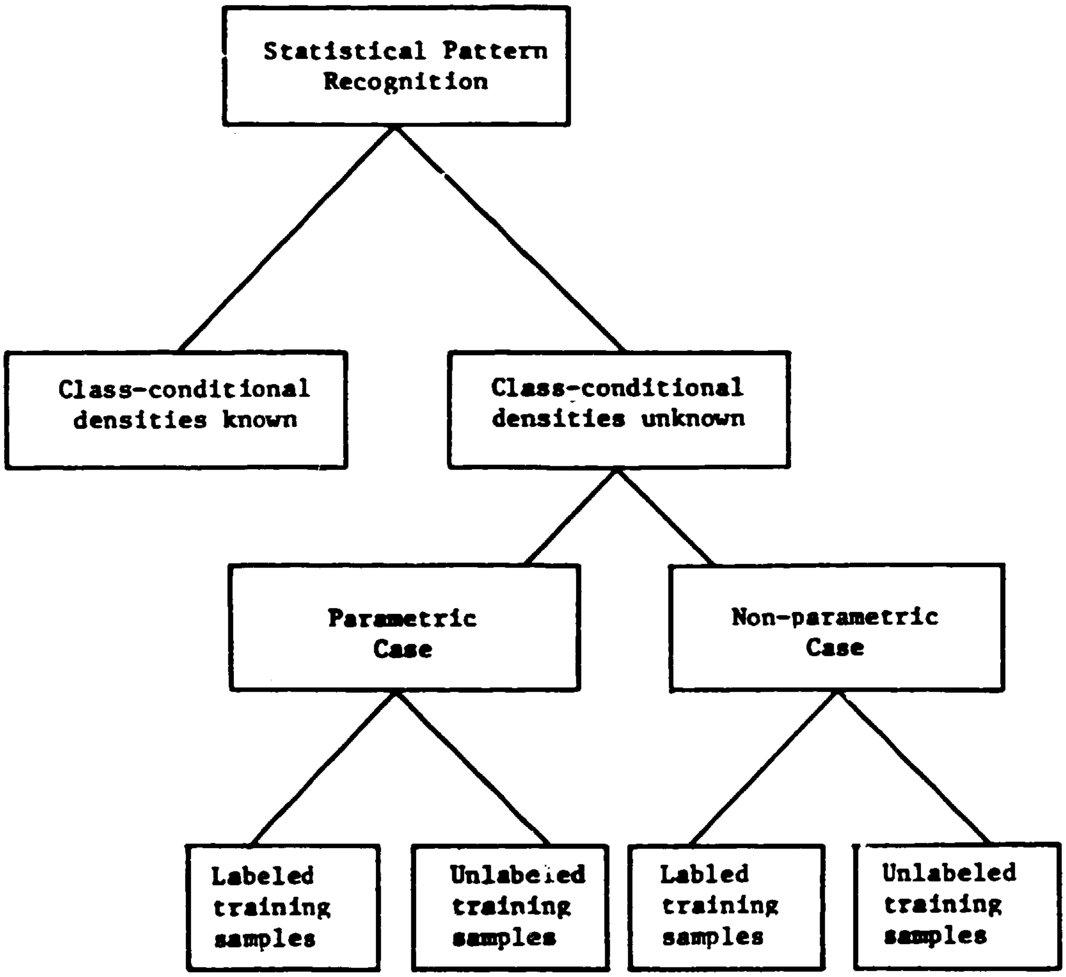


Figure 4.5 Dichotomies in Statistical Pattern Recognition

[20]. It is recommended that the number of training samples per class be at least five to ten times the number of features.

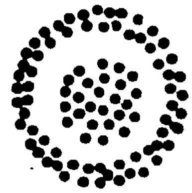
4.3.4. Supervised Versus Unsupervised Learning

In many applications of pattern recognition it is difficult or expensive to label the training samples. Consider, for example, an application in remote sensing. In order to get 'ground truth' information, somebody must visit the physical site associated with the pixel to determine its land-use category. Similar problems exist in nondestructive evaluation. Often the labels are not very reliable. This is especially true in some medical decision-making applications where different experts may give different opinions about, say, an EKG waveform.

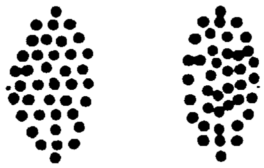
Unsupervised learning refers to situations where training samples are not labelled. This is a difficult problem and, in practice, one often has to rely on techniques of cluster analysis [17] to identify natural groupings (categories) present in the data. Cluster analysis attempts to assess the interaction among patterns by organizing them into groups or clusters. The results of cluster analysis are useful for forming hypotheses about the data, classifying new data, testing for the homogeneity of the data, and for compressing the data.

A large number (hundreds) of clustering algorithms have been developed and implemented. But there is no 'best' clustering algorithm - the performance of a clustering algorithm is data dependent. Figure 4.6 shows some of the possibilities for the shapes and sizes of clusters in two dimensions. Different clustering algorithms will partition these data sets differently [22]. This points to the difficulty of interpreting the results of a clustering algorithm, particularly in high dimensions. Some heuristics are available to assess cluster tendency and cluster valiity [23].

1(a) Compact and well-separated clusters

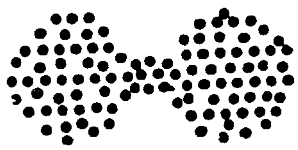


1(g) Concentric clusters

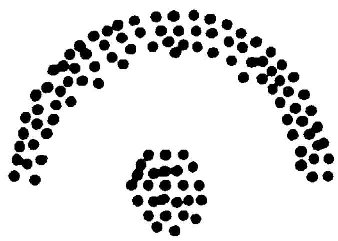


1(b) Elongated Clusters

1(h) Clusters within clusters



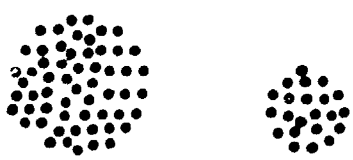
1(c) Touching clusters



1(d) Linear non separable
clusters

Figure 4.6 Data sets exhibiting different clustering structures

1(e) Non-compact clusters



1(f) Unequal cluster size

4.3.5 Feature Selection

For computational reasons and cost considerations, one often faces the problem of deciding which of the available features are good for classification [5]. It has been shown that the problem of determining a best subset of features of size M from the N available features requires an exhaustive search over all possible subsets of size M [20]. A recent feature selection technique utilizes the branch and bound method to avoid exhaustive search [21].

4.3.6 Error Estimation

The available samples must be partitioned into training and test sets. The classifier is designed using the training set and evaluated on the samples belonging to the test set. In other words, the classifier is cross-validated. Depending on the number of samples available, the error rate or the performance of the classifier is estimated using the hold-out method or leave-one-out method [5].

4.4 Pattern Recognition Software

There are several general-purpose statistical packages (e.g. SPSS, BMDP) which contain programs for discriminant analysis, feature selection, and clustering, but the choice is very limited. A few software packages are available specifically for cluster analysis. The most well-known cluster analysis package is CLUSTAN [24]. A good interactive pattern classification software package is ISPAHAN [25]. It has been in operation for a number of years and has been utilized successfully in a variety of problems such as the classification of white blood cells. The algorithms implemented in ISPAHAN cover the following areas: scatter plots, supervised classification, unsupervised classification, mapping algorithms, and feature vector transformations. While ISPAHAN is a good software package to get a research group started, substantial additions have to be made to the package (quadratic classifier, feature selection) for any rigorous analysis and interpretation of the data.

4.5 Recent Developments in Pattern Recognition

The fundamentals of the decision theory are now quite well understood. Therefore, recent concerns have been about the complexity of the recognition algorithms, their implementation in VLSI, and incorporating domain-specific knowledge or context in improving the recognition accuracy.

4.5.1 Computational Complexity

In pattern recognition, decisions often have to be made about patterns residing in a spatial region. For example, the decision of whether a given pattern is the Kth nearest neighbor of another given pattern is used frequently in cluster analysis and decision making. However, making such decisions through naive approaches can render an algorithm impractical for moderate sized data sets. Hence there has been considerable development of the subfield of computational geometry, which concerns itself with the implementation and complexity of geometric decisions. Through these efforts, efficient algorithms are now available for computing nearest neighbors, minimal spanning tree, convex hull etc. which are useful both in pattern recognition and image processing.

Computational savings in designing pattern recognition systems can be made by using efficient search techniques or by studying the property of density functions. A recent paper reduces the processing time of a maximum-likelihood classifier by establishing a set of fixed thresholds, which if exceeded by one probability density function, makes it unnecessary to evaluate a competing density function [27].

4.5.2 Use of VLSI systems

Many pattern classification algorithms are inherently time consuming and require large amounts of storage. Special VLSI chips can now be designed and built for a specific algorithm which speed up the operation by taking advantage of pipelining and parallelism [28]. VLSI designs are now available for matrix operation, minimum-distance classification, clustering, string matching, and several graph algorithms. This will have the effect that the designers of a pattern recognition system need no longer discard an algorithm or an approach just because it takes too much time.

Several special computer architectures have also been implemented to solve pattern recognition problems in real-time. One such system, called Massively Parallel Processor (MPP), is being tested at NASA Goddard. Preliminary runs indicate that a 512x512 multispectral image containing 4 channels of LANDSAT D data can be partitioned into 16 clusters using the ISODATA clustering algorithm (100 iterations) in 18 seconds.

4.5.3 Utilizing context and domain-specific knowledge

Our experience with the design of recognition systems indicates that maximum similarity alone is not adequate to recognize patterns. A satisfactory solution to most recognition problems requires the utilization of context and domain-specific knowledge in decision making [29]. The role of context is to resolve ambiguities. It has been successfully used as a preprocessing step in text recognition (regular occurrences of a particular combination of letters qu, ee, est, tion) and remote sensing (no wheat fields in residential areas) [30]. Context is often incorporated through the use of compound decision theory.

Domain-specific knowledge has always played an important role in designing a classifier. The choice of features, and the grouping of pattern classes in forming a decision tree are based on this knowledge. This knowledge is especially important when the difficulty of the classification problem increases (contrast the situation when only the machine printed numerals have to be recognized against the recognition of two thousand hand written chinese characters) [31]. Lately there has been an emphasis on building expert systems tailored to a specific domain. An expert system is generally knowledgeable about the actions it must take to achieve one of its goals.

4.6 Summary

Pattern recognition has emerged as a mature discipline over the past twenty five years. The earliest applications of pattern recognition included character recognition and classification of white blood cells. Since then pattern recognition techniques have been successfully applied to remote sensing, speaker identification, non-destructive testing (x-ray, eddy current, ultrasound), target identification, and various biomedical applications [32,33]. There are several approaches to classification, none of them being 'optimal'. Therefore, it is important to try out various approaches in designing a classification system for a new application.

Current research emphasis in pattern recognition is on designing efficient algorithms, implementing the algorithms on novel computer architectures, and incorporating context and domain-specific knowledge in decision making.

References

1. R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, Wiley, 1973

2. K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, 1972.

3. J. Tou and R. Gonzalez, Pattern Recognition Principles, Addison Wesley, 1974.

4. Y. T. Chien, Interactive Pattern Recognition, Dekker, 1978.

5. P. Devijver and J. Kittler, Statistical Pattern Recognition, Prentice Hall, 1982.

6. E. Patrick, Fundamentals of Pattern Recognition, Prentice Hall, 1972.

7. C. H. Chen, Statistical Pattern Recognition, Hayden, 1973.

8. K. Castleman, Digital Image Processing, Prentice Hall, 1979.

9. T. Pavlidis, Structural Pattern Recognition, Springer Verlag, 1977.

10. K. S. Fu, Syntactic Methods in Pattern Recognition, Academic Press, 1974.

11. R. C. Gonzalez and M. G. Thomason, Syntactic Pattern Recognition, Addison Wesley, 1978.

12. K. S. Fu, Syntactic Pattern Recognition and Applications, Prentice Hall, 1982.

13. B. G. Batchelor (Editor), Pattern Recognition: Ideas in Practice, Plenum, 1978.

14. B. Everitt, Cluster Analysis, Heinemann Educational Books, 1977.

15. M. R. Anderberg, Cluster Analysis for Applications, Academic Press, 1973.

16. J. Hartigan, Clustering Algorithms, Wiley, 1975.

17. R. Dubes and A. K. Jain, 'Clustering Methodology in Exploratory Data Analysis', in Advances in Computers, Vol. 19, M. Yovits (Ed.), Academic Press, 1980.

18. J. K. Mui and K. S. Fu, 'Automated Classification of Nucleated Blood Cells Using Binary Tree Classifier', IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 2, Sept., 1980, pp. 429-443.

19. A. K. Jain and B. Chandrasekaran, 'Dimensionality and Sample Size Con-

sideration in Pattern Recognition Practice', in Handbook of Statistics, Vol. 2, P. R. Krishnaiah and L. N. Kanal (Eds.), North Holland, 1982, pp. 835-855.

20. T. M. Cover and J. M. Van Campenhout, 'On the Possible Orderings in the Measurement Selection Problem', IEEE Trans. System, Man and Cybernetics, Vol. 7, 1977, pp. 657-661.

21. P. M. Narendra and K. Fukunaga, 'A Branch and Bound Algorithm for Feature Subset Selection', IEEE Trans. Computers, Vol. 26, Sept. 1977, pp. 917-922.

22. R. Dubes and A. K. Jain, 'Clustering Techniques: The User's Dilemma', Pattern Recognition, Vol. 8, 1976, pp. 247-260.

23. R. Dubes and A. K. Jain, 'Validity Studies in Clustering Methodology', Pattern Recognition, Vol. 11, 1979, pp. 235-254.

24. D. Wishart,'CLUSTAN 1C User's Manual', Program Library Unit of Edinburgh University, Edinburgh, 1978.

25. E. S. Gelsema, 'ISPAHAN, An Interactive System for Pattern Recognition', Proc. BIOSIGMA Conf., 1976, pp. 469-477.

26. D. T. Lee and F. P. Preparata, 'Computational Geometry: A survey', IEEE Trans. Computers, Vol. 33, 1984, pp. 1072-1101.

27. A. H. Feiveson, 'Classification by Thresholding', IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 5, Jan. 1983, pp. 48-54.

28. K. S. Fu (Editor), VLSI For Pattern Recognition and Image Processing, Springer Verlag, 1984.

29. G. T. Toussaint, 'The Use of Context in Pattern Recognition', Pattern Recognition, Vol. 10, 1978, pp. 189-204.

30. P. H. Swain, et al., 'Contextual Classification of Multispectral Image Data', Pattern Recognition, Vol. 13, 1981, pp. 429-441.

31. M. Nagao, 'Control Strategies in Pattern Analysis', Pattern Recognition, Vol. 17, 1984, pp. 45-56.

32. P. T. Krishnaiah and L. N. Kanal (Editors), Handbook of Statistics, Vol. 2, Classification, Pattern Recognition and Reduction of Dimensionality, North Holland, 1982.

33. K. S. Fu (Editor), Applications of Pattern Recognition, CRC Press, 1982.

34. J. R. Ullmann, 'Advances in Character Recognition', in Applications of Pattern Recognition, K. S. Fu (Editor), CRC Press, 1982, pp. 197-236.

V. RECOGNITION SYSTEMS: EXAMPLES

The work in pattern recognition and image processing in the last twenty years has provided a wealth of new methods and techniques, but commercial applications lag behind primarily because of cost and efficiency factors. An edge detection or object recognition algorithm may provide excellent final results, but if the processing time is excessive, it cannot currently be considered for applications. In 1972 the Machine Perception group at General Motors Research constructed an automatic wheel-mounting machine with a vision system [1]. The process included locating the holes in a wheel and aligning them with the studs on a hub. The demonstration proved to be a technical success with a wide range of types of wheels, but was a practical failure, for, according to the report. "the total time required for one wheel mounting was approximately one minute" [1]. Cost and speed must be factors in any practical vision system!

We briefly review three recognition systems here. These systems represent some of the early successful industrial applications of pattern recognition and image processing. Many current recognition systems are patterned after these systems.

5.1 SIGHT 1

In 1977 an automated vision inspection system for integrated circuit chips was installed at Delco Electronics in Kokomo, Indiana [2]. The tasks accomplished by the system are:

(a) determine the position and orientation of a Darlington IC chip mounted on a heat sink;

(b) inspect the chip for flaws or defects;

(c) align test probes for circuit testing.

The size of the chip (5.25mm x 5.26mm) presented a problem both with manual inspection and vision system design because of the final detail in the chip. Following are the important processing stages in the design of this system:

(i) Image Sensing: A solid-state array camera with 50x50 spatial resolution and 16 gray levels is used to record the image.

(ii) Image Segmentation: The 3x3 Sobel gradient operator is applied at every pixel. The angular orientation of the chip is estimated. If the 'tilt' is more than 30° from the horizontal the chip is rejected because of testing and later assembly operations.

(iii) Object Description: The 'exact' position of the chip is determined by finding the four corners of the chip. If the location of the chip is found to be off of the center of the heat sink or too close to the weld cup it can be rejected at this time.

(iv) Object Recognition: Flaw detection is the final inspection task. The contrast along each edge of the located chip is calculated. A broken or missing chip will be missing part of an edge and the contrast should fall below a pre-set threshold and be rejected. For chips not rejected, the position and orientation information allows test probes to be positioned for component testing.

(v) <u>Performance</u>: Currently, the entire procedure requires less than one second per chip. As for reliability, sample tests resulted in the following [2].

(a) 270 acceptable chips were all found by the SIGHT-1 process;

(b) 33 chips out of alignment or out of position were all rejected by the process;

(c) Out of 39 defective assemblies, 16 were rejected by the vision system. However, such assemblies are not commonly encountered, and subsequent electrical testing identifies the defective chips.

SIGHT-1 has been used for many industrial models since it has been proven to be both reliable and efficient (much faster than manual inspection and testing).

5.2 <u>KEYSIGHT</u>

Automotive engine heads are inspected for the presence of valve spring cap keys prior to final assembly. A missing key can cause severe engine damage. In 1979 a computer vision system was installed at Chevrolet's engine plant in Flint, Michigan, which inspects 400 heads per hour [3]. The rate of inspection is crucial because the assembly process cannot be slowed without serious cost implications. The basic vision methods are similar to those developed in SIGHT-1 and we discuss each aspect of the vision design.

(i) Image Sensing: Bright light is concentrated vertically downward so the presence of metallic objects (such as the cap keys) will reflect light, whereas a missing key will produce a dark region.

A linear 64 pixel array camera records the image of each assembly as the engine head passes under the camera. The assembly line speed is controlled and correlated with the camera scan rate so that the images will contain a 64 x 64 two-dimensional image of each valve spring assembly. Sixteen (4 bits) gray levels are used for the possible intensities at each pixel.

(ii) Image Segmentation: Once again a Sobel gradient operator is used to find the magnitude point at all points. A thinning operation is used prior to thresholding.

(iii) Object Description: The pixels obtained thus far should roughly represent a circular region. The center of symmetry is found. An edge mask of an idealized assembly is applied to confirm the configuration obtained. Any large deviation from the edge mask causes immediate rejection of the assembly.

(iv) Object Recognition: For the final inspection the intensity profile of the original image is observed along a circle centered at the now identified center of symmetry with radius chosen so the circle passes through the suspected key positions. If more than 20% of the pixels on this circle have intensity less than half the maximum intensity, the assumption is made that at least one key is missing and the assembly is rejected.

(v) Performance: The algorithm is written in assembly language and requires less than 0.8 seconds per assembly (6.4 seconds per head) [3] on a minicomputer. In the nearly four years since installation, the vision system has proved to be highly reliable and accurate.

5.3 CONSIGHT

A potentially further reaching project than the previously discussed inspection systems is CONSIGHT, a computer vision system which identifies objects on a conveyor belt to enable a robot to transfer the objects to locations depending on the object [4].

(i) Image Sensing: As in KEYSIGHT, a linear array camera encodes the image information as the object passes beneath the line of sight of the camera. One innovative feature of the system is the lighting scheme. In many industrial settings, contrasts between objects and the background may not be high enough for easy and efficient analysis. In CONSIGHT, this problem is avoided through the use of a directed beam from two sources focused directly beneath the line of sight of the camera on the conveyor belt. An object on the belt will deflect the beam and a dark region will be recorded by the camera as the object passes, thus obtaining essentially a binary image.

(ii) Image Segmentation: The actual segmentation is trivial because a binary image is obtained.

(iii) Object Description: The features or descriptors included for each object include the zeroth through third moments of the two dimensional image as well as key numbers such as minimum and maximum coordinate values. The actual features used depend on the nature of the parts to be identified. Some examples of attributes which can be calculated from the moment information already obtained which are independent of position and rotation are:

1) Total area;
2) Number of holes;
3) Area of the holes;
4) Ratio of the minimum to maximum moments of inertia.

(iv) Object Recognition: The features analyzed are previously determined for known objects, and a nearest-neighbor decision rule is used for identification. If the similarity (judged by distance in feature space) is not sufficiently high, the recognition is voided and the part is labelled as "unknown".

One complication is that three-dimensional objects are being analyzed in a two-dimensional representation. It is very likely that an object has more than one silhouette image, arising from different stable positions on the conveyor belt. The solution is to have all possible stable configurations of every part which may be encountered treated as different models and stored in memory prior to the actual implementation.

Once identified, the part's position is represented as (x,y,θ) where x is the x-coordinate of the center of area, y is the y-coordinate of the center of area, and θ is an orientation angle calculated from the major axis (of least moment of inertia), minor axis (of largest moment of inertia), center of area coordinates, and other factors. The exact algorithm is part-dependent and is related to the fact that the information must be passed to a robot arm which has been "taught" how to pick up the identified part from its orientation and place it in a desired location.

(v) Performance: CONSIGHT is more a testing project for future robotic vision systems than an end of itself. A major limitation is its inability to distinguish overlapping parts. Nevertheless, the system capabilities are impressive. The current configuration can handle loads on a "30 cm wide conveyor travelling at 20 cm/sec with up to eight parts in the field of view simultaneously" [4].

5.4 Summary

Pattern recognition and image processing applications are not only technologically feasible but economically practical at the current time as has been demonstrated by the systems discussed (as well as many other commercial applications not discussed here) [5]. Future applications are virtually unlimited, as computer hardware and architecture advances will make even more techniques commercially possible. Even greater flexibility in systems is likely to appear in the near future. One example is a multi-camera system which uses one camera in a robot hand and another camera to monotor the entire scene [6].

- 39 -

References

1. J. Olsztyn, L. Rossol, R. Dewar, N. Lewis, "An Application of Computer Vision to a Simulated Assembly Task", Proc. First Int'l Joint Conf. on Pattern Recognition, Washington, D.C., 1973, pp. 505-513.

2. M. L. Baird, "SIGHT-1: A Computer Vision System for Automated Chip Manufacture, "IEEE Trans. Syst. Man & Cyber., Vol. SMC-8, No. 2, 1978, pp. 133-139.

3. W. Perkins, D. Juliette, R. Dewar, "KEYSIGHT: A Computer Vision System for Inspecting Valve Spring Assemblies", GM Research Publication #GMR-3601, 1981.

4. S. Holland, L. Rossol, M. Ward, "CONSIGHT-1: A Vision Controlled Robot System for Transferring Parts from Belt Conveyors", Computer Vision & Sensor Based Robots, Plenum Publ., pp. 81-97.

5. K. Fu, "Pattern Recognition for Automatic Visual Inspection", Computer, December, 1982, pp. 34-40.

6. G. Agin, "Computer Vision Systems for Industrial Inspection and Assembly", Computer, May 1980, pp. 11-20.

VI CURRENT STATUS AND RELEVANCE OF PATTERN RECOGNITION AND IMAGE PROCESSING IN DEVELOPING COUNTRIES

6.1 Current Status

The number of developing countries where there is at least some capability of expertise in pattern recognition and image processing is very small. This is evident from the membership of the International Association of Pattern Recognition (IAPR) which is concerned with the promotion of pattern recognition and image processing. Its member societies are located in the following countries: Austria, Belgium, Canada, Peoples's Republic of China, Denmark, Finland, France, Federal Republic of Germany, Hungary, India, Israel, Italy, Japan, Mexico, The Netherlands, Spain, Sweden, Switzerland, United Kingdom, and the United States of America. Even in the developing countries with some activity in pattern recognition and image processing, it is by and large restricted to the academic institutions or the government research laboratories. Further, most of the academic research groups have been concentrating on theoretical problems because of the lack of adequate computing resources. Therefore, the industrial organizations in a developing country have not been able to benefit from the academic community in terms of feasibility studies before investing a lot of money in adopting pattern recognition and image processing methodology. Another problem in a developing country is that typically its industry has very little money earmarked for research and development.

India is a good example of this phenomenon. There are a number of research groups at Indian Institutes of Technology, Indian Institute of Science and Indian Statistical Institute which are working on interesting problems related to classifier design, image enhancement, segmentation, and coding. Several good PhD theses have been written on these topics at these institutes. However, no research group in India is investigating the various important applications arising due to industrial automation and robotics, e.g. automatic inspection of parts, printed circuit board inspection, and automatic welding. This lack of industrial activity in pattern recognition and image processing has resulted in a large number of trained scientists and engineers to migrate to developed countries where there are numerous opportunities in this field. India has a diverse and strong manufacturing sector. But even the companies which make electronic components and devices or printed circuit boards, where quality control is crucial, have not started to think in terms of adopting machine vision technology. This is partly due to the lower priority given to productivity and quality control and partly due to the lack of awareness about the advances in machine vision technology.

Developing countries have made serious efforts to acquire expertise in remote sensing. For example, several countries in Latin America have acquired hardware and software to process satellite images. Most of these activities are, however, handled by either the government research laboratories or the government user agencies, and there is very little involvement of the industrial sector or even the academic institutions. Quite often, this transfer of technology in remote sensing is not complete in the sense that the users in the developing countries only know how to run a particular software package imported from developed countries but do not completely understand the principles involved in implementing or modifying these processing or classification techniques. Current emphasis on merging imaging data with map data and utilizing time-varying information requires the use of even more sophisticated pattern recognition and image processing techniques.

6.2 <u>Relevance</u>

During the last fifteen years or so, many industries in developed countries have set up assembly plants in developing countries to cut down the labor costs. However, now with increased automation, these companies are beginning to find it cheaper to assemble the products in the home country. Thus, a migration of jobs from developing country to developed country is taking place. This makes it imperative for the industries in the developing countries to adopt the current manufacturing technology. The same argument applies in situations involving quality control and hazardous environment for human workers. A machine vision system used to spray paint an automobile body offers three advantages over manual operation: 1) better finish 2) less consumption of paint and 3) elimination of the human from the hazardous environment (paint fumes). Thus there is a definite need for the industrial organization in developing countries to adopt pattern recognition and image processing methodology.

Pattern rcognition and image processing are relevent in developing countries for the same reasons that they are relevant in developed countries. Image processing and pattern recognition should be considered whenever two conditions exist:

(i) An image or a signal is required to make an important decision quickly

(ii) Raw image or signal from the sensor is inadequate (noisy) in some way.

The processed image or signal can be either interpreted by a human or by a computer. Examples of pattern recognition and image processing systems already in use include taking an inventory of forest lands from remotely sensed images, determining whether a seismic wave was generated by an earthquake or other events, analyzing acoustic signatures of truck motors to predict impending failure, inspection of brake assemblies in an automotive plant, inspection of glass containers, measurement and inspection of steel bars in a steel mill, alignment and placement of automobile windscreens, and inspection of painted surfaces for defects. The number of such applications is endless and they are all relevant to developing countries.

Adoption of pattern recognition and image processing methodology has been motivated by the following considerations:

(i) Increased productivity

(ii) Lower cost

(iii) Better quality of finished product

(iv) Faster and reliable decision making

(v) Ability to incorporate data and information from different sources or sensors in decision making

(iv) Removal of human operator from a monotonous or hazardous environment

One vendor of machine vision systems boasts of increased productivity with zero defects because the vision system never blinks. The popularity of these systems is growing because they can determine distances in three dimensions, making it a valuable tool for giving sights to robots. A major auto manufacturer in the United States is using a vision system to guide glass and trim installation. The system identifies the glass, monitors prime and adhesive application and guides robots for glass installation. The vision system registers and compensates for allowable manufacturing variations. It also automatically compensates for small variations in the position of the car body on the carrier as it comes down the line. The electronics industry is also making increased use of machine vision systems. As circuit boards grow smaller, the number of components on them is increasing. This higher density along with increased throughput requirements creates a need for automatic inspection. Recent UNIDO reports [1,2] highlight the importance and role of microelectronics in a developing country.

6.3 Summary

It is clear that pattern recognition and image processing is affecting virtually every sector of manufacturing industry as well as other scientific and engineering disciplines in developed countries. It is finding applications in the food industry, pharmaceutical companies, agricultural sector, and in numerous inspection and measurement tasks. Only a very small number of developing countries are familiar with the potential of this high technology area. It is important that they take steps to adopt this technology.

References

1. 'A Silicon Foundary to Service Developing Countries Needs: A Preliminary Approach', UNIDO Report IS.444, February 1984.

2. J. M. Oliphant, 'Microprocessor Applications in Developing Countries', UNIDO Report IS.351, November, 1982.

VII EXAMPLES OF SPECIFIC APPLICATIONS OF PATTERN RECOGNITION AND IMAGE PROCESSING IN DEVELOPING COUNTRIES

Pattern recognition and image processing techniques are currently used in numerous applications and no one doubts that these techniques will gradually find many more applications. Images need to be processed and analyzed for robot vision , target acquisition and recognition, handwritten character recognition, and medical cell and tissue analysis. igital signals which arise in applications involving detection of unsafe conditions in machines and industrial plants, analysis of electrocardiograms, radar and sonar fingerprinting, seismic analysis, and speaker identification need to be interpreted. As the cost of computing decreases and computer power increases, the need for efficient and reliable processing of these images and signals for making better, reliable, and consistent decisions or increasing productivity will make the application of this science irresistible. More and more organizations are already beginning to examine the feasibility of applying pattern recognition and image processing to manufacturing processes in their own factories.

A developing country should be informed of these various applications so that it can determine a suitable application area for automation based on its own needs and requirements. It is important to understand the complex relationships among cost, benefit, speed, accuracy, and flexibility to build a successful automated system. In some applications, for example estimating crop yield from remotely sensed images, utilization of pattern recognition and image processing is the only way to obtain the desired information accurately. It is important that the purpose of the image or signal analysis be specified explicitly before a specific and satisfactory solution can be developed. The recognition system components (lighting, sensors, optics, electronics, software, and user interface) must be chosen or developed keeping in mind the requirements of the particular application.

New and exotic applications of pattern recognition are being reported every day and each one can offer potential benefits to developing countries. A recent application will revolutionize the way manufacturers of solid wood products produce their parts. The proposed system will consist of the following processing stages:

(i) Axial Tomography: The tree log is scanned by x-ray to detect such defects as knots and cracks. This information is fed into the computer which then determines how to rotate the log in such a way as to produce the longest and best boards when the log is cut. This procedure gives 15% greater yield than conventional log processing methods.

(ii) Optical Scanning: At this stage, various surface defects (knots, decay, worm hole, discolorations, splits) on boards are identified and removed. The conventional method of removing the defects using a human operator is inefficient and leads to lot of wastage. Boards are scanned by an optical scanner, and the defects are identified and classified using pattern recognition and image processing techniques.

(iii) Optimal Cutting Strategy: The defect classification information is used to calculate an optimal cutting strategy.

(iv) Laser Saws: Laser cutting of boards offers several advantages over mechanical saw cutting. For example, laser cutting reduces noise, eliminates sawdust, has low energy consumption and greater operator safety, and, most importantly, it has the ability to 'blind cut' (it can start and stop a cut anywhere on the board).

While no such Automated Lumber Processing System manufacturing plants are in operation currently, this example illustrates the importance of pattern recognition and image processing in various manufacturing tasks.

To get started, it is recommended that a developing country choose a specific problem from the application domains of remote sensing, parts inspection, or data and text entry. Again, considerable effort is needed to identify a viable application of pattern recognition and image processing. In the following, we list a number of recognition problems for each one of these application domains.

7.1 Remote Sensing

Remote sensing techniques derive information about our land, water, and atmospheric resources from airborne and satellite sensor systems. These sensors provide information in several portions of the electromagnetic spectrum from the microwave to the ultraviolet region. Some of these sensors (infrared, passive-microwave, radar) operate during day as well as night, and radar sensors provide reliable data even in the presence of clouds and bad weather [1]. Different objects on earth return different amounts of energy in various bands of the spectrum. Pattern recognition and image processing techniques are used to detect these differences which enables the identification of various earth surface features. These processing techniques have provided fast and accurate knowledge about natural resources and cultivated land which is crucial to the social and economic development of developing countries [2].

Examples of the use of remote sensing are summarized below [3-4]:

(i) Agriculture: Plant disease; monitoring crop growth; crop inventory; crop yield; identifying field boundareis; soil type and characteristics; vegetation stress analysis

(ii) Cartography & Geography: Geographic information system development; mapping of rivers, lakes, delineation of wet lands; monitoring land-use change and urban developments

(iii) Environmental Control: Monitoring of atmospheric pollution; forecast of cyclone development; study of aquatic eco-systems

(iv) Forests: Monitoring cutting practices; mapping fire damage; forest cover mapping

(v) Geology: Geothermal mapping; detecting surface alteration effects associated with mineral deposits

(vi) Hydrology: Soil moisture; flood mapping; snow cover; ice accumulation and their changes

(vii) Oceanography: Monitoring wave patterns; monitoring oil spills

It is clear from the above examples that the computer processing of remotely sensed data can lead to large-area, cost-effective resource inventory and assessment procedures [5]. Recent developments in imaging spectrometry will allow the acquisition of remotely sensed data in hundreds of spectral bands simultaneously compared to the 4-7 bands in current Landsat satellites [6]. This data will allow a direct identification of earth surface materials on the basis of their unique spectral reflectance features. Of course, with an order of magnitude increase in the amount of data, new processing techniques will have to be developed and utilized.

7.2 Industrial Inspection

A recent survey of end user's applications of machine vision equipment identified the following areas of interest:

(i) Inspection

(ii) Gauging

(iii) Sorting

(iv) Process Control

(v) Robotics

Among these applications, inspection is by far the most popular. It is anticipated that machine vision has a much bigger future in quality-control inspection than guiding a robot. Even though the human visual system is very effective at pattern recognition, the same system is especially susceptible to fatigue and boredom when doing repetitive tasks common in industrial inspection. Seiko watch company in Japan, for example, is using machine vision systems in its factories to sort watches by style and then sets each timepiece to the correct time in the part of the world to which it will be shipped at the rate of one watch every second.

Advantages of machine vision for automatic inspection are listed below:

(i) Non-contact inspection

(ii) No moving part

(iii) Flexible system implementation

(iv) Reusable or reprogrammable from one product to another

(v) Allows 100 percent parts inspection

(vi) Improved product quality

(vii) Significant cost savings

(viii) Product inspection at multiple phases in a manufacturing cycle

In high-volume production, cost savings are obtained by eliminating or reducing expensive manual inspection. But even in low-volume lines, improved yield or product quality offer sufficient incentives to adopt automatic inspection. With the help of the inspection techniques we can control a manufacturing process by continually measuring the process and upgrading out of control elements. Thus, early detection avoids costly unscheduled downtime.

Some of the examples of inspection problems which have been automated are listed below:

(i) Identify faults in metal cutting inserts for machine tools [7].

(ii) Inspect beans for extraneous matter such as stems and leaves before they are sent for canning process [7].

(iii) Detect bone matter in beef products before final packaging [7].

(iv) Find scratches, voids or pits in plastic eyeglass lens material [7].

(v) Locating cracks in sealed contact switch [8].

(vi) Inspect printed wiring boards for flaws (additional unwanted material, small holes or cracks in the desired conductor pattern, and insufficient widths or clearances) [8].

(vii) Measurement of hole diameter, location, and wall thickness in metal castings [9].

(viii) Inspect integrated circuit chips for structural integrity [9].

(ix) Measure the internal clearances between the knife edge of a rotating blade and a honeycomb seal in a gas turbine engine [9].

(x) Inspecting bonding conditions between two disks from ultrasonic scans.

Automatic inspection techniques will play an increasing role in the electronics industry, particularly in the manufacture of surface mounted devices [7]. Specific inspection problems include solder application, misformed solder joints, and component placement.

7.3 Data and Text Entry

Pattern recognition and image processing is also being used to input and output various kinds of information (voice, data, text, drawings, images). For example, in the field of printing, most phototypesetters are now under computer control and hence require information in digital form. Fonts of type must be stored in the computer for easy access. Until recently, these master characters were hand digitized which required enormous cost and time. With the help of image digitizers, these characters as well as logos, artwork and continuous-tone photographs can be easily stored in a computer. The newspaper industry as well as the publishing business are already using this technology.

Another fast growing application of image processing is in the field of computer aided design. Line drawings, printed circuit boards, and integrated circuit layouts can be digitized, displayed on a system, manipulated by a human operator, and the final product then printed or stored in computer memory for later use. Automatic document readers which cost in the range of $5,000-$10,000 are increasing office productivity by automatically reading text into word processors, personal computers, or phototypesetters. This avoids re-keying already typed material. These systems can input a page of text in 15 seconds, which is up to thirty times faster than the average typist. Other applications of optical character readers (OCR) include mail sorting at post offices, processing bank checks, and inspecting and routing packages in a warehouse.

Voice input systems are also useful in industrial applications, especially in the following situations [10]:

(i) The worker's hands are busy so he cannot use a keyboard.

(ii) Mobility is required during the data entry process.

(iii) The worker must constantly monitor a display or an instrument gauge and cannot take his eyes off them.

(iv) The environment is not too clean to allow the use of a keyboard.

The main advantage of voice input is that it does not need hands or eyes. So the worker may be tracking or controlling a process and still give verbal commands or instructions, or provide data about the behaviour of the process. Several automatic speech recognition systems are available commercially. These systems must be trained for individual speakers and words. Current systems can recognize only isolated words but systems for recognizing connected speech are in the development stage.

7.4 Summary

Pattern recognition and image processing offer tremendous potential for simplifying or automating various decision-making tasks. Automation in remote sensing, industrial inspection, and data and text entry should benefit a developing country in gathering and utilizing information about its resources, increasing productivity, and improving the quality of its products to stay competitive in the world market. Pattern recognition and image processing techniques have also contributed to medical research and diagnostics. It has been used to search for brain tumors in CAT scan images, for brain artery aneurysms in angiogram x-rays, for white blood cell differential count, and for early detection of blood vessel blockage. It is important that a developing country stay informed about these developments so that it can adopt this emerging technology to its best use.

References

1. R. M. Hord, Digital Processing of Remotely Sensed Data, Academic Press, 1982.

2. M. Hernandez and A. Flores, 'Machine Processing of Remotely Sensed Data: Three Applications in Mexico', IBM J. Res. Develop., Vol. 22, Sept. 1978 pp. 455-463.

3. P. R. Pisharoty, 'Introduction to Remote Sensing', Proc. Indian Acad. Sci., Vol. 6, June 1983, pp. 97-107.

4. F. T. Ulaby, 'Radar Signature of Terrain: Useful Monitors of Renewable Resources', Proc. IEEE, Vol. 70, Dec., 1982, pp. 1410-1428.

5. G. Schultnik, 'Integrated Remote Sensing and Information Management Procedures for Agricultural Production Potential Assessment and Resource Policy Design in Developing Countries', Canadian Journal of Remote Sensing, Vol. 9, July 1983, pp. 4-18.

6. A. F. M. Goetz et. al., 'Imaging Spectrometry For Earth Remote Sensing', Science, Vol. 228, June 7, 1985, pp. 1147-1153.

7. Machine Vision International Newsletter, Vol. 1, Spring 1985.

8. J. F. Jarvis, 'Visual Inspection Automation', Computer, Vol. 13, May 1980, pp. 32-38.

9. W. Myers, 'Industry Begins to Use Visual Pattern Recognition', Computer, Vol. 13, May 1980, pp. 21-31.

10. J. R. Welch, 'Automatic Speech Recognition-Putting it to Work in Industry', Computer, Vol. 13, May 1980, pp. 65-73.

VIII. INITIATIVES NEEDED TO START ACTIVITIES IN PATTERN RECOGNITION AND IMAGE PROCESSING

Industrial applications of pattern recognition and image processing have started to multiply in developed countries. There is a need for the industrial sector of the developing country to get interested in the possibility of adopting this technology. This requires that an industrial organization be informed about the recent developments in pattern recognition and image processing and have access to hardware and software to design such a recognition system. Only then can a company select a specific problem where pattern recognition and image processing might help. So the best initiative to start activities in pattern recognition and image processing in a developing country is to first make the scientists and engineers in that country aware of the potentials of this high technology area.

Any serious effort in the application of image processing and pattern recognition methodology requires a critical mass of hardware, software, and personnel. This is an interdisciplinary activity requiring close interaction among computer scientists, engineers, and experts in the application area. For example, in order to develop a system to analyze and interpret satellite and aerial images for meteorological applications, pattern recognition and image processing experts must work in close cooperation with meteorologists. Processing images requires complex software consisting of algorithms for two-dimensional filtering, edge detection, region growing, shape analysis, texture, and classification. A typical image is composed of 512x512 pixels, each with 256 gray levels, requiring approximately 2 million bits of storage. Even a simple operation on an image, say 3x3 averaging or smoothing, can take several minutes on a superminicomputer. Therefore, a dedicated and interactive computing facility is required for Pattern Recognition and Image processing development work. Quite often to speed up the processing, many of these algorithms are implemented in hardware. We briefly review the requirements for training, hardware, and software to initiate a research and development program in pattern recognition and image processing.

8.1 Education and training

Different aspects of pattern recognition and image processing require different kinds of background. In order to design or completely understand sensors (image digitizers) and their interface to a digital computer, electrical engineering background is necessary. To design and evaluate various decision-making algorithms, strong background in mathematics or statistics is useful. Finally, to implement these algorithms, one must be familiar with programming languages, data structures, and other related computational issues.

Most major universities in developed countries offer courses at the graduate level (M.S. or PhD level) in the areas of pattern recognition and image processing. Typical prerequisites or background needed to take these courses are: college-level calculus, linear algebra, probability and statistics, high-level programming language (Pascal or FORTRAN), and data structure. With the exception of programming language and data structure, most college graduates in Mathematics, Statistics, Physics, and Engineering have the necessary background to learn the concepts of image processing and pattern recognition. However, familiarity with computers, programming, and computer technology is essential to understanding the cost and effort needed to design a recognition system.

It is possible to design a three-month full-time training program in pattern recognition and image processing for scientists and engineers from developing countries. The program can be broken up into three parts.

(i) 4-week training in pattern recognition

(ii) 4-week training in image processing

(iii) 4-week project

Classroom instruction needs to be supplemented by hands-on experience on an interactive computing facility. The project portion of the program will involve designing a recognition system and evaluating its performance utilizing real data in a problem domain of interest to developing countries. This will help students integrate various concepts, understand the difficulties encountered with real data, and evaluate competing approaches to solving a problem.

In addition to the large number of books on pattern recognition and image processing (listed in chapters 3 and 4), the students should also have access to the conference proceedings and journals where papers on these and related topics appear. Thus a good library facility is important for training. Some of these journals and conference proceedings are listed below.

Journals

(i) Pattern Recognition, Pergamon Press

(ii) Pattern Recognition Letters, North Holland

(iii) Computer Vision, Graphics and Image Processing, Academic Press

(iv) IEEE (Institute of Electrical & Electronics Engineers) Transactions on Systems, Man & Cybernetics

(v) IEEE Transactions on Pattern Analysis and Machine Intelligence

(vi) IEEE Transactions on Geoscience and Electronics

In addition to these journals, which publish general research papers, there are specialized journals which concentrate on applications of pattern recognition and image processing in a specific area. For example, Robotics Today and Manufacturing Engineering routinely publish papers related to machine vision applications in industry.

Proceedings

There are a large number of conferences held every year on various aspects of pattern recognition and image processing, indicating the popularity of this subject in both academic and industrial sectors. The most well-known of these conferences are:

(i) International Joint Conference on Pattern Recognition held in Washington D.C. (1973), Lyngby, Denmark (1974), Coronado, California (1976), Kyoto, Japan (1978), Miami Beach (1980). Montreal (1984), Paris (1986)

(ii) IEEE Conference on Computer Vision and Pattern Recognition formerly known as Pattern Recognition and Image Processing held in Troy, N.Y. (1977), Chicago (1978,1979), Dallas (1981), Las Vegas (1982), Washington, D.C. (1983), San Francisco (1985), Miami Beach (1986).

The Society of Manufacturing Engineers (SME) organizes an annual conference on machine vision applications in industry. Scientists and engineers from the developing countries should be given an opportunity to attend some of these international meetings on a regular basis.

8.2 Hardware

Commercial pattern recognition and image processing equipment is available in a variety of forms. They range from simple image manipulation capabilities to high-volume processing power as needed, for example, in remote sensing. The price of these systems ranges from $10,000 to $500,000. The more a system costs, the faster and more flexible it will be in handling the date [1].

An image processing system, irrespective of its size and cost must have the capability to handle the following tasks:

(i) Image digitization - conversion of analog pictorial information into digital form

(ii) Image storage - large disks and tape drives are needed to store digital images

(iii) Image display - to view raw and processed images

(iv) Processor - To execute various enhancement and classification tasks

If one is interested in classifying 1-dimensional signals such as speech, then the image digitizers can be replaced by an analog/digital converter.

An interactive computer environment is essential for image processing and pattern recognition. This facilities the interplay between man and machine to interpret, understand, and classify images, signals, or objects [2-5]. Early image processing systems used in industrial application were binary systems because of limitations in speed, cost, and computer memory. Current computer technology allows gray scale processing at a reasonable cost. Color systems are now appearing which will further increase application potential, especially in agriculture.

One can get started in image processing al ost immediately with an investment of about $5,000 for a personal computer, $300 for a camera, and $200 for a board to "grab" and filter the frames from the camera. Some vendors will sell a complete desk top image processing system for about

$30,000. A large multiuser system for a research group could easily cost $100,000 or more. For example, a typical system in such a price range ($100,000) can be configured using the following components:

MC 68010 CPU

80 MB Winchester

1 MB Floppy Disk (5 1/4")

1/2" 9 Track Tape Drive

2 MB Memory

Floating Point Processor

Console CRT

19" Color Monitor

AP-512 Analog Board

FB-512 Frame Buffer

256x256 CCD Camera

UNIX Operating System with FORTRAN & C

P-Cubed Image Processing Software

In fact, most inspection systems installed in industry cost about $100,000 and up. Information about such systems can be obtained from advertisements in the trade journals covering image processing (such as Electronic Imaging). Introductory information, lists of companies, and dates of exhibitions of vision and robotics systems can be obtained from the Machine Vision Association of SME, One SME Drive, P.O. Box 930, Dearborn, MI

48121 [1]. Some well-known vendors of image processing systems and their addresses are given below:

(i) Analog Devices, Inc.
Two Technology Way
Norwood, MA 02062

(ii) Machine Vision International
325 East Eisenhower
Ann Arbor, MI 48106

(iii) Spatial Data Systems, Inc.
P.O. Box 249
500 S. Fairview Avenue
Goleta, CA 93017

(iv) VICOM
101 Flames Road
Marshfield, MA 02050

(v) ESL, Inc.
495 Jave Drive
P.O. Box 510
Sunnyvale, CA 94086

(vi) Gould De Anza
1870 Lundy Avenue
San Jose, CA 95131

(vii) Image Technology
Box 5047
S-580 05 LINKOPING
SWEDEN

(viii) Image 100
General Electric Company
P.O. Box 2500
Daytona Beach, FL 32015

(ix) MICRO CONSULTANTS LTD.
Kenley House, Kenley Lane
Kenley, Surrey CR25YR
ENGLAND

Important factors which should be considered in selecting a system are:

(i) cost

(ii) suitability to the particular application

(iii) ease of expanion

(iv) training program

(v) maintenance contract

(vi) interface to existing digital computers in the organization.

A recent paper [3] makes several recommendations for a potential buyer.

8.3 Pattern Recognition and Image Processing Software

Processing strategies needed for signals and images available from different domains are different. For example, the nature and source of noise and distortion in ultrasonic images in biomedical applications is different from those present in x-ray images of welds in nondestructive testing. This explains the large number of processing and classification techniques available in the literature. There is no 'optimal' sequence or set of processing steps. The given images or signals influence the choice of image processing and pattern recognition techniques. Therefore, it is important that a user has access to a large variety of algorithms. The best way to get started is to purchase a comprehensive software package. This will eliminate the large amount of resources needed to write and test the software inhouse. The necessary manpower required to write many of the complex routines may also not be available in a developing country. Further, purchasing an already proven package can cut down the valuable start-up time.

Various image processing operations can be broadly classified into the following categories.

(i) Basic Operations (Histogram, thresholding, etc.)

(ii) Image Transforms (Fast Fourier transform, convolution, etc.)

(iii) Registration (Correlation, affine transform, etc.)

(iv) Enhancement and Smoothing (Histogram transform, edge preserving smoothing, etc.)

(v) Restoration (Inverse filter, Wiener filter,etc.)

(vi) Edge and Line Detection (Differential, template matching type, Hough transform, etc.)

(vii) Texture Analysis (Cooccurrence matrix, Fourier features, etc.)

(viii) Region Segmentation (K-S test, split and merge, etc.)

(ix) Geometry (Boundary description, thinning, etc.)

(x) Shape (Moments, Fourier descriptors, etc.)

A number of image processing software packages which include algorithms for the above operations are available in the market. Most of them are written in FORTRAN language and can be easily run on a variety of computers, ranging from minicomputers to main-frames. Two of these image processing packages are:

(i) SPIDER

Joint System Development Corporation
Yuseigojyokai- Kotohira Bldg. 14-1, 1-Chome
Toranomon, Minato-Ku
Tokyo
JAPAN

(ii) P-CUBED

Advanced Technology Transfer
5179 Overland Avenue
Culver City, CA 90230
USA

The cost of these systems vary a lot depending on how many subroutines are purchased, but the price is in the neighborhood of $20,000.

Pattern recognition operations can be broadly classified into the following categories:

(i) Exploratory techniques

(ii) Parameter estimation

(iii) Parametric classifier

(iv) Density estimation

(v) Nearest-neighbor rules

(vi) Feature selection

(vii) Error estimation

(viii) Clustering

The software package ISPAHAN incorporates many of the above pattern recognition techniques. Further information about this package can be obtained from:

Department of Medical Information
Vrije Universiteit
Amsterdam
THE NETHERLANDS

Both pattern recognition and image processing are computation-intensive. Therefore a user must understand the various options and routines contained in these software packages.

8.4 Summary

The following are the key requirements to initiate an activity in pattern recognition and image processing:

( ) A team of 4-5 scientists with backgrounds in Electrical Engineering/Computer Science/Statistics/Mathematics/Physics.

At least some of them should have graduate degrees. A three-month full-time intensive course on pattern recognition and image processing will give sufficient background to undertake an application project.

(ii) A dedicated computing facility consisting of a superminicomputer with tape drives and disks, digitizer, high-resolution raster color display, film recorder, and printer. A number of companies (COMTAL, VICOM, De Anza) sell a complete system, starting from $100,000. Such a facility can also be used for scientific computing in addition to pattern recognition and image processing work.

(iii) A Comprehensive and proven software package containing commonly used pattern recognition and image processing algorithms. Quite often the hardware vendors provide a software package with limited capabilities. A good package sells for about $20,000.

(iv) A library facility with a good collection of books and periodicals on pattern recognition and image processing.

Once the hardware, software and manpower requirements are met, the successful application of this technology will depend on the choice of specific problems and the necessary funds for exploratory projects. Collaborative research efforts between industry and academia will alleviate the manpower to identify relevant problems for automation in developing countries.

References

1. G. Stockman and A. K. Jain, 'Imaging Environment and General Image Processing Algorithms' Paper No. 85-3034, Proceedings of the 1985 Summer meeting of American Society of Agrigultural Engineers, June 1985, East Lansing, MI

2. R. Jarvis, 'An Interactive Minicomputer Laboratory for Graphics,

Image Processing and Pattern Recognition', Computer, Oct. 1974, pp. 49-60.

3. H. C. Andrews, 'An Educational Digital Image Processing Facility', Computer, August 1977, pp. 48-53.

4. R. Jarvis, 'A Computer Vision and Robotics Laboratory', Computer, June 1982, pp. 9-22.

5. G. Gini and M. Gini, 'A Software Laboratory for Visual Inspection and Recognition', Pattern Recognition, Vol. 18, 1985, pp. 43-51.

6. A. Novini, 'Before You Buy a Vision System ...', Manufacturing Engineering, March 1985, pp. 42-48.

APPENDIX: Glossary of some pattern recognition and image processing terms

Charged-Couple Device (CCD) Camera	An image sensor consisting of an array of semiconductor elements; the electric charge at the output of an element depends on the input light intensity
Coding	Procedures to represent an image to minimize storage requirements.
Component	A collection of picture elements which share some common properties. A component usually corresponds to some physical entity in an image.
Computer-Aided Design	The use of computers in design with interactive graphics capabilities.
Computer-Aided Manufacturing	The use of computers to manage and control the

	operations of a manufacturing facility.
Digitization	The process of converting a two-dimensional continuous brightness function (e.g. a photograph) into an array of pixels, each pixel takes one of a finite number of gray values.
Enhancement	A collection of image processing operations to improve the appearance of an image. These operations are general in nature and do not utilize any specific properties of the image.
Features	Properties of individual picture elements (e.g. gray level, edge intensity) or a collection of pixels (e.g. size, texture).
Frequency domain	A one-dimensional signal

	(waveform) or a two-dimensional signal (image) can be converted to frequency domain representation by Fourier transformation. Frequency domain representation allows a more efficient implementation of some image processing operations.
Gray values	It is a quantized or discretized value of the brightness of a picture element. Typically 256 different gray values are used with 0 denoting complete black and 255 denoting complete white.
High Pass Filtering (Sharpening)	An image processing operation which highlights the boundaries of the objects present in an image.
Histogram	A graph representing the frequency of occurence of each gray level in an image.

Image	A digital representation of a scene by an image sensor as an array of pixels.
Image Processing	A set of operations to improve the quality of an image, or to locate and determine the properties of objects or patterns of interest in the image.
Low Pass Filtering (Smoothing)	An image processing operation which eliminates noise. The operation involves replacing the gray value of a pixel by the average of its neighboring pixel's gray values.
Machine Vision Systems (Computer Vision)	Systems that automatically inspect components and aid robots in assembly operations.
Optical Scanner	An image sensor where an electron beam scans the surface of photoelectric

material where the scene to be sensed is projected

Non-parametric decision rule — The decision rule assume no parametric form for the statistical distribution of the feature vector. Thus its distribution must be estimated from training samples

Parametric decision rule — The decision rule assumes a certain parametric form for the statistical distribution of the features (e.g. Gaussian distribution). The parameters of the distribution are estimated from training samples

Pattern Recognition — The collection of techniques that classify objects or signals into one of a set of predetermined categories

Pixel (picture elements) — The individual elements in a digitized image. This is the smallest distinguish-

	able area of an image and hence pixel size is inversely related to the resolution of the image sensor.
Remote Sensing	The science of gathering information about (sensing) objects or areas at great distances.
Restoration	These enhancement techniques model the source of noise or distortion present in the image as well as the properties of the image to improve its appearance.
Robotics	A reprogrammable manipulator to perform a variety of intelligent tasks (e.g. inspection, sorting, assembly).
Segmentation	Decomposition of an image into components, each of which may correspond to a separate object or part of

	an object in the image.
Test Samples	Patterns belonging to various categories which are used to estimate the performance of the decision rule.
Training Samples	Patterns belonging to various categories which are used to estimate the parameters of the decision rule.