Def Sci J, Vol 37, No. 4, October 1987, pp 457-468

Image Processing and its Military Applications

V.V.D. Shah

Military College of Telecommunication Engineering, Mhow–453 441

ABSTRACT

One of the important breakthroughs, image processing is the stand alone, non-human image understanding system (IUS). The task of understanding images becomes monumental as one tries to define what understanding really is. Both pattern recognition and artificial intelligence are used in addition to traditional signal processing. Scene analysis procedures using edge and texture segmentation can be considered as the early stages of image understanding process. Symbolic representation and relationship grammers come at subsequent stages. Thus it is not reasonable to put a man into a loop of signal processing at certain sensors such as remotely piloted vehicles, satellites and spacecrafts. Consequently smart sensors and semi-automatic processes are being developed. Land remote sensing has been another important application of the image processing. With the introduction of programmes like Star Wars this particular application has gained a special importance from the Military's point of view. This paper provides an overview of digital image processing and explores the scope of the technology of remote sensing and IUSs from the Military's point of view. An example of the autonomous vehicle project now under progress in the US is described in detail to elucidate the impact of IUSs.

1. INTRODUCTION

Image processing – a moving horizon! Walking towards a horizon is open ended. The horizon never gets any closer to you, but continually recedes from you. Thus it has been with the growth of image processing, as a technical discipline. Constant progress is being made – but the potential is far from exhausted. In the early years of image processing the concern was of basic phenomena, for example, making models for image data compression, image restroation and image enhancement. Currently there is a great interest in moving beyond physical phenomena and into the realms that are wrapped with psychology, perception and cognition. The research in this branch of image processing is often called as 'Image Understanding'.

Received 27 June 1987

The ability to place earth oriented sensors into orbit, because of its large economic potential has received considerable research emphasis, and many operational systems have been evolved. Land remote sensing has been evolving since 1960s and its various applications in urban land utilization, forestry and food commodity production forecasting has been producing remarkable results. An off shoot of remote sensing by satellites has been the use of this data for military purposes.

1.1 Defining Image Processing

The traditional view of image processing tends to embrace one or more of picture processing, pattern recognition, image interpretation, or even graphics. 'Graphics' deals with the generation of images from non-pictorial information and covers diverse applications. In the order of increasing complexity, production of plots of functions, composition of displays for the computer games and scenes used in flight simulators are some of the examples of displays. Picture processing deals with problems in which both input and output are pictures. Over exposed, under exposed or blurred pictures can be improved with contrast enhancement techniques.

'Pattern Recognition' deals with methods for producing either a description of the input picture or an assignment of the picture to a particular class. In a sense, it is the inverse problem of computer graphics. It starts with a picture and transforms it into an abstract description, a set of numbers, a string of symbols, etc. Further processing of these forms results in assigning the original picture to one of several classes. An automatic mail sorter that examines the postal code written on an envelope and identifies the digits is a typical example of the application. However, the term 'Image Processing' should be used as a catch all for all these activities and in a much broader context with the implicit understanding that the fundamental underlying activity is that of 'Information Processing'. Increasingly, in the future, much information is going to be represented, and subsequently processed, as digital images, be they X-ray scans, satellite images, video films or whatever. This is no more than a reflection of the fact that our information processing channel with the highest band width, by a long way, is the visual one. It is this primacy of images in information representations that renders a digression into the possible social impact of information processing.

2. BASIC CONCEPTS

2.1 Enhancement and Restoration

The formation of an image; or its conversion from one form into another; or its transmission from one place to another often involves some degradation of image quality; with the result that the image then requires subsequent improvement, enhancement or restoration. Image restoration is commonly defined as the reconstruction or estimation of an image to correct for image degradation and approximate and ideal degradation free image as closely as possible. Image enhancement involves operations that improve the appearance of an image to a human viewer, or convert an image to a format better suited to machine processing. The basic distinction between enhancement and restoration is that with the former no attempt is made to establish the underlying degradation, while with the latter, specific known or estimated degradation processes are assumed. The former embraces such techniques as contrast modification, deblurring, smoothing and noise removal, while restoration tends to revolve around formalism of filter theory.

Contrast modification involves two types of basic operations: 'Grey Scale Correction' and 'Grey Scale Transformation'. The former attempts to compensate for uneven contrast, usually the consequence of non-uniform exposure, by modifying grey levels of individual pixels. The latter is concerned with increasing the dynamic range of image contrast and includes techniques like 'Histogram Modification'. Blurring is averaging or integration process in that high spatial frequencies are attenuated. Thus blurring can be corrected by the use of differential operations. Smoothing is essentially concerned with noise removal. The noise may be systematic or random, the former being easier to deal with in practice – for example, with Fourier techniques. Noise in an image usually possesses significant high spatial frequency components in its spectrum relative to the image under consideration. As a consequence simple low-pass spatial filtering can be quite effective.

2.2 Segmentation

Segmentation is a generic term for those techniques which involve taking an image and extracting information relevant to specific picture 'Segments', such as lines, regions and objects, and their inter-relationship. It is basically a process of data compression by pixel classification, in that an image is segmented into subsets by assigning individual pixels to particular classes. Some of the techniques used for segmentation are clustering, pattern classification and grey level thresholding. Clustering and pattern classification involve classification of a set of pixels according to the value(s) of some image property which can be measured or evaluated for each relevant set of pixels. Grey level thresholding is a segmentation technique which classifies pixels into classes on the basis of their grey levels. In its simplest form it involves classification into two classes, black and white, on the basis of a chosen threshold.

2.2.1 Image Feature Extraction

An image feature is a distinguishing primitive characteristic or attribute of an image field. Some features are natural in the sense that such features are defined by visual appearance of an image while other so-called artificial features result from specific manipulations or measurements of an image. Natural features include the brightness of a region of pixels, edge outlines of objects, and grey scale textural region. Image amplitude histograms and spatial frequency spectra are examples of artificial features.

3. MILITARY APPLICATIONS

There are two compelling reasons why image processing will play an increasingly important role in future defence systems. These are the need for autonomous operation and the need to make greater use of the outputs from a diverse range of sophisticated sensors.

3.1 Remote Sensing

Remotely sensed imagery has now been available from the Landsat series of satellites for over 10 years in US. During this period, the data use has expanded from investigations performed by a group of sponsored investigators to attempted routine use by private companies, agribusiness, and resource planners in that country. The data is becoming a powerful tool in meeting the needs of the world in resource exploitation and management. New analysis methods are being developed to take advantage of the new types of data. The Seasat program, a short lived one, provided a source of synthetic aperture radar data which responds to a different set of parameters (surface angle, surface roughness, dielectric constant, which are affected by factors such as soil moisture, vegetative cover or water surface geometry). It also provided other instruments such as radar altimeter and a microwave scatterometer. All these instruments are widely used in the military. Recent intelligence reports of the Iran-Iraq War and how the satellites were used for the reconnaissance of each other's territory is a case which brings to light the vulnerability of today's Armies to the remote sensing technology. At the tactical level even sensing of the enemy minefields may be done by satellites. On the strategic level, verification of the arms control agreements strongly depends on image processing to identify and count missile silos from reconnaissance images.

3.1.1 Remote Sensing Technology

The first step in remote sensing is the data acquisition. This is done by placing a multichannel, multispectral scanner (MSS) with high resolution. The image information either transmitted directly to earth or, when the satellite is out of range of ground receiver, stored on-board magnetic tape recorder for subsequent transmission. When on-board storage is not used, real time processing of the data is alsc possible by Tracking and Data Relay Satellite System (TDRSS). Each geosynchronous TDRSS satellite will have high band width transponders (300 Mbps) and relays the data to a single predetermined earth station. After reception the data is achieved on high density tapes from which both laser recorded photographic data and computer compatible tapes are generated. The digital images are system corrected at the image processing centre on the basis of data received from the tracking stations, including attitude, and altitude. Radiometric correction based on on-board calibration sources removes sensor and digitizer anomalies. Geometric correction compensates for satellite altitude and attitude changes, skewing introduced by satellite motion, and the effects of the angle subtended by satellite's field of view. The images are resampled, i.e. the grey value assigned to a pixel in the target image is a weighted average of the points corresponsing to the target point in the resource image. Haze removal, compensates for atmospheric scatter. Contrast stretching, allows use of the full dynamic range for low-contrast images. Edges enhancement, which exaggerates the difference in intensity between a given pixel and a user specified neighbourhood. Users who prefer to avoid the loss of discrimination due to resampling may only have radiometrically corrected data; this results in loss of positional accuracy.

Another important process is the process of classification. The incorporation of previously defined digital image representation into systematic methods of determining the required attributes of object scenes is the subject of classification. Classification includes concepts like categorization, identification, recognition clustering, partitioning, taxonomy and segmentation of concern are supervised and unsupervised learning, teaching or training, estimation of parameters, distributions and error rates. Various algorithms have been made for these processes. However it is not reasonable to place a man into a loop of signal processing at certain sensors, such as remotely piloted vehicles, satellites and space crafts. Consequently many a breakthroughs can be expected in what is called as 'Image Understanding' – the learning of some basic information about an image without a human in a visual role. Smart sensors and semi-automatic processes are being evolved. A machine that can segment an image into regions of homogenous content; consistent with human segmentation is the answer.

3.1.2 Understanding Images

For obvious reasons, rapid response is becoming increasingly important in Defence systems. This forces decisions to be made without explicit operator intervention. Indeed, in many cases operator intervention will not be physically feasible. Targetting, surveillance, command and control activities all need rapidly to make sense out of large amount of desparate and possibly unreliable information. This is bound to require the application of advanced Artificial Intelligence (AI) methods, especially those relating to image understanding. Subsequent control may be passed to an autonomous system which will attempt to select an appropriate target from captured image data set, and initiate an appropriate response. One obvious area of concern is the role of autonomous systems during a transition from peace to hostilities and vice versa, when the rules of engagement may alter very quickly. An extremely cogent example of such a dilemma is the case of any projected defence against strategic nuclear missiles, where systems act so rapidly that almost complete reliance may have to be placed on automated systems in circumstances where the complexity and unpredictability of factors affecting irreversible decisions is very great indeed. The task of understanding images becomes monumental as one tries to define what 'Understanding' really is. Both pattern recognition and AI techniques are used, in addition to signal processing. Scene analysis procedures using edge and texture segmentation, as well as object recognition, can be considered early stages of an Image Understanding System (IUS). On the other hand, symbolic representation, relationship grammers, and syntactic manipulations are the later stages. Cultural experiences are built into the task of understanding an image. For example, the recognition and understanding of the airport in a reconnaissance image is extremely difficult, if any one who has not flown at high altitudes or observed aerial photography. The simple segmentation of an image into its homogenous parts, a task performed consistently by humans is still relatively difficult , if not impossible, for a machine. The long experience of a professional viewer, photo interpreter, criminologist, radiologist – is probably not possible to duplicate in machines in near future. But pay-offs and rewards will be great once success is achieved. As an illustration of the difficulty of automating even the simplest task, consider the problem of edge detection. Humans constantly experience edges in their viewing experiences, but how does one define an edge? It is a difference in grey shades, but how different? Perhaps it is different in texture, but how does one define texture? Such problems have to be overcome before a fully automatic IUS is developed.

3.2 Image Understanding System

An IUS uses visual data to generate descriptions that are useful for desired applications. The description generated can be at very different levels and degrees of details. The image is represented by an array of numbers characterising the brightness at each point on a rectangular grid. These brightness elements are called pixels. Image descriptions of these forms the starting point for IUSs.

The system generates a series of descriptions that are progressively more general, until a descriptive level is reached that satisfies the system requirements. It has been observed that the successive levels of abstraction require that higher levels of system

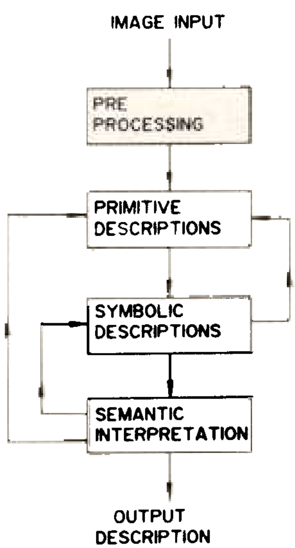

Figure 1. Image understanding system.

interact with lower levels, based on current descriptions. This processing approach is called 'Hierarchial'. The IUS can therefore be conceptualised as having a hierarchy of processing levels shown in Fig 1. The primitive description level extracts local features that are not related to context. The primary or first order features of pixel in a monochrome image are its brightness and spatial location. All other features are of higher order, i.e. they describe how the pixel is related to surrounding pixels in an image. These features describe such primitive local attributes of the picture as brightness, texture and colour. A proper primitive description on level of the IUS would transform the features into a co-ordinate system where numerical distance would be related to human perceptual difference. The symbolic description level of the system takes the primitive descriptions and forms more global and symbolic description of the image. Segmentation of the image takes place at this level. The initial segmentation is based purely on perceptual difference. After analysis by the semantic interpretation level, the symbolic level may be directed to merge or to further divide the image. Thus the decisions about dividing the scene into similar or homogenous regions are made at this level. Feedback from the semantic interpretation level is necessary to ensure that the symbolic descriptions are consistent with the goals of the IUS. The semantic interpretation level of the system generates hypothesis for the contents of the image based on the symbolic descriptions, and then further directs the lower processing levels until the symbolic description confirms one of the hypotheses.

3.3 Computational Vision

Most of the information humans receive, learn, remember and use is in picture like form. The largest projection area on the cortex of our brain corresponds to vision; this accounts for extraordinary capacity that we have for processing pictorial information strategically, i.e. according to the specific contextual needs of the observer. As a consequence a great challenge is posed if computers must substitute partially or totally the human operator in some of the tasks in which vision plays a significant role.

Vision is an information processing task with well defined input and output. It is an active process that imposes an interpretation on sensory data based on certain relations. The interpretation of data according to prior model or knowledge is the key to the signal understanding or image understanding. Determining what knowledge is needed, and how it should be represented and efficiently applied, is the essence of IUS. In computer vision, the input is in the form of images, perhaps a single image or images from different view points, or more generally, a time series of images. The output is a useful interpretation of the image data in terms of a symbolic description of the scene. Computer vision focuses on understanding the process of vision in particular by concentrating on the identifiable functional modules of the human visual system.

3.3.1 A Computational Model of the Vision

Experience with building machine vision systems, taken together with insights gained by studying human vision provide the basis of a computational model of visual processing. The design of any vision system is primarily constrained by the characteristics of the scene domain and the overall goals of the system. The fundamental considerations are :

(a) What information is sought from the image?

(b) How is it manifested in the image?

(c) What prior knowledge is needed to recover it?

(d) What is the nature of computational process?

(e) How should the information and knowledge be represented?

The goals determine what information is sought. Human vision support a wide range of tasks ranging from simple alerting (for example, sudden movement) to complex event understanding (following a movie). In between are tasks like navigation and object manipulation, analogic reasoning (for example, determining whether two objects make together) and object recognition.

For each of the above tasks there is an appropriate representation that makes the relevant information explicit in a suitable geometric or relational reference frame. An array containing the image locations of significant spatial or temporal intensity discontinuities is sufficient for alerting tasks. Navigation and manipulation require a description of the layout of the space (i.e., range and orientation) in a viewer centred co-ordinate system. Higher levels of processing require an object centred representation of surface and volumes, and so forth.

A vision system is naturally structured as a succession of levels of representation. The initial levels are constrained by what is possible to compute directly from the image data, while higher levels are dictated by the information required to support the ultimate goals. In between, the order of representation is constrained by the information available at preceding levels and the information required by succeeding levels. The processing that transforms each level of representation to the next requires knowledge from models of the scene, the illumination, and the imaging process. At lower levels, these models help resolve the ambiguity inherent in going from three dimensional world to a two dimensional image. At higher levels, they provide the basis for organising surface fragments into recognizable objects. It can be argued that any perceptual system, natural or artificial, operating in the same visual environment, must rely upon the same models to obtain consistent interpretation.

A plausible general purpose system computational vision system is shown in Fig. 2. The sensor encodes the physical characteristics of the scene into a two dimensional array of brightness values, an input image, which is the initial level of representation. Known geometric or photometric distortions introduced by sensors may be removed to produce a corrected image that facilitates subsequent processing. Information in the image is manifested primarily through spatial or temporal intensity changes (and their location in the image), which correspond to changes in some physical characteristics of the scene. The next step is to detect these changes and represent them explicitly as two dimensional arrays of feature descriptors. Edges, for example, might be described by their orientation, width and contrast. Local patterns

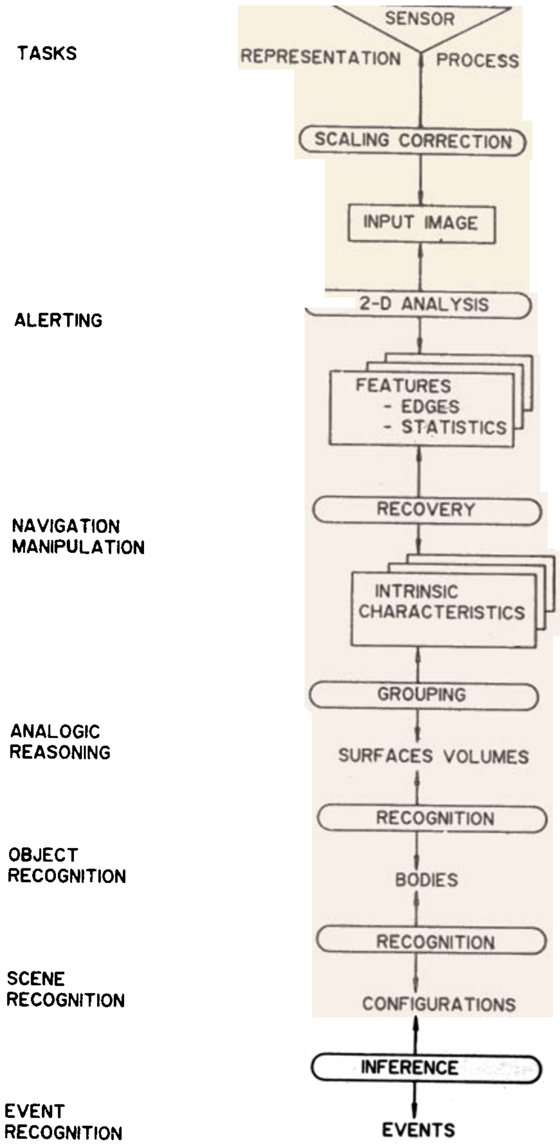

Figure 2. Computational model of a general purpose vision system

of image features yield information about the three dimensional structure of the scene in the form of texture and shading gradients, occlusion cues (for example, line endings), contour shapes and so forth. From these arrays of surface orientation, distance, and other intrinsic characteristics of the surface element visible at each point in the image can be recovered. The information represented in these arrays directly support tasks such as navigation, manipulation, and material identification. Simple iconic groupings processes operating on the arrays of intrinsic characteristics can recover regions of homogenous properties corresponding to three dimensional surfaces. The notion of a surface is a symbolic abstraction, and this is the level at which the transition from iconic representation to symbolic representation can naturally occur. Symbolic grouping processes organise surfaces into sets corresponding to distinct bodies. Bodies so described are then recognised as known objects or new ones, and symbolic representations constructed to describe them. Object representations include information about three dimensional location and orientation, and pointers to generic descriptions that provide information about characteristics that may not be directly visible (for example, weight).

3.3.2 The Autonomous Vehicle

One of the most important applications of image processing currently under progress by the US strategic computing programme, managed by the Department of Defense's Advanced Research Projects Agency (DARPA) is 'The Autonomous Vehicle'. The proposed vehicle will contain a small modular computer control system. Vision modules will be included that provide basic scene processing and object recognition capabilities. With vision modules as input devices, symbolic processor modules will then be able to directly process fragments of pictorial, graphic and three dimensional scenic images when further supported by rule based inferencing and image understanding in a compact but powerful symbol processor and interfaced with specialised motor controlled systems, these vision modules will enable the computer controlled autonomous vehicle to 'See', to move about and to interact intelligently with its environment. The resulting vision, scene interpretation, and motor control processes will be, it is anticipated, at the very least analogous to those found in lower animals.

To develop an autonomous land vehicle with the capabilities envisaged, requires an expert system for navigation as well as a sophisticated vision system. The expert navigation system must plan routes using digital terrain and environmental data, devise strategies for avoiding unanticipated obstacles, estimate the vehicle's position from other data, update the on-board digital terrain data base, generate moment-to-moment steering and speed commands, and monitor vehicle performance and on-board systems. All these functions must be accomplished in real-time to near-real-time while the vehicle is moving at speeds upto 60 km/h. Scaling up from laboratory experiments indicates that such an expert system would require of the order of 6500 rules firing at a rate of 7000 rules/sec. Current systems contain approximately 2000 rules and fire at a rate of 50–100 rules/sec.

The vision system must take in data from imaging sensors and interpret these data in real-time to produce a symbolic description of the vehicles environment. It must recognise roads and road boundaries; select, locate and dimension fix moving obstacles in the roadway; detect, locate and classify objects in open or forested terrain; locate and identify man-made and natural landmarks; and produce thematic maps of the local environment; while moving at speeds upto 60 km/h. Such a system would require computing capabilities of 10-100 GIPS. This compares with the capabilities, for example, of 30-50 MIPS in today's most powerful Von Neumann type of computers.

Of equal importance with these required computing capabilities is the weight, space and power required. For a land reconnaissance vehicle, for example, the computers should occupy no more than about a cubic meter, should weigh less than 250 kg, and should consume less than 1 kW of power, including environmental support. The requirements represent at least 1-4 orders of magnitude reduction in weight, space and power over today's computing systems. For aerospace and undersea autonomous vehicles, the constraints and requirements will be tighter and need to include the capability to operate in high radiation environments.

As a foretaste of what is to come, in June 1984 the dummy warhead of the Minuteman I ICBM, launched from Vandenberg Air Force Base in California, was intercepted, and destroyed high over the Pacific Ocean by an optically guided missile launched from another test range. A long wave length infrared sensor in the interceptor was activated after the second stage booster rocket had been fired. With the aid of the computers, processing data at some 18 MIPS, the interceptor homed in on the dummy warhead and destroyed it while it was some 160 km above ground. This was a convincing demosntration of the power of sophisticated imaging guidance system in a space environment and definitely the beginning of 'Star Wars'.

BIBLIOGRAPHY

1. Andrews, Harry C., *IEEE Spectrum,* **16** (4), (1979), 38-49.
2. Landgrebe, David A., *IEEE Proc.* **69** (5), (1981), 1236-1256.
3. Billingsley, F.C., Remote sensing: A view of integration, *In* Digital Image Analysis, S. Levialdi, (Ed.), (Pitman Publishing Inc., London), 1984, p.54.
4. Nagy, G., A critical analysis of remote sensing technology, *In* Digital Image Analysis, S. Levialdi, (Ed.), (Pitman Publishing Inc., London), 1984.
5. Cobeman, G.B. & Andrews, H.C., *IEEE Proc.,* **67** (5), (1979), 723-748.
6. Barrow & Tannenbaum, *IEEE Proc.,* **69** (5), (1981), 1272-1273.
7. Pavlidis, Theo, Algorithms for Graphics and Image Processing, (Prentice Hall, Englewood Cliffs), 1986.