International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977, pp. 635-641.

[40] H. Barrow and J.M. Tenenbaum, "Recovery of Intrinsic Scene Characteristics from Images," in *Computer Vision Systems*, A. Hanson and E. Riseman, (eds.), Academic Press, New York, 1978, pp. 3-26.

[41] R. J. Woodham, "Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings" MIT Artificial Intelligence Laboratory Memo AI-TR-457, June 1978.

[42] B. K.P. Horn and B. L. Bachman, "Using Synthetic Images to Register Real Images with Surface Models," *Communications of ACM*, Vol. 21, No. 11, November 1978, pp. 914-924.

[43] J. J. Gibson, *The Perception of the Visual World*, Houghton Mifflin, Boston, Mass., 1950.

[44] K. Stevens, "Surface Perception from Local Analysis of Texture and Contour," MIT Artificial Intelligence Laboratory Memo AI-TR-512, February 1980.

[45] R. Bajcsy, "Computer Identification of Visual Surface," *Computer Graphics and Image Processing*, Vol. 2, No. 2, 1973, pp. 118-130.

[46] R. Bajcsy and L. Liebermann, "Texture Gradient as a Depth Cue," *Computer Graphics and Image Processing*, Vol. 5, No. 1, March 1976, pp. 52-67.

[47] A. P. Witkin, "A Statistical Technique for Recovering Surface Orientation from Texture in Natural Imagery," *Proceedings of the First Annual National Conference on Artificial Intelligence*, Stanford, Calif., August 1980, pp. 1-3.

[48] J. R. Kender, "Shape from Texture," Carnegie-Mellon University, Computer Science Technical Report CMU-CS-81-102, Pittsburgh, November 1980.

[49] T. Kanade, "Recovery of the Three-Dimensional Shape of an Object from a Single View," Carnegie-Mellon University, Computer Science Report CMU-CS-79-153, Pittsburgh, October 1979.

[50] D. G. Lowe and T. O. Binford, "The Interprettion of Three-Dimensional Structure from Images," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, Vancouver, Canada, August 1981, pp. 613-618.

[51] T. O. Binford, "Inferring Surfaces from Images," *Artificial Intelligence*, Vol. 17, 1981, pp. 205-244.

10

SYSTEMS AND APPLICATIONS

Visual perception systems may be "general purpose" or tailored for special tasks. A general-purpose system is expected to have capabilities similar to the human visual system and handle a wide variety of scenes under a variety of viewing conditions. Under normal viewing conditions, humans may expect of certain objects to be present in the scene, but our ability to perceive seems to be almost as good when we are presented with a randomly chosen photograph. We are able to generate high-quality descriptions of unfamiliar objects, such as pictures of new planets, or photomicrographs of molecules. General-purpose vision may be defined to have capabilities similar to human perception; a more fundamental definition is difficult, owing to the inherent ambiguity of the images.

Our understanding of the perceptual processes needed to achieve general vision is poor, and the performance of the techniques discussed in the previous chapters is low in comparison to human performance. Fortunately, a great many applications of practical importance do not require this generality, as the domain of objects is often small, and significant knowledge of the scene is available a priori. Special-purpose *knowledge-based systems* aim to maximize the utilization of such knowledge.

In this chapter we examine some requirements for a general-purpose system and describe some knowledge-based systems with applications.

10.1 GENERAL SYSTEMS

As visual information is inherently ambiguous, some knowledge and assumptions are required for its interpretation. These assumptions may occasionally lead to incorrect conclusions, but the human system's perfomance is amazingly accurate in almost all instances in our daily experience. General-purpose and knowledge-based systems differ not only in the range of objects they encounter, but also in the type of knowledge they use. The systems attempting to be general tend to use *generic* rather than *specific* knowledge. Generic knowledge includes restrictions due to physical phenomena, such as surface reflectivity, continuity, object coherence, and support requirements. Specific knowledge refers to the knowledge of particular objects likely to be present in the scene, their properties, and the specific viewing conditions. Also, the general systems tend to defer use of specific knowledge, whereas the knowledge-based systems tend to utilize such knowledge early in the processing hierarchy. A general-purpose system needs at least the following abilities:

1. To perceive lightness and color of surfaces under a variety of illumination conditions;

2. To detect significant changes in intensity and perform 2-D segmentation into useful regions, even in the presence of texture;

3. To infer 3-D structure of the surfaces of a scene from a variety of monocular cues and also from a sequence of stereo or motion images;

4. To organize the surfaces and regions into objects of interest;

5. To generate descriptions of objects and recognize them among a potentially large class of objects; and,

6. To make nonvisual intelligent inferences about the scene based on the visual processing.

The above processes may be considered to form a hierarchy of abstraction levels. Two methods of data flow among these levels are *bottom-up* and *top-down.* In bottom-up processing, the information flows from one level to the next higher level without any influence from the expectations of this higher level. In top-down control, the processing at a lower level is specifically directed to satisfy an expectation or goal of a higher level. At the highest level, the goal may be to verify if a certain object is present in the scene. It seems that our ability to perceive unexpected or unfamiliar scenes requires capabilities of bottom-up

processing, to the level of meaningful object descriptions. However, extensive communication between the various levels is needed, hence the processing is not strictly bottom-up. Humans are capable of top-down processing also, as indicated by our ability to see a suggested object in an otherwise confusing scene (for example, a suggested pattern in the clouds in the sky).

There are currently no systems even approaching the level of general-purpose performance of the human system. However, the desire for future generality leads to very different design strategies than those for knowledge-based systems for specific tasks, as discussed below.

10.2 KNOWLEDGE BASED SYSTEMS

Consider the typical office scene as shown in Fig. 10-1. A general system might proceed by attempting to segment the image using edge and/or region methods, possibly aided by range information to describe the objects and surfaces by the chosen representations. It is likely that the scene is too complex for resulting descriptions to be directly in terms of the objects that we perceive. However, specific tasks can be performed if scene knowledge is utilized judiciously. Consider the sample task of locating the telephone on the table. We could search the image for a region of known properties (known color and approximate size). However, it may be more efficient to locate the table top first, which is easily located if range data is available, and constrain the search for the telephone to the table top.

Tenenbaum has described a system to perform such tasks [1]. The knowledge in his system consists of the properties of the objects, such as color, size, and shape, and their relationships to other objects in the scene, such as the table having to be on the floor and the telephone on the table. The program operates in two phases, an acquisition and a validation phase. Acquisition is based on attributes that are easy to compute and are distinct. Other attributes and relations are used to validate the initial hypotheses.

Garvey extended this approach to automatic generation of a plan to locate objects, given the properties and the relations of the objects [2]. These systems may be viewed since following top-down processing, since each step of low-level processing is to satisfy a specific high-level goal. Such processing is likely to be useful if the scene has small variations and if the properties of the objects and even their locations are known approximately.

Bajcsy and Tavakoli have described a system to locate objects, such as roads, rivers, and bridges, in aerial images using models of

Figure 10-1: A typical office scene (from Tenenbaum [1])

these objects in a scene [3, 4]. Road models at the scene level are inferred from more abstract properties (by the designers and not by the program). A functional definition of a road is a path to allow for the passage of certain vehicles, persons, or animals. From this follow physical and geometrical properties requiring a road to be relatively smooth and firm and to have bounded steepness, width, and curvature. These requirements in turn define image properties such as roads being narrow strips of bounded width and curvature and spectral properties of materials such as concrete, asphalt, and rocks. Also, roads must be connected to other roads or other cultural features. Similarly, a bridge is defined to be over water, but connected to land masses on either side. Note that such models are generic and not for a particular scene.

Nevatia and Price have described a system to locate specific objects, such as airports in aerial images, by first locating larger and easier to locate objects [5]. In the image of the San Francisco area, shown in Fig. 1-9, the San Francisco International Airport can be recognized more easily if it is known to be along the edge of the bay, and south of the city of San Francisco. The city may in turn be located by the bridges, which are distinct. This system uses an approximate model of the scene, in the form of a rough map without precise distance

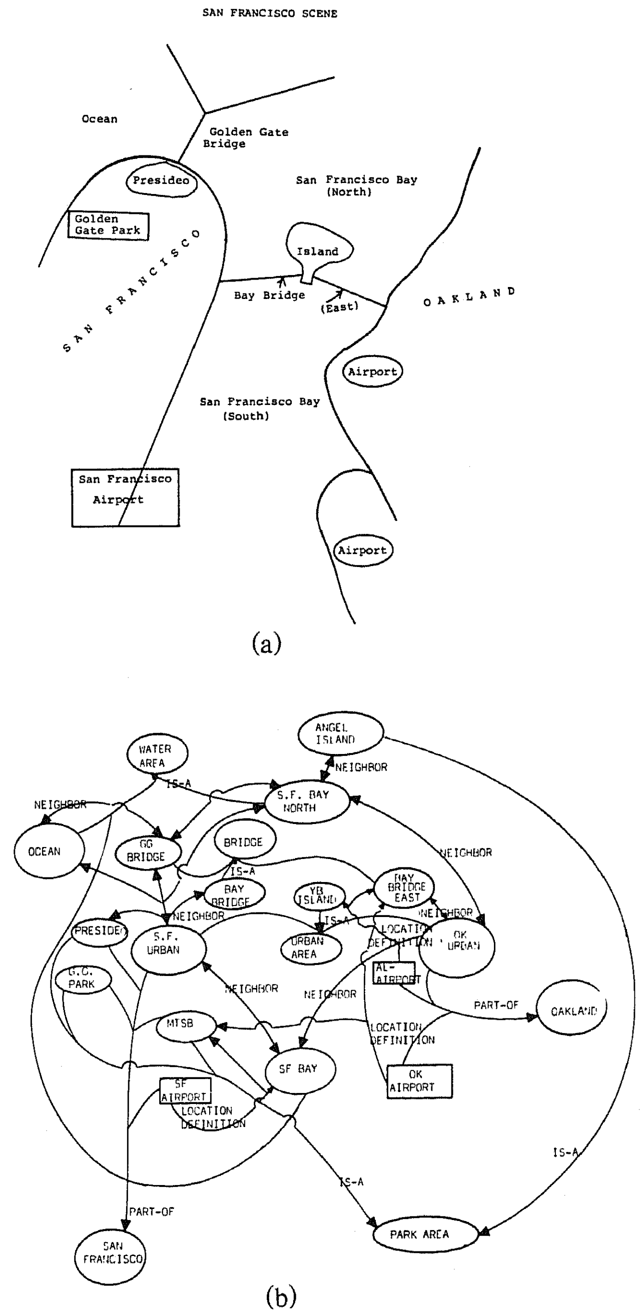

Figure 10-2: (a) A map of the San Francisco area and (b) its graph representation

information as shown in Fig. 10-2(a). Figure 10-2(b) shows the corresponding graph model used in the analysis. Segmentation of the scene uses both region and edge segmentation and is basically bottom-up, but it also employs some model information, such as the maximum width of the roads and bridges, and the features to use for segmentation of water areas. Figure 10-3 shows the segmented regions and roadlike and bridgelike features, with labels associated by matching with the map. Expected locations of the airports can now be estimated from the known features and these locations examined in more detail for verification.

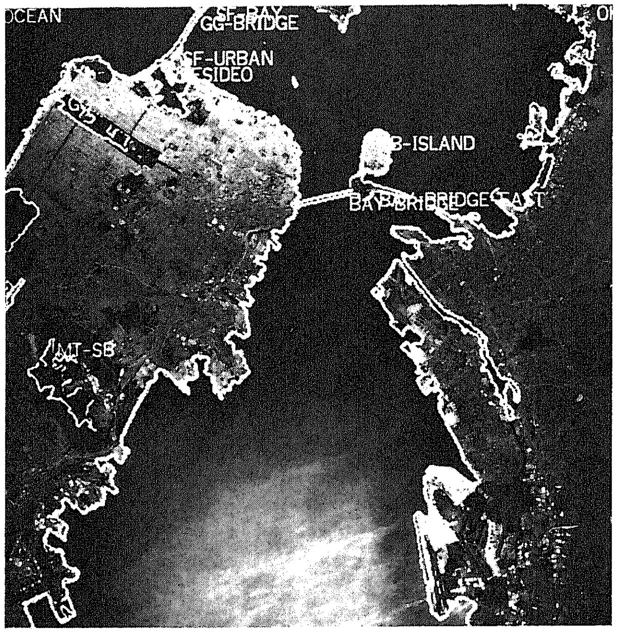

Figure 10-3: Segmented scene with recognized elements

Freuder has proposed that vision systems should utilize knowledge at the earliest levels of processing, even when a specific goal is not specified [6]. Thus, if a region of constant intensity can be identified to be a part of one or a few objects, this knowledge should be used to guide further processing. The main difficulty here is that of *indexing*—that is, of retrieving the set of likely objects based on their low-level descriptions. Without contextual information, indexing requires good high-level descriptions.

A system called ACRONYM having many modules of general utility has been developed at Stanford University [7]. This system has

abilities for top-down predictive processing and also for bottom-up descriptions. At the top level, object models are described as 3-D volumes in an *object graph.* The modeling system is quite general and allows generic as well as specific objects to be described. For a given model geometric reasoning techniques are used to predict feature which will be invariantly observable; that is, features which are visible over a range of viewing conditions. A *prediction graph* is generated whose nodes are predictions of image features and the arcs specify relations which must hold between them. A *picture graph* is generated from the image which describes the features in that image. Generation of picture graph may be guided by inputs from the prediction graph. Interpretation of the scene and recognition of objects takes place by matching between the prediction and the picture graphs. Much of this system is implemented in form of rules and is extensible. ACRONYM has been applied to tasks of complex object recognition; a typical example is recognition of a commercial aircraft parked at an airport terminal, in an aerial image. For this application, the prediction and picture graphs are essentially in terms of "ribbons" which are two-dimensional projections of three-dimensional generalized cones. The major performance limitations of the system are because of the weaknesses of the lower level description modules.

Another system intended for a broad range of scenes, called VISIONS, is described in [8]. The systems referred to here are not the only knowledge-based systems that have been designed, but they suggest the range of tasks that have been approached.

10.3 APPLICATIONS

Potential application areas of a general purpose machine vision system are vast. However, current application tasks must be carefully chosen to match with the current capabilities of machine techniques. The successful applications tend to be for specific and well defined tasks, within a limited domain of objects, and with adequate a priori knowledge. Because of the large and increasing number of application tasks, only the major areas of applications and some typical tasks are described below.

10.3.1 Industrial Applications

A major area of industrial applications is the visual inspection of manufactured parts. The inspection tasks can range from detection of major flaws such as missing parts to detection of small defects, misalignments, size measurement, subtle color changes, and so on. For humans, these tasks tend to be dull, and their repetitiveness leads to decreased performance.

Successful practical applications require fast processing, inexpensive hardware, and high reliability. To simplify the image-analysis problems, lighting may be controlled to give high-contrast images. Back lighting or fluorescent conveyor belts may enable thresholding to provide satisfactory segmentation. Special lighting may also be required to make visible some defects, such as cracks in a glass seal [9].

Major successes have been achieved in the inspection of electronic printed circuit boards (PCBs) and integrated circuits (ICs). A direct approach is to compare the images of the patterns to stored images of defect-free patterns on a pixel by pixel basis. However, difficulties are caused by the variability of the images, alignment errors, changes in size, and so on. Another approach is to examine small neighborhoods, such as 5-by-5 or 7-by-7 binary windows, and classify the patterns as defective or not [9]. Ejiri and others developed an interesting system to detect small defects, defined to be patterns with widths less than the board conductors. These defects appear as thin blobs of extra or missing metal [10]. A region of the image is first expanded and then shrunk by the same amount. This removes convex blobs of a certain size. The processed pattern is then compared with the original to detect the convex defects. Concave defects are detected by first shrinking and then expanding.

Baird describes a system developed at General Motors for orienting IC chips correctly before they are bonded [11]. The orientation is basically determined by a histogram of edges detected in the image. Other systems for orienting IC chips are described in [12, 13]. Some of these systems are reported to be in large scale production use.

Agin has described laboratory experiments at detection of flaws in castings such as missing or incorrectly dimensioned holes, and inspection of other industrial objects [14]. The orientation of the parts is constrained so that the perspective variations are avoided.

Another important area of applications is in materials handling. Here, the parts are usually in a heap or a bin with other similar or different parts. A part needs to be identified and grasped by a mechanical manipulator at the appropriate points. Partial success has

been obtained, if the possible orientation of the parts can be restricted, and they are unoccluded or else occluded in small areas only [15] (a typical scene is shown in Fig. 1-7).

Visual feedback for automated assembly is made difficult by the complexity and cluttered nature of the scenes. Some simplification can be obtained by marking objects by patches of specific shape or color (for example, see [12, 16]). Use of more elaborate markings is limited by visibility and space requirements. We can also take advantage of the slow and predictable variations in the scene for incremental analysis (see [17]).

More examples of industrial applications may be found in [18, 19].

10.3.2 Photo Interpretation and Change Detection

Images taken from airplanes or orbiting satellites provide a rich source of information for monitoring changes on the surface of the earth. The applications include surveying crops, forests, pollution, and other natural resources, military surveillance and monitoring of new construction of roads and other man-made structures, and automatic mapping from the images. Manual analysis of aerial images is not only tedious and labor intensive, but also the large volume of data transmitted to earth causes a major communications bottleneck that could be eliminated by on-board processing.

Image analysis of aerial scenes is more complex than analysis of industrial scenes, owing to the presence of fine texture and a variety of objects in a single scene, and the requirements of high resolution for observing the fine details. However, the scenes are largely two-dimensional with little occlusion, even though the effects of shadows and presence of mountainous terrain can be significant. Generally, unguided segmentation of aerial images is error prone. Some regions—for example, uniform water areas such as lakes— can be extracted easily and reliably.

A simple processing approach, commonly used for crop identification, is classification of each pixel by its multispectral properties (the sensors may include infrared, radar, and so on in addition to the visual sensors). This approach suffers from its ignoring of the context that helps in distinguishing between ambiguous cases. Other applications have concentrated on the development of specific techniques for the extraction of specific features, such as roads and railroads [20].

Some photo interpretation tasks are extremely labor intensive and tedious; for example, making maps from aerial images involves manual tracing of small features with high accuracy. For such applications, partial automation with interactive human control would still result in large savings. Some instances of interaction are: correcting machine errors, pointing to certain examples (such as, a central area of a region is to be extracted and its properties used to extract the larger region), and following a road if its intensity profile is initially traced interactively. Descriptions of some implementations are given in [20-22].

Change detection requires comparison of two regions, taken at two different times, and possibly with different sun angles and weather conditions, different viewing angles, and maybe even different sensors. Changes can be detected by simple grey-level correlation of the two images if the changes caused by these factors are small [23]. For larger changes, matching at the symbolic level, by first extracting features such as uniform regions and roads, has been more successful [24]. Again, the main limitation is in computing adequate symbolic descriptions for complex scenes. Change detection may also be aided by the use of a previous map of the scene.

10.3.3 Guidance, Navigation, and Scene Registration

A major use of vision by humans and animals is for navigating in the surrounding environment, avoiding obstacles, and reaching desired locations. Much of the current automatic navigation of machines, such as commercial aircraft, is by the use of special navigation sensors, such as radar, located along the desired route. Some applications—for example, exploration of distant planets —require navigation without modification of the environment and with little or no human intervention, and without a priori knowledge of the terrain.

For surface vehicles, a major concern is to avoid obstacles and hazards in the path of a moving vehicle. For space applications, the obstacles are rocks, craters, loose surface materials, and the like. Experimental studies for a lunar rover have been conducted at the Jet Propulsion Laboratories. Moravec constructed a vehicle that navigated in simple indoor and outdoor scenes and avoided obstacles perceived by a stereo vision system [25]. Roving vehicles in urban environments have a more complex task, owing to the large variety of the objects in the scene and other moving objects. Specially prepared roadways may be of help here.

Following a chosen trajectory is of major concern for airborne vehicles. Here, a "map" of the trajectory and surrounds is available, either in symbolic form or as a sequence of images along the path. Because of the variability in the images under different flight conditions, simple grey level correlation is unlikely to be effective. More success has been obtained by correlating the terrain elevation profiles.

Processes using matching of features extracted from a perceived image with features extracted from stored images are likely to be more tolerant to changes in images. One implementation for matching scenes having different scale and some perspective changes, taken from possibly different sensors, with different seasonal and viewing conditions, using lines extracted from the two scenes is described in [26]. Surveys of various approaches may be found in [27].

A human pilot, flying under visual guidance, navigates by locating specific, distinguished features on the earth and relating them to the map symbols. Some of the above mentioned techniques can also be used for such *image-to-map correspondence* (for example, see [5, 24, 26]). Another interesting iterative technique is described in [28]. Here, selected features in a map, with 3-D data, are projected onto the image, using an estimate of the camera position. These projected points are matched to the closest similar points in boundaries extracted from the image. The new matches now define a new camera transform, and the process is repeated until the projected points and the matched points are within an acceptable range.

10.3.4 Medical Applications

Many medical diagnostic procedures use images—for example, chest and other x-rays, microscope photographs of blood cells, acoustic images of various organs, and three-dimensional "images" of organs obtained by computer tomography. Highly trained medical personnel are required for interpretation of such images, and shortage of such people is a deterrent to a more widespread use of these techniques.

Medical images have some significant characteristics. These images frequently use illumination that penetrates the objects; that is, the object surfaces are not opaque. The defects to be detected may be small and subtle, sometimes characterized by smooth changes in the grey levels only (as for bone structures). Also, the allowed margin of error is small, owing to the potentially serious consequences of an incorrect diagnosis. However, significant a priori information is available in many cases from human anatomy, and viewing conditions can be controlled.

Analysis of chest x-rays has received considerable attention. A method for the diagnosis of pneumoconiosis (coal miner's disease) by an analysis of the texture of dark blobs in the x-rays is described in [29]. Other x-ray analysis techniques are described in [30-31]. For chest x-rays, it is useful to be able to first isolate ribs and other bone structures, as certain types of tumors do not occur there. However, the ribs are typically of low contrast and only partially visible (see Fig. 1-8).

Knowledge of the approximate shape and relative locations of the ribs has been used for their detection [32-34].

Analysis of blood-cell images and human chromosomes for genetic defects has also been popular. In some cases, as for chromosome analysis, new biochemical techniques can simplify or replace visual techniques.

Use of computers to obtain three-dimensional data of brain and other organs, by use of multiple two-dimensional views, has been hailed as a major advance in the diagnostic procedures. Much of this work has concentrated on rapid and accurate reconstruction of the three-dimensional data from the two-dimensional views, rather than the automatic analysis of the resulting 3-D data.

The area of medical applications has grown so that several symposia are entirely devoted to them. A good source of current progress is the proceedings of these symposia (for example, see [35]).

10.3.5 Hardware Requirements

The large computational requirements of visual processing are a constraint on the range of practical applications. The resolution of the images is limited to far less than that of human vision or the optical resolution of an average camera. The complexity of the processing algorithms must be limited to meet response time requirements. The more sophisticated algorithms take several minutes (or hours) to run on modern large computers (executing 1 to 5 million operations per second). The continuing increase in the speed of generral-purpose computers, with simultaneous reduction in cost and size, will be helpful in the future. However, it is unlikely that complex high-resolution visual processing will be feasible by using sequential, general-purpose machines alone.

Fortunately, the structure of visual processing is suited to *parallel* implementation. Much of the processing time is taken for simple, low-level processes, such as convolution for edge detection and histograms for region threshold selection. Usually at the higher levels of processing the amount of data to be processed decreases dramatically, hence the processing time decreases even though the operations on each data are more complex. The lower-level processing is usually simple, uses information from small neighborhoods only, and hence can be easily implemented using parallel processors. Some efforts at developing parallel machines for visual processing are described in [36, 37].

Simplicity of the lower-level processes also allows for their implementation in special-purpose hardware. Algorithms such as edge

detection, histogramming, and thresholding have been implemented on single chips using CCD (charge coupled devices) technology and operate at or near TV frame rates [38-40]. The deployment of the VLSI (very large scale integrated circuits) technology should have a major impact on the complexity of algorithms that can be implemented in a few IC chips [38, 41-42].

10.4 SUMMARY AND FUTURE

This chapter has outlined the state of the art in applications of machine perception techniques and the range of the tasks to which they are applied. Successful applications have been where the tasks are well defined and the domain is restricted. The simplicity of the algorithms currently used, in comparison to the known techniques, and the rapid advances in the hardware technology of general-purpose computers and special-purpose devices assures evolutionary progress in the use of machines to perform more and more complex tasks. The science fiction fantasy (or fears) of "super-human" robots replacing and outperforming humans in complex perceptual tasks remains far from reality.

REFERENCES

[1] J. M. Tenenbaum, "On Locating Objects by Their Distinguishing Features," *Computer Graphics and Image Processing*, Vol. 2, 1973, pp. 308-320.

[2] T. Garvey and J. M. Tenenbaum, "On the Automatic Generation of Programs for Locating Objects in Office Scenes," in *Proceedings of the Second International Joint Conference on Pattern Recognition*, Copenhagen, August 1974, pp. 162-168.

[3] R. Bajcsy and M. Tavakoli, "Computer Recognition of Roads from Satellite Imgaes," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 6, No. 9, September 1976, pp. 623-637.

[4] R. Bajcsy and M. Tavakoli, "Image Filtering—A Context Dependence Process," *IEEE Transactions on Circuits and Systems*, May 1975, pp. 463-474.

[5] R. Nevatia and K. Price, "Locating Objects in Aerial Images," *Proceedings of the Fourth International Conference on Pattern Recognition*, Kyoto, Japan, November 1978, pp. 686-690.

[6] E. Freuder, "A Computer System for Visual Recognition Using Active Knowledge," MIT Artificial Intelligence Laboratory Memo AI-TR-345, June 1976.

[7] R. A. Brooks, "Symbolic Reasoning Among 3-D Models and 2-D

Images," Stanford Artificial Intelligence Laboratory Memo AIM-343, June 1981.

[8] A. R. Hanson and E. M. Riseman, "VISIONS: A Computer System for Interpreting Scenes," in *Computer Vision Systems*, A. R. Hanson and E. M. Riseman (eds.), Academic Press, New York, 1978, pp. 303-333.

[9] J. F. Jarvis, "Visual Inspection Automation," *IEEE Computer*, May 1980, pp. 32-39.

[10] M. Ejiri, T. Uno, M. Mese and S. Ikeda, "A Process for Detecting Defects in Complicated Patterns," *Computer Graphics and Image Processing*, Vol. 2, 1973, pp. 326-339.

[11] M. L. Baird, "SIGHT-1: A Computer Vision System for Automated IC Chip Manufacture," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 8, February 1978, pp. 133-139.

[12] M. Yachida and S. Tsuji, "Industrial Computer Vision in Japan," *IEEE Computer*, May 1980, pp. 50-64.

[13] B. K.P. Horn, "A Problem in Computer Vision: Orienting Silicon Integrated Circuit Chips for Lead Bonding," *Computer Graphics and Image Processing*, Vol. 4, No. 3, September 1975., pp. 294-303.

[14] G. J. Agin, "Computer Vision Systems for Industrial Inspection and Assembly," *IEEE Computer*, May 1980, pp. 11-20.

[15] W. A. Perkins, "A Model-Based Vision System for Industrial Parts," *IEEE Transactions on Computers*, Vol. 27, 1978, pp. 126-143.

[16] A. Gill, "Visual Feedback and Related Problems in Computer Controlled Hand Eye Coordination," Stanford Artificial Intelligence Laboratory Memo AIM-178, October 1972 (Ph.D. thesis).

[17] R. Bolles, "Verification Vision for Programmable Assembly," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, Mass., August 1977, pp. 569-575.

[18] W. Thompson (ed.), Special Issue on Machine Perception, *IEEE Computer*, Vol. 13, No. 5, May 1980.

[19] G. G. Dodd and L. Rossol (eds.), *Computer Vision and Sensor-Based Robots*, Plenum Press, New York, 1979.

[20] H. G. Barrow, "Interactive Aids for Cartography and Photo Interpretation," SRI Technical Report 5300, Menlo Park, Calif., November 1976.

[21] L. H. Quam, "Road Tracking and Anomaly Detection in Aerial Imagery," *Proceedings of Image Understanding Workshop*, Cambridge, Mass., May 1978, pp. 51-55.

[22] J. M. Tenenbaum and H. G. Barrow, "Interpretation Guided Segmentation," *Artificial Intelligence*, Vol. 8, 1977, pp. 241-274.

[23] R. L. Lillestrand, "Techniques for Change Detection," *IEEE

Transactions on Computers, Vol. 21, July 172, pp. 654-659.

[24] K. Price and D. R. Reddy, "Matching Segments of Images," *IEEE Pattern Analysis and Machine Intelligence*, Vol. 1, No. 1, January 1979, pp. 110-116.

[25] H. P. Moravec, "Obctacle Avoidance and Navigation in the Real World by a Seeing Robot Rover," Stanford Artificial Intelligence Laboratory Memo AIM-340, September 1980.

[26] C. S. Clark, W. O. Eckhardt, C. A. McNary, R. Nevatia, K. E. Olin, and E. M. Van Orden, "High-accuracy Model Matching for Scenes Containing Man-made Structures," in [27], pp. 54-62.

[27] *Proceedings of Symposium on Digital Processing of Aerial Images*, SPIE, Vol. 186, 1979.

[28] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Cambridge, Mass., August 1977, pp. 659-663.

[29] R. P. Kruger, W. B. Thompson, and A. F. Turner, "Computer Diagnosis of Pneumoconiosis," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 4, 1974, pp. 40-49.

[30] H. Ballard and J. Sklansky, "A Ladder Structured Decision Tree for Recognizing Tumors in Chest Radiographs," *IEEE Transactions on Computers*, Vol. 25, 1976, pp. 503-513.

[31] C. A. Harlow and S. A. Eisenbeis, "The Analysis of Radiographic Images," *IEEE Transactions on Computers*, Vol. 22, July 1973, pp. 678-689.

[32] H. Wechsler and J. Slansky, "Finding the Rib Cage in Chest Radiographs," *Pattern Recognition*, Vol. 9, 1977, pp. 21-30.

[33] D. H. Ballard, "Model-directed Detection of Ribs in Chest Radiographs," *Proceedings of the Fourth International Joint Conference on Pattern Recognition*, Kyoto, Japan, November 1978, pp. 907-910.

[34] J. Toriwaki, Y. Suenaga, T. Negoro, and R. Fukumura, "Pattern Recognition of Chest X-ray Picture Patterns," *Computer Graphics and Image Processing*, Vol. 2, December 1973, pp. 252-271.

[35] *Proceedings of Symposium on Computer-Aided Diagnosis of Medical Images*, Coronado, Calif., November 1976 (IEEE Computer Society Publication 76CH1170-0C).

[36] G. R. Allen and P. G. Juelten, "SPARC-Symbolic Processing Algorithm Research Computer," *Proceedings of Image Understanding Workshop*, Palo Alto, Calif., April 1979, pp. 171-174.

[37] C. Rieger, "ZMOB: A Mob of 256 Cooperative 280-A Based Microcomputers," *Proceedings of Image Understanding Workshop*,

Los Angeles, Calif., November 1979, pp. 25-30.

[38] G. Nudd, "Image Understanding Architectures," *Proceedings of the National Computer Conference*, Anaheim, Calif., 1980, pp. 377-390.

[39] G. R. Nudd, G. D. Thurmond, and S. D. Fouse, "A Charge Coupled Device Image Processor for Smart Sensor Applications," *Proceedings of the SPIE Symposium*, San Diego, August 1978, pp. 15-22.

[40] T. J. Willett and G. E. Tisdale, "Hardware Implementation of a Smart Sensor: A Review," *Proceedings of Image Understanding Workshop*, Cambridge, Mass., May 1978, pp. 1-8.

[41] T. J. Willett, A. R. Helland, and G. E. Tisdale, "Higher Level Algorithms: Evaluation and Implementation," *Proceedings of Image Understanding Workshop*, Los Angeles, Calif., November 1979, pp. 15-24.

[42] W. L. Eversole and D. J. Mayer, "Investigation of VLSI Technologies for Image Processing," *Proceedings of the Image Understanding Workshop*, College Park, Md., April 1980, pp. 182-189.

INDEX

Machine Perception

Ramakant Nevatia

Ramakant Nevatia, in his **Machine Perception**, explores the growing field of visual perception. The potential applications for such systems include tasks like automation of industrial processes of inspection and assembly, automated medical x-ray diagnosis, vehicle guidance, and automatic photo-interpretation, and have generated a fast-growing interest in this field.

Among its many highlights:

- a comprehensive coverage of the field, assuming familiarity with digital programming but no previous knowledge of the field
- a general grouping of techniques by their common themes with descriptions of basic concepts, details of major approaches, and a guide to the literature
- numerous illustrations including examples of typical systems
- extensive bibliography and a summary concluding each chapter

Chapter headings: Introduction; Pattern Classification Methods; Simple Polyhedral Scenes; Complex Scenes of Polyhedra; Shape Analysis and Recognition; Perception of Brightness and Color; Edge and Curve Detection; Region Segmentation and Texture Analysis; Depth Measurement Analysis; Knowledge Based Systems and Applications.

PRENTICE-HALL, Inc., Englewood Cliffs, N.J. 07632

ISBN 0-13-541904-2