

ARTIFICIAL INTELLIGENCE LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Working Paper No. 157

May 1978

VISION REVIEW

Berthold K. P. Horn

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643.

Underestimating the Difficulties

To us, vision is an immediate experience, not subject to careful introspection. We cannot write down protocols of processing steps that lie between the raw image intensities and our vivid impression of the surrounding scenery. Furthermore, we find ourselves in possession of this faculty long before we learn to master the more sequential processing tasks involved in the use of language, for example. Consequently, the difficulties of the vision process are often not appreciated. This is as true today, when many inroads have been made on the problem of understanding this process, as it was earlier when it was thought that vision could be understood simply in terms of some general ideas of artificial intelligence.

Divergence of Objectives

This difficulty is further compounded by the fragmentation of efforts resulting from widely varying motivations which bring researchers to this problem. These range from an intense desire to understand naturally occurring vision systems to an interest in industrial application of machine vision. Somewhere in between we find those pursuing the information inherent in an image without regard to the implementation details of particular vision systems. It is not too surprising then that one finds widely diverging criteria for judging the importance of a particular piece of work.

Machine Vision is not the "I/O of A.I."

Still others use the vision domain only as a test bed to illustrate some general mechanisms currently favoured in other artificial intelligence work, or think of machine vision and manipulation merely as the "I/O of A.I.", the interfaces which allow the smart machine to interact intelligently with its environment. I contend that this is unreasonable since vision appears to have interesting features which do not have counterparts elsewhere, certainly not in the serial, linguistic kind of reasoning pursued in other areas of artificial intelligence. It is these features which make vision worthy of study in its own right. One such aspect which has not been approached seriously elsewhere is that of spatial reasoning which cannot be conveniently handled using the kinds of data structures explored so far and found so useful in other domains.

Task of a Vision System

What is the task to be tackled by a vision system? Despite widespread disagreement on many other aspects of vision, most would agree that a vision system is expected to produce a description of what is being viewed. The input may be one or more images, each a two-dimensional distribution of scene radiance values obtained from some sensing device. There is less agreement on the form of the output. What kind of description is acceptable? Clearly two criteria must be satisfied: The description must

- (1) reflect some aspects of the three-dimensional reality, and
- (2) be useful in carrying out a specified task.

Usually it is expected that the description take a symbolic form. Quite different kinds of descriptions are likely to be considered adequate when the system is part of a device which lines up integrated circuit chips for automated lead bonding [Horn, 1975b], as opposed to a stage of a system meant to express an opinion about the merits of a work of art. As is so common, the representation of the given information must be matched to the task at hand.

Task Independence

It would be much nicer if one common mode of description could be employed,

since the system dealing with visual inputs then could be designed in isolation, without considering the overall task. Perhaps this will turn out to be possible, at least for early stages of an image analysis system. Certain kinds of operations on images appear to be dictated by the image rather than the task and ought to be done without consideration for the task. At this point, however, it seems that task-dependent representations will be with us for a while.

Complexity of Machine Vision

An important lesson to be learned from the work on vision so far is that the problems are profound and not likely to succumb to the application of a bag of tricks from some other field, such as communications theory, statistics or linear systems theory. Machine vision merits its own methodology. Attention must be paid to both the physics of image formation as well as the information processing techniques which can produce the desired internal descriptions of what is being viewed. It now seems that this latter endeavour can be helped along by careful consideration of the human visual system and its strength and weaknesses.

Unfortunately biological vision systems are extremely complex and one can easily be led astray while studying them in isolation, without adequate tests of hypothesis one develops. Similarly, knowledge of certain physiological or chemical detail, for example, may not turn out to be very illuminating. What happens to the conformation of the rhodopsin molecule in the first few pico-seconds after a photon hits it is very exiting, but not helpful in the understanding of vision in a broader sense. One thing one learns quickly from even casual study of natural vision systems is that prodigious amounts of computation are involved in the processing of the image information.

A Peculiar Dichotomy

For a mobile biological entity above a certain size, vision is vital. It is hard to survive in a world where others have this faculty and use it in competition for food and in predator avoidance. Similarly, machine vision holds great promise for artificial systems. Many tasks cannot be done, or can be done only slowly or clumsily without it. So, vision, while difficult, is also very useful. As a result people will push the technology hard to get working systems. This has resulted in a peculiar dichotomy. There are two kinds of systems:

- (1) systems which work at reasonable speeds, and
- (2) systems which work reasonably well.

"Automated" Stereo

Illustrations of this curious phenomena abound. There are, for example, a variety of machines which extract topographic information at reasonable speeds from stereo pairs of aerial photographs. These devices use special purpose hardware to implement rather simple correlation techniques, and, as a result, require significant human assistance. First, the operator is obliged to help the system out of "trouble spots" where the correlation technique fails because either there is no detail, as on smooth sand or a lake, or because the two views are too different, perhaps because the slope is large. Many times the machine does not even note that it is in trouble and so records bad information. These "glitches" then have to be tediously removed in an interactive editing process if the data is to be at all useful.

On the other side of the coin, one finds many good ideas in the machine vision community which require sophisticated hardware and software for their implementation and which are slow on computers of today's ilk. Methods recently developed at Stanford [Quam 1971, Gennery 1977, Arnold 1978] and at M.I.T. [Marr 1974, Marr & Poggio 1976, Marr & Poggio 1977] are considerably more robust, but require staggering computing power on machines of standard architecture.

"Automatic" Terrain Classification

Systems have been developed for classification of terrain based on the application of pattern recognition techniques on a point by point basis. Special hardware has even been built to implement this simple process, perhaps prematurely, since the performance of this method leaves much to be desired. The classifier has to be trained anew for each image, it cannot deal with hilly terrain and changes in the lighting angles. Typically, the classifier is only used as a step in an iterative refinement process with the human operator making the real decisions. Even then many points are incorrectly classified if the separation between classes is not very distinct.

Yet at the same time work at Purdue [Landgrebe 1973, Landgrebe 1975, Gupta et al 1973, Swain 1973] and the University of British Columbia [Starr & Mackworth 1978],

has demonstrated the advantages of several methods for including contextual information. Growing small regions of similar spectral signature and classifying the regions, rather than individual points helps, as does the use of even rather primitive textural measures [Bajcsy 1973]. Amongst several other promising ideas are those recently expounded at the University of Maryland [Rosenfeld 1977b, Rosenfeld 1978] regarding the use of relaxation methods, also known as cooperative computation methods. While all of these methods produce results far superior to those generated by the point by point methods, it must be admitted that they require considerably more computing power.

Line-finding

As a last example we may look at edge-detection and line-finding. Many fast systems, some even running at full video speeds [Nudd 1978], use simple operations such as Robert's gradient, discrete approximations to the Laplacian operator or Sobel gradients. These produce visually pleasing result, but the edge fragments produced tend to be too noisy and ill-defined to succumb to concerted efforts to glue them into reasonably continuous lines and well-defined vertices.

Systems which do produce useable symbolic edge information such as those developed at M.I.T. [Griffith 1970, Horn 1971, Shirai 1975, Marr 1976, Marr 1978] require vast amounts of computing power both in terms of storage and machine cycles. Very similar sorts of things can be said about approaches which depend on scene segmentation using region growing techniques instead of edge-finding [Brice & Fennema 1970, Ohlander 1975, Tenenbaum & Barrow 1977].

Vision hardware

Part of the explanation for this dichotomy then lies in the impatience of the implementers and the real world need for solutions to pressing problems involving processing of visual information. It is natural to think in terms of special purpose hardware suited to particular algorithms. For specialized tasks it is possible to realize one to three orders of magnitude speed-up in processing with affordable special purpose devices. In the past the development of such hardware was perhaps inappropriate since no one had enough confidence in any particular scheme to commit resources to an implementation effort. Also, the existence of fast systems that use very

simple methods, has discouraged further work, since "the problem has been solved". A look at the results quickly convinces one that this is not so.

Roots

Several fields may be identified as having contributed major ideas to machine vision. I will single out just three here for discussion.

- (1) Image Processing
- (2) Pattern Recognition
- (3) Scene Analysis

Each field has now matured sufficiently to have its basic tools documented in a number of monographs, collections and text books [Andrews 1970, Biberman 1973, Gonzalez & Wintz 1977, Huang 1975, Lipkin & Rosenfeld 1970, Rosenfeld 1969b, Rosenfeld 1976b], [Cheng 1968, Fu 1974, Fu 1976, Grasselli 1969, Tou & Gonzalez 1974, Watanbe 1969], [Duda & Hart 1973, Hanson & Riseman 1978, Winston 1969, Winston 1977], as well as hundreds of papers. Indeed, I cannot begin to do justice to these here, but instead refer the reader to A. Rosenfeld's excellent bibliographies issued annually [Rosenfeld 1969a, Rosenfeld 1972, Rosenfeld 1973, Rosenfeld 1975, Rosenfeld 1976a, Rosenfeld 1977a]. In order to see where we are and to discern possible future trends we should analyze the strength and weaknesses of each of these paradigms in tackling the basic task we have set out for a machine vision system.

Image Processing

Image processing, as the name suggests, is something one does with images. Herein lies both its strength and its weakness. Of the three fields mentioned, this is the only one which deals with images as input. Indeed the basic operations apply to arrays of raw image intensities. Many useful transfers of ideas to machine vision can be traced to this emphasis. Unfortunately, image processing also produces images as output -- not descriptions. Only in so far as these new, possibly enhanced, smoothed or sharpened images are easier to process are these techniques useful. Since, in the case of image processing, the final product is intended for human viewing, this is rarely the case. Further, one finds an unfortunate emphasis on linear, shift-invariant methods and consequently assorted transform techniques. Such ideas have only played a limited

role in machine vision.

Pattern Recognition

At the core of this field is a method, pattern classification, which is concerned neither with images nor with descriptions thereof. Pattern classification instead deals with the mapping of vectors into integers -- the vectors having components which represent measurements of some entity, the integers denoting the classes to which this entity might belong. This paradigm of feature extraction followed by pattern classification is of interest here, however, because many of its applications have involved features extracted from visual data. Several techniques used in the calculation of the numerical feature values for the classification process have found other applications in machine vision. Much of the sophisticated mathematical paraphernalia used to analyze the pattern classification stage has not.

Scene Analysis

Scene analysis concerns itself with the processing of descriptions of images into more sophisticated, or perhaps more useful, descriptions. In this category one finds much of the blocks-world work on line drawings [Roberts 1965, Clowes 1971, Huffman 1971, Waltz 1975]. As it turns out, obtaining the line drawings in the first place from the raw image information was the more difficult task; in fact, no system produces the perfect descriptions needed by early scene analysis systems [Winston, 1972, Grape 1972, Falk 1972].

More recently, discouraged by the complexity of the distributions of raw image intensities, researchers have turned to methods which exploit prior knowledge about the likely contents of the scene being viewed [Reddy et al 1973, Tenenbaum & Barrow 1976]. In Max Clowes' words: "Vision is controlled hallucination". The image contributes a small "controlling" influence on the vision system's "hallucinations" based on expectations and predictions. Similar ideas have taken hold in other areas such as speech, where researchers despair of dealing with the complexities of the raw acoustic waveform without guidance from various "knowledge sources". There is however an ever-present danger of "controlled hallucination" turning into "hallucination". I think we may have closed our eyes to the raw image for too long.

Things to avoid

The early years of any field tend to be characterized by a wide variety of approaches, many false starts and techniques based on inappropriate analogies. We can learn from these mistakes if we wish. Here are some things to avoid:

- (1) Using a mechanism-oriented approach, instead of a problem-oriented one.
- (2) Applying a known bag of tricks from another field.
- (3) Believing that complexity will automatically give rise to interesting behaviour.
- (4) Hoping that "learning" will provide a boot-strapping mechanism.
- (5) Believing what works in a simple situation can be easily extended to a more complex one.
- (6) Suffering from theorem-envy -- introducing unwarranted mathematical hair.
- (7) Working only on the "interesting" sub-problem -- often not the weakest link.
- (8) Following the latest fad. Create your own instead!
- (9) Taking a random path through a maze of possibilities without explanation.
- (10) Admiring the King's new clothes.

Current Trends

Attempts are being made to apply machine vision methods to so many different problems using so many different methods that it is impossible to give any kind of coherent summary. Furthermore, progress is being made in understanding several

important fundamental issues which cut across the spectrum of applications domains. It seems appropriate to concentrate attention to some of these issues.

Representation of Objects

If the task of the vision system is to produce useful descriptions of the scene being viewed, it is naturally important to pick a good representation for three-dimensional objects. If such a description is then to be used for recognition or in the determination of an object's position and orientation, it must capture information about the shape of the object and its disposition in space. This is an important problem, which does not occur in the processing of two-dimensional patterns such as microscopic image of bio-medical interest or in other areas such as finger-print identification or character recognition. A number of representations are currently being explored. One uses generalized cylinders or cones to approximate parts of objects after segmenting them into suitable pieces [Agin & Binford 1973, Nevatia 1974, Binford 1971a, Hollerbach 1976, Nevatia & Binford 1977, Marr & Nishihara 1977]. The information needed to construct such representations may be obtained by a variety of techniques including laser range finding [Nitzan et al 1977] and stereo disparity calculations.

SPINES

Another representation of the shape of an object uses surface normals or "spines". This was suggested [Horn 1977], as a more appropriate representation than one in terms of elevations above some reference plane [Horn 1975a], in part because surface normals undergo a simpler transformation under rotation. Indeed, human performance on shaded images suggests that we are rather poor at establishing relationships in elevation, but have a pretty good idea about the local surface orientation. Fortunately, methods for determining this kind of information exists, ranging from photometric stereo [Woodham 1977] to the shape from shading algorithm. More recently this representation has been suggested as a half-way step to representation in terms of generalized cones [Marr 1978].

Early Symbolic Description

From the discussion of the roots of machine vision it must be clear to the reader

that the crucial thing missing from all three ancestor fields is the lack of a method which takes one from raw image intensities to symbolic descriptions. Little thought had been given even to the problem of where the appropriate point for this transformation would be. Recent work suggests that the first symbolic description be obtained at an early stage [Marr 1976] of the processing of the visual information. That is, the initial symbolic description contains very many items, each of a rather simple nature. Further analysis is then carried out using symbolic information processing techniques on this initial data base.

This is a considerable departure from vision work in the blocks world, where the first real symbolic description was a complete line drawing. Even then it was clear that this was inappropriate, and crude symbolic description and the mechanisms for manipulating them, existed hidden in huge assembly language programs [Horn 1971].

Many of the ideas regarding the use of early symbolic descriptions have come from a better understanding of human vision. Conversely, computer implementations provide an outstanding way of testing emerging theories about visual perception [Marr 1978]. Without such checks speculation runs rampant.

Understanding Image Formation

It is not uncharacteristic of computer science to tackle a new domain with total disdain for the details of the mechanisms evident in that domain. Of more interest to the computer scientist are questions of computational structures and efficiency and whether a proposed algorithm will apply in the new domain. Machine vision is no exception in this regard. It seems that for a long time there was very little interest in the origins of the arrays of numbers given as input to a machine vision system. Recently it has been found that many constraints due to the physics of the real world situation can be successfully exploited, once understood [Waltz 1975, Horn 1975a]. This enables processing not otherwise possible.

Understanding the process of image formation is helpful in inverting the image formation. That is, it is useful to know how objects are imaged, if one wishes to build a symbolic description of what is being viewed from the image. This kind of consideration has focused attention on smooth variations in intensities in an image [Horn 1977, Horn 1978]. Previously, image intensities were processed only to extract regions of more or less uniform properties or to locate points of more or less rapid

intensity change. At that point the image intensities themselves were discarded. This is unfortunate since a great deal of information about the objects being imaged is available there. This is quite different from the situation which applies in the case of binary images, useful in character recognition and printed circuit inspection for example.

Basically, what one is after is information about the permanent properties of the objects, such as reflectance, color and shape. This information is present in the raw image, but only in a coded fashion [Barrow & Tenenbaum 1978]. One may, for example, have to also deal with illumination conditions and shadowing. It is possible to extract all of this information from the raw image intensities, once the basic laws of image formation are understood. It is time to break the code.

Conclusion

Progress has been made -- at least we now know more about what we are up against. Much remains to be done. There is no shortage of good ideas right now, so we can discard some that no longer serve us well.

References

Agin, G. J. & Binford, T. O. (1973) "Computer Description of Curved Objects," Proc. 3rd Int. Joint Conf. on Art. Intell., Stanford University, Stanford.

Andrews, H. C. (1970) Computer Techniques in Image Processing, Academic Press, N.Y.

Arnold, R. D. (1978) "Local Context in Matching Edges for Stereo Vision," Proc. DARPA Image Understanding Workshop, May 1978, Baumann, L. (Ed), Science Applications, Inc.

Bajcsy, R. (1973) "Computer Identification of Visual Surfaces," Computer Graphics and Image Processing, Vol 2, No. 2, pp 118-130.

Barrow, H. G. & Tenenbaum, J. M. (1978) "Recovering Intrinsic Scene Characteristics from Images," in Hanson, A. & Riseman, E. Computer Vision Systems.

Biberman, L. M. (Ed) (1973) Perception of Displayed Information, Plenum Press, N.Y.

Binford, T. O. (1971a) "Visual Perception by Computer," Proc. IFIP Conf., Dubrovnik, Yugoslavia.

Binford, T. O. (1971b) "Visual Perception by Computer," IEEE Conf. Systems and Control, Miami, Dec 1971.

Brice, C. & Fennema, C. (1970) "Scene analysis using regions," Artificial Intelligence, Vol 1, No. 3, pp 205-226.

Clowes, M. B. (1971) "On seeing things," Artificial Intelligence, Vol 2, No. 1, pp 79-112.

Cheng, G. C. et al (Eds) (1968) Pictorial Pattern Recognition, Thompson Book Co., Washington D. C.

Duda, R. O. & Hart, P. E. (1973) Pattern Classification and Scene Analysis, John Wiley & Sons, N.Y.

Falk, G. (1972) "Interpretation of imperfect line data as a three-dimensional scene," Artificial Intelligence, Vol 4, No. 2, pp 101-144

- Fu, K. S. (1974) Syntactic Methods in Pattern Recognition, Academic Press, N.Y.
- Fu, K. S. (Ed) (1976) Digital Pattern Recognition, Springer, N.Y.
- Gennery, D. B. (1977) "A Stereo Vision System for an Autonomous Vehicle," Proc. 5th Int. Joint Conf. Art. Intell., Cambridge, Mass.
- Gonzalez, R. C. & Wintz, P. (1977) Digital Image Processing, Addison-Wesley, Reading, Mass.
- Grape, G. R. (1973) "Model Based (Intermediate-Level) Computer Vision," Stanford Univ. A.I. Memo 201.
- Grasselli, A. (Ed) (1969) Automatic Interpretation and Classification of Images, Academic Press, N.Y.
- Griffith, A. K. (1970) Computer Recognition of Prismatic Solids, M.I.T. Project Mac, TR-73, M.I.T., Cambridge, Mass.
- Gupta, J. N., Kettig, R. L., Landgrebe, D. A., & Wintz, P. A. (1973) "Machine boundary finding and sample classification of remotely sensed agricultural data," Machine Processing of Remotely Sensed Data, Purdue University, 4B25-4B35.
- Hanson, A. & Riseman, E. (Eds) (1978) Computer Vision Systems, Academic Press, N.Y.
- Hollerbach, J. (1976) "Hierarchical Shape Description of Objects by Selection and Modification Prototypes," M.I.T. A.I. T.R. 346.
- Horn, B. K. P. (1971) "The Binford-Horn Line-Finder," M.I.T. A.I. Memo 285.
- Horn, B. K. P. (1974) "Determining Lightness from an Image," Computer Graphics and Image Processing, Vol. 3, No. 1, pp 277-299.
- Horn, B. K. P. (1975a) "Determining Shape from Shading," Chapter 8, in Winston, P. H. (Ed) The Psychology of Computer Vision
- Horn, B. K. P. (1975b) "A Problem in Computer Vision: Orienting Silicon Integrated

Circuit Chips for Lead Bonding," Computer Graphics and Image Processing, Vol 4, No. 3, pp 294-303.

Horn, B. K. P. (1977) "Understanding Image Intensities," Artificial Intelligence, Vol. 21, No. 11, pp 201-231.

Horn, B. K. P. & Bachman, B. L. (1978) "Using Synthetic Images to Register Real Images with Surface Models," C.A.C.M. (in the press)

Hueckel, M. H. (1971) "An operator which locates edges in digital pictures," J. Assoc. Comp. Mach., Vol 18, pp 113-125.

Hueckel, M. H. (1973) "A local visual operator which recognizes edges and lines," J. Assoc. Comp. Mach., Vol 20, pp 634-647.

Huffman, D. A. (1971) "Impossible Objects as nonsense sentences," in Machine Intelligence, Vol 6, Metizer B. & Michie, D. (Eds), Edinburgh University Press, Edinburgh, pp 295-323.

Huang, T. S. (Ed) (1975) Picture Processing and Digital Filtering, Springer, N.Y.

Landgrebe, D. A. (1973) "Machine Processing for Remotely Acquired Data," LARS Information Note 031573. Laboratory for Applications of Remote Sensing, Purdue Univ., Lafayette, Ind.

Landgrebe, D. A. (1975) "NASA Contract NAS9-1416 Final Report," Laboratory for Applications for Remote Sensing, Purdue Univ., Lafayette, Ind.

Lipkin, B. S. & Rosenfeld, A. (Eds) (1970) Picture Processing and Psychopictorics, Academic Press, N.Y.

Mackworth, A. K. (1973) "Interpreting pictures of polyhedral scenes," Artificial Intelligence, Vol 4, pp 121-138.

Marr, D. (1974), "A note on the computation of binocular disparity in a symbolic, low-level visual processor," M.I.T. A.I. Memo 327.

Marr, D. (1976), "Early processing of visual information," Phil. Trans. Roy. Soc. B

275, 483-524.

Marr, D. & Poggio, T. (1976), "Cooperative computation of stereo disparity," Science 194, 283-287.

Marr, D. & Poggio, T. (1977), "A Theory of Human Stereo Vision," M.I.T. A.I.Memo 451.

Marr, D. & Nishihara, H. K. (1977), "Representation and recognition of the spatial organization of three-dimensional shapes," Proc. Roy. Soc. B (in the press).

Marr, D. (1978), "Representing visual information," in Computer Vision Systems, Hanson, A. & Riseman, E. (Eds)

Nevatia, R. (1974) "Structured Descriptions of Complex Curved Surfaces and Visual Memory," Stanford University, A.I. Memo 250.

Nevatia, R. (1976) "Depth Measurement by Motion Stereo," Computer Graphics and Image Processing, Vol 5, No 2., pp 203-214.

Nevatia, R. & Binford, T. O. (1977) "Description and Recognition of Curved Objects," Artificial Intelligence, Vol 8, No. 1, pp 77-98.

Nitzan, D. & Brain, A. E. & Duda, R. O. (1977) "The measurement and use of registered reflectance and range data in scene analysis," Proc. IEEE, Vol 65, No. 2, pp 206-220.

Nudd, G. R., Nygaard, P. A. & Thurmond, G. D. (1978) "Charge-Coupled Device Technology for Smart Sensors," Proc. DARPA Image Understanding Workshop, May 1978, Baumann, L. (Ed), Science Applications, Inc.

Ohlander, R. B. (1975) "Analysis of Natural Scenes," Carnegie-Mellon University, Department of Computer Science, Ph.D. Thesis.

Quam, L. H. (1971) "Computer Comparison of Pictures," Stanford University A.I. Memo 144.

Reddy, D. R., Erman, L. D. & Neely, R. B. (1973) "A model and a system for machine perception of speech," IEEE Trans. Audio and Electronics, AU-21, Vol 3., pp 229-238.

Roberts, L. G. (1965) "Machine Perception of Three-Dimensional Solids," in Optical and Electro-optical Information Processing, Tippet, J. T. et al (Eds), M.I.T. Press, Cambridge, Mass.

Rosenfeld, A. (1969a) "Picture Processing by Computer," Computing Surveys, Vol 1, pp 147-176.

Rosenfeld, A. (1969b) Picture Processing by Computer, Academic Press, N.Y.

Rosenfeld, A. (1972) "Picture processing: 1972," Computer Graphics and Image Processing, Vol 1, pp 394-416.

Rosenfeld, A. (1973) "Progress in picture processing: 1969-71," Computing Surveys, Vol 5.

Rosenfeld, A. (1975) "Picture Processing: 1974," Computer Graphics and Image Processing, Vol 4, No. 2, pp 133-155.

Rosenfeld, A., Hummel, R. A., & Zucker, S. W. (1976) "Scene Labelling by Relaxation Operations," IEEE Transactions on Systems, Man and Cybernetics, SMC-6, pp 420-433.

Rosenfeld, A. & Kak, A. C. (1976) Digital Picture Processing, Academic Press, N.Y.

Rosenfeld, A. (1976a) "Picture Processing: 1975," Computer Graphics and Image Processing, Vol 5, No. 2, pp 215-237.

Rosenfeld, A. (Ed) (1976b) Digital Picture Analysis, Springer, N.Y.

Rosenfeld, A. (1977a) "Picture Processing: 1976," Computer Graphics and Image Processing, Vol 6, No. 2, pp 157-183.

Rosenfeld, A. (1977b) "Iterative Methods in Image Analysis," Proc. IEEE Conf. on Pattern Recognition and Image Processing, June 1977, pp 14-18.

Rosenfeld, A. (1978) "Some Recent Results Using Relaxation-Like Processes," Proc. DARPA Image Understanding Workshop, May 1978, Baumann, L. (Ed), Science Applications, Inc.

Shirai, Y. (1975) "Analyzing Intensity Arrays using Knowledge about Scenes," Chapter 2,

in Winston, P. H. (Ed) The Psychology of Computer Vision.

Starr, D. W. & Mackworth, A. K. (1978) "Exploiting Spectral, Spatial and Semantic Constraints in the Segmentation of LANDSAT images," University of British Columbia, Computer Science T.R. 78-1.

Swain, P. H. (1973) "Pattern Recognition: A Basis of Remote Sensing Data Acquisition," LARS Information Note III572. The Laboratory for Applications of Remote Sensing, Purdue Univ., Lafayette, Ind.

Tenenbaum, J. M. & Barrow, H. G. (1976) "IGS: a paradigm for integrating image segmentation and interpretation," in Pattern Recognition and Artificial Intelligence, Academic Press, N.Y.

Tenenbaum J. M. & Barrow, H. G. (1977) "Experiments in interpretation-guided segmentation," Artificial Intelligence, Vol 8, No 3. pp 241-274.

Tou, J. T. & Gonzalez, R. C. (1974) Pattern Recognition Principles, Addison-Wesley, Reading, Mass.

Waltz, D. L. (1975) "Understanding Line Drawings of Scenes with Shadows," Chapter 2, in Winston, P. H. (Ed) The Psychology of Computer Vision.

Watanabe, S. (1969) Methodologies of Pattern Recognition, Academic Press, N. Y.

Winston, P. H. (1972) "The M.I.T. Robot," Machine Intelligence 7, Edinburgh University Press, Edinburgh.

Winston, P. H. (Ed) (1975) The Psychology of Computer Vision, McGraw-Hill, N.Y.

Winston, P. H. (1977) Artificial Intelligence, Chapter 8, Addison-Wesley, Reading, Mass.

Woodham, R. J. (1977) "A cooperative algorithm for determining surface orientation from a single view," Proc. 5th Int. Joint Conf. on Art. Intell., M.I.T., Cambridge, Mass.