

AD-A145 908

AUTOMATIC RECOGNITION AND TURBULING OBJECTS(U) TEXAS
UNIV AT AUSTIN COLL OF ENGINEERING J K AGGARWAL
31 DEC 83 AFOSR-TR-84-0771 F49620-83-K-0013

1/1

UNCLASSIFIED

F/G 20/6

NL

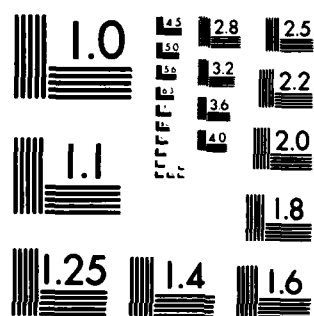
END

DATE

FORWED

10 84

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSS-TR- 61-0781	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Automatic Recognition and Turbuling Objects		5. TYPE OF REPORT & PERIOD COVERED ANNUAL
7. AUTHOR(s) Professor J.K. Aggarwal		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS The University of Texas at Austin College of Engineering Austin, Texas 78712		8. CONTRACT OR GRANT NUMBER(s) F49620-83-K-0013
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bolling AFB, DC 20332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2305 B3
12. REPORT DATE 1 Dec 82 31 Dec 83		13. NUMBER OF PAGES pp 1-15.
14. SECURITY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
16. STATEMENT (of this Report) Approved for public release; distribution unlimited.		15a. DECLASSIFICATION DOWNGRADING SCHEDULE
17. STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. KEY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) During the period December 1, 1982 through December 31, 1983, the research supported by AFOSR under Contract F49620-83-K-0013 resulted in the 11 presentations at national technical meetings, 10 published papers in referred journals, 7 published papers in conference proceedings, and 3 technical reports of the Laboratory for Image and Signal Analysis. The report describes briefly the following three projects: (1) A Normalized Quadtree Representation and A Volume Surface (Conte...		

AD-A145 908

DTIC
ELECTE
SEP 21 1984
S D E

DTIC FILE COPY

Representation;

4(2) Analysis of a Model for Parallel Image Processing; and

3) Determining Motion Parameters Using Intensity and Range Information.

A List of publications and presentations is provided.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

11/10

Automatic Recognition and Tracking of Moving Objects

Scientific Report for the Period

1 December 1982 - 31 December 1983

AFOSR Contract

No. F49620-83-K-0013

The University of Texas at Austin

Department of Electrical Engineering

Austin, Texas 78712

Principal Investigator: Professor J. K. Aggarwal

Approved
Gifting

84 09 17 066

Scientific Report

During the period December 1, 1982 through December 31, 1983, our group made 11 presentations and published 10 papers in refereed journals and 7 papers in conference proceedings. In addition, 3 technical reports were prepared. The group also prepared 7 papers which have been accepted for journal publication. A complete listing of these activities is provided at the end of this report. The group devoted its efforts to three areas of research briefly described in the next section of this report.

1. A Normalized Quadtree Representation and A Volume/Surface Octree Representation

Quadtrees are hierarchical data structures used for compact representations of two dimensional images. A quadtree is generated by dividing an image into quadrants and repeatedly subdividing the quadrants into sub-quadrants until each quadrant has uniform color (e.g., '1' or '0' in a binary image). The root of a quadtree corresponds to the image it represents. A node in a quadtree either is a leaf (terminal node) or has four son-nodes (non-terminal) node). Each son-node is associated with a quadrant of the block corresponding to its father-node.

Several excellent reviews of the subject are available in the literature. The advantage of the quadtree representation for images is that simple and well-developed tree traversal algorithms allow fast execution of certain operations such as superposition of two images, area and perimeter calculations,

moments computation, etc. Other researchers have shown that the quadtree representation of images yields substantial data compression over a variety of source images. In their experiments, image compression ratios ranging between three to one and thirty-three to one were found, with five or six to one being the general compression factor.

However, the quadtree representation has certain disadvantages. The quadtree representation of an object in an image is heavily affected by its location, orientation and relative size. A small change in these parameters will generate different quadtrees. One may eliminate the effect due to the translation of objects in an image by defining a normal form for quadtrees. Assuming that the size of an image lies between 2^{N-1} and 2^N , the image is moved around a region of size 2^{N+1} to find a minimal cost quadtree in terms of the number of nodes. This quadtree representation is unique for any image over the class of translations. However, the problem arising from rotations and size change still remains.

In this investigation, we propose a representation scheme, the normalized quadtree representation, which is invariant to object translation, rotation and size change. Instead of generating a quadtree for the entire image, a normalized quadtree is generated for each object in the image. The object is normalized to an object-centered coordinate system, with its centroid as the origin and principal axes as coordinate axes, and then scaled to a standard size (a $2^N \times 2^N$ image). In this way, the normalized quadtree of an object is dependent only on the shape of the object, but not affected by its location, orientation or relative size. In other words, the normalized quadtree representation can be utilized as a shape descriptor. In addition, information

related to the size, the position and the angle of the major principal axis of the object in the image may be retained enabling reconstruction of the object as it appeared in the image. According to Pavlidis' classification, the normalized quadtree representation is an information preserving shape descriptor.

A report has been prepared which briefly reviews the notions of moments and principal axes. The algorithm for the generation of quadtrees is then described followed by the description of the normalized quadtree representation and some examples. Finally, we describe the application of the normalized quadtree representation to two problems: shape matching and computation of principal moments. The report considers several examples of the generation and application of the normalized quadtree representations of a set of eight airplanes. The normalized octree representation and possible applications of the normalized quadtree representation for the modelling of three dimensional objects are also discussed.

The octree structure for the representation of 3-D objects is an extension of the quadtree representation of 2-D images. In general, it is generated from the 3-D binary array of the object it represents. However, the acquisition of 3-D array itself is a non-trivial problem. In the continuation of the above study, we generate the octree of an object from its three orthogonal views exploiting the volume intersection technique. To incorporate the surface information into the octree representation, which is basically a volume description, we propose a multi-level boundary search algorithm to find all the interfaces between black and white blocks. The information of each such interface is then stored in one of the two corresponding nodes. This makes the octree representation compact, informative, and especially useful for continuous

displays and object recognition tasks. All the algorithms developed in this study are essentially tree traversal procedures and therefore are suitable for implementation on parallel processors. In addition to developing the octree representation, we are investigating its applicability to recognition tasks.

2. Analysis of a Model for Parallel Image Processing

In the recent past there has been a great deal of attention focused on the design and implementation of parallel architectures to meet the computation demands of image processing and system design has followed two distinct paths. Low level operations which typically form the bulk of the computation and possess comparatively, a few well characterized operations, have been investigated for dedicated implementation. Based on descriptions of the algorithm, advanced VLSI design automation techniques are in the process of being developed to realize optimal dedicated low level architectures. On the other hand, high level processing characterized by widely varying data structures and computations, requires more flexible architectures for efficient operation. Multiple processors connected by dynamically reconfigurable networks have been investigated for the implementation of several algorithms established the utility of parallel image processing and provided some insight into the general problems encountered in constructing parallel tasks.

However, at this point it is still not evident what architectures are "best" suited for image processing in general, or for a given application in particular. It is equally uncertain how one may go about determining those which are. Though dedicated implementation provides high speed performance, their range of application is extremely narrow. On the other hand, the flexibility required of high level processes is achieved with some sacrifice in processing speed and increased system complexity. For general purpose image processing, what is required is the capability for uniform treatment of applications by way of a methodology, so that alternate choices can be compared

and contrasted in determining the best solutions. Techniques should take into account the underlying parallel tasks. Therefore, the overall problem of determining suitable parallel architectures for general purpose image processing encompasses the following:

- 1) being able to determine the requirements of parallel image processing tasks,
- 2) characterization of the capabilities of an architecture that is to host the computation and
- 3) techniques for "matching" the results of 1) and 2).

The present investigation addresses the problem 1) above and discusses how it affects choices in 3). Some solutions to problem 2) above are discussed elsewhere. The basic assumptions made about the architecture that will perform the processing are embodied in a parallel architecture model that is considered here. With this specification of the target machine, a synchronous parallel processing model is proposed for image analysis. Given a description of the algorithm and data structures to be operated upon, techniques are presented for constructing parallel tasks and specifying the communication requirements between them. Such a construction would produce an instance of the proposed processing model. An analysis of the parallel tasks constructed in this manner to illustrate how the model may be useful in determining a high level specification of the "best" architecture for a given application is considered. In addition, given pre-defined performance levels such as real time processing, it is shown how one may arrive at initial estimates of the capabilities of the components of the architecture, that are required to achieve these performance

goals, e.g., number and speed of the processing elements, relative speed of operation of communication and processing, etc. These estimates would provide a starting point for a designer prior to the refinement down to a complete detailed design. Examples are considered to illustrate the utility of the model and the techniques used to analyze an instance of the design. A report has been prepared and been submitted for publication.

3. Determining Motion Parameters Using Intensity and Range Information

The use of image processing techniques to derive motion parameters has been the subject of several recent research efforts. The use of intensity domain information has typically encountered difficulty imposed by the fact that from a single two-dimensional projection of three-dimensional scene, one may not determine the 3-D coordinates of any point in a scene. This arises since the projected point may result from any point lying on a line connecting it and the lens center. To determine the position of a point from intensity, at least two views are necessary. For motion parameter calculation correspondence among feature points in various views must be established and a complex systems of equations (which are frequently nonlinear) must be solved. We briefly review several of these methods to illustrate the complexity in the multi-view intensity domain.

Ullman developed trigonometric equations that derive object structure from two views of four points if it is known a priori that the transformation is a translation and a rotation about a single axis. Our early work has developed a method by which motion parameters may be derived from two views of five points but we are forced to solve a system of nonlinear equations. Nagel notes similar results.

For purposes of robot guidance and obstacle avoidance, Moravec presents an integrated system of vision hardware and software. His mobile vehicle has a camera mounted on a horizontal track so that when the robot is stationary it can slide the camera to any one of nine positions along the track. By determining

correspondence in the separately derived images and knowing the relative positioning of each of the camera locations, his system is able to determine the distance of objects while the robot is stopped. By examining several views (as seen by the robot at different positions) it is able to deduce the amount of motion the robot has undergone.

All of the multiple view intensity methods use the advantages afforded by rapid acquisition of data but must deal with the considerable difficulties of projection models. That is to say, intensity images can be gathered quickly at rates on the order of 30 frames per second or more. However, information contained in such projections must be supplemented by the establishing of correspondence of points in multiple views before three-dimensional structure or motion may be computed. On the other hand, several studies have made use of range sensing devices that can determine three-dimensional coordinates of a single point. The devices that perform such sensing are significantly slower in the data gathering process than those used for intensity imaging. The acquisition of a typical range image requires from several minutes to several hours.

In this investigation, we adopt the philosophy that the two domains should be used in a complementary fashion. Since intensity images can be obtained quickly but lack directly obtainable 3-D information, they should be used to guide sensing of a limited number of range image points, thus minimizing the data acquisition time in the range domain. As it will be seen the approach taken also avoids many of the problems associated with solving large systems of nonlinear equations and gives a good approximation to motion parameters.

The underlying philosophy of motion parameter extraction is that information from the intensity image of each view of an object should be used to guide range image sensing for the purpose of directly distilling 3-D coordinates that can be used to derive motion transformations. To this end, a set of feature points is identified in the intensity image. The nature of the feature points is not particularly important, but their extractability should be relatively insensitive to object orientation, assuming that they are not occluded. A second consideration is that it is desirable to link at least some of the feature points into a graph structure that partially represents the object that is being observed. Although this graph structuring is in some cases not absolutely necessary, it can greatly constrain the search when comparing the feature points of an object with those of various models for recognition and correspondence purposes. Such graph structures will generally represent geometric or spatial relationships between feature points.

After determining the basic (partial) graph structure for an object's feature points based solely on intensity information, the range image is sensed to determine the range values at these points. The 3-D coordinate corresponding to each feature point is associated to each node in the graph. A comparison of the information in the object's graph is made with all model graphs and the object is identified. The comparison itself is based upon arc lengths (distances between connected feature points) and angular differences between the arcs. Checking for consistency among the connections between nodes in the model and the candidate node matches in the object further filters out impossible labelings.

After an object is identified then the transformation required to move its model to the current object orientation is computed. This transformation is a 4×4 matrix that is the product of rotation and translation matrices. It is derived by treating the feature points for which good graph matching results were obtained as correspondence points and applying a least squares curve fit to compute the elements of the transformation matrix. The transformation matrix itself may be readily decomposed into a three-dimensional translation vector and a single rotation angle if it can be assumed that the moving object has traveled in a plane orthogonal to one of the principal axes. The imposition of such a constraint would be applicable in the case of a vehicle moving on a flat road which was perpendicular to the Z-axis.

If there are several views of the same object, the above process may be applied to each view and the relative (object centered) motion parameters extracted. This is achieved by deriving a sequence of model-to-object transformation matrices as previously described and relating adjacent elements in the sequence to derive object-to-object transformations.

The above methodology for the determination of motion parameters is being implemented in conjunction with the (Technical Arts) laser ranging device recently acquired by our laboratory. A report outlining preliminary results is under preparation.

PRESENTATIONS, PROCEEDINGS AND PUBLICATIONS**A. Presentations**

1. J. K. Aggarwal, "Dynamic Scene Analysis," at Carnegie-Mellon University, January 1983.
2. J. K. Aggarwal, "3-D Computer Vision - An Introduction," at the ACM Siggraph/Sigart Workshop on Motion - Representation and Perception, Toronto, Canada, April 1983.
3. W.N. Martin, B. Gil and J. K. Aggarwal, "Volumetric Representation for Object Model Acquisition," at the NASA Symposium on Computer Aided Geometry Modeling, Hampton, VA, April 1983.
4. J. Courtney and J. K. Aggarwal, "Robot Guidance Using Computer Vision," at Trends and Application 1983 Conference, Gaithersburg, MD, May 1983.
5. S. Yalamanchili and J. K. Aggarwal, "A Model for Parallel Image Processing," at the International Society for Optical Engineering Annual Symposium, San Diego, CA, August 1983.
6. M. Magee and J. K. Aggarwal, "Intensity Guided Range Sensing Recognition of Three-Dimensional Objects," at the 3-D Workshop of the American Association for Artificial Intelligence, Washington, DC, August 1983.
7. Y.C. Kim and J. K. Aggarwal, "Rectangular Coding of Binary Images," at IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983.
8. M. Magee and J. K. Aggarwal, "Intensity Guided Range Sensing Recognition of Three-Dimensional Objects," at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983.
9. C.H. Chien and J. K. Aggarwal, "A Normalized Quadtree Representation," at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983.
10. J. K. Aggarwal, "Pattern Recognition Vs. Artificial Intelligence," at the 12th Applied Imagery Pattern Recognition Workshop at University of Maryland, College Park, MD, September 27, 1983.
11. M. Magee and J. K. Aggarwal, "Robot Vision for Location Determination and Obstacle Avoidance" at COMPCON Fall '83, Arlington, Virginia, September 25-29, 1983.

B. Papers

1. W. N. Martin and J. K. Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 2, pp. 150-158, March 1983.
2. Amar Mitiche and J. K. Aggarwal, "Detection of Edges Using Range Information," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 2, pp. 174-178, March 1983.
3. J. A. Webb and J. K. Aggarwal, "Shape and Correspondence," Computer Vision, Graphics and Image Processing, Vol. 21, No. 1, pp. 145-160, January 1983.
4. B. Gil, A. Mitiche and J. K. Aggarwal, "Experiments in Combining Intensity and Range Edge Maps," Computer Vision, Graphics and Image Processing, Vol. 21, No. 3, pp. 395-411, March 1983.
5. W. N. Martin and J. K. Aggarwal, "Volumetric Descriptions from Dynamic Scenes," Pattern Recognition Letters, Vol. 1, pp. 107-113, December 1982.
6. A. Mitiche, B. Gil and J. K. Aggarwal, "On Combining Range and Intensity Data," Pattern Recognition Letters, Vol. 1, pp. 87-92, December 1982.
7. J. K. Aggarwal, Guest editor for Volume 21, Nos. 1 and 2, January and February 1983 issues of Computer Vision, Graphics and Image Processing, on Motion and Time-Varying Imagery.
8. W. N. Martin and J. K. Aggarwal, "Dynamic Scene Analysis," in the book Image Sequence Processing and Dynamic Scene Analysis edited by T. S. Huang, Springer-Verlag, 1983, pp. 40-73.
9. A. Mitiche and J. K. Aggarwal, "Contour Registration by Shape-Specific Points for Shape Matching," Computer Vision, Graphics and Image Processing, Vol. 22, pp. 396-408, December 1983.
10. J. K. Aggarwal, "Motion and Time-Varying Imagery," Computer Graphics, Vol. 18, No. 1, January 1984, pp. 20-21.

C. Papers Prepared to Appear

1. S. Yalamanchili, K. V. Palem, L. S. Davis, A. J. Welch and J. K. Aggarwal, "Image Processing Architectures: A Taxonomy and Survey," to appear in Progress in Pattern Recognition, Vol. 2, Edited by L. N. Kanal and A. Rosenfeld.
2. C. H. Chien and J. K. Aggarwal, "A Normalized Quadtree Representation," to appear in Computer Vision, Graphics and Image Processing.

3. M. Magee, J. W. Courtney and J. K. Aggarwal, "Robot Guidance Using Computer Vision," to appear in Pattern Recognition.
4. Y. F. Wang, M. Magee and J. K. Aggarwal, "Matching Three-Dimensional Objects Using Silhouettes," to appear in IEEE Transactions on Pattern Analysis and Machine Intelligence.
5. M. Magee and J. K. Aggarwal, "Determining Vanishing Points from Perspective Images," to appear in Computer Vision, Graphics and Image Processing.
6. S. Yalamanchili and J. K. Aggarwal, "Analysis of a Model for Parallel Image Processing," to appear in Pattern Recognition.
7. S. Yalamanchili and J. K. Aggarwal, "Formulation of Parallel Image Processing Tasks," to appear in Pattern Recognition Letters.

D. Conference Proceedings

1. Y.C. Kim and J. K. Aggarwal, "Rectangular Coding of Binary Images," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983, pp. 108-113.
2. C.H. Chien and J. K. Aggarwal, "A Normalized Quadtree Representation," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983, pp. 121-126.
3. M. Magee and J. K. Aggarwal, "Intensity Guided Range Sensing Recognition of Three-Dimensional Objects," at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 1983, pp. 550-552.
4. J. Courtney and J. K. Aggarwal, "Robot Guidance Using Computer Vision," at Proceedings of IEEE Computer Society Conference on Trends and Applications 1983, Gaithersburg, MD, May 1983, pp. 57-62.
5. S. Yalamanchili and J. K. Aggarwal, "A Model for Parallel Image Processing," at the International Society for Optical Engineering Annual Symposium, San Diego, CA, August 1983.
6. W.N. Martin and B. Gil and J. K. Aggarwal, "Volumetric Representation for Object Model Acquisition," Proceedings NASA Symposium on Computer Aided Geometry Modeling, NASA Conference Publication No. 2272, pp. 87-94, April 1983.
7. M. Magee and J. K. Aggarwal, "Robot Vision for Location Determination and Obstacle Avoidance," proceedings of COMPCON, Fall '83, Arlington, Virginia, September 25-29, 1983, pp. 201-210.

E. Technical Reports

1. J. K. Aggarwal, "Three-Dimensional Description of Objects and Dynamic Scene Analysis," Report No. TR-83-1-20, Laboratory for Image and Signal Analysis, The University of Texas, Austin, TX, January 1983.
2. M. Magee and J. K. Aggarwal, "Intensity Guided Range Sensing Recognition of Three-Dimensional Objects," Report No. TR-83-3-22, Laboratory for Image and Signal Analysis, The University of Texas, Austin, TX, February 1983.
3. M. Magee and J. K. Aggarwal, "Determining Vanishing Points from Perspective Images," Report No. TR-83-2-21, Laboratory for Image and Signal Analysis, The University of Texas, Austin, TX, February 1983.