AD-A264 335

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED FINAL/01 JUN 91 TO 30 SEP 92

4. TITLE AND SUBTITLE
QUALITATIVE METHODS IN COMPUTER VISION (U)

DTIC ELECTE MAY 14 1993 S C D

5. FUNDING NUMBERS
2304/A7
AFOSR-91-0239

6. AUTHOR(S)
Professor Azriel Rosenfeld

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Maryland
Dept of Mathematical Sciences
College Park MD 20742

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 93 0214

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 Duncan Ave Suite 100
Bolling AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-91-0239

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release;
Distribution unlimited

12b. DISTRIBUTION CODE
UL

13. ABSTRACT *(Maximum 200 words)*

Current object recognition systems can only "recognize" a limited class of objects. Objects having variable numbers of parts and only loosely constrained shapes cannot be modeled and recognized by these systems. The PI proposed the use of a data structure called the VAPOR (Variable APpearance Object Representation) model to represent objects with these kinds of variable appearances and develop a search procedure called MOSS (MOdel Space Search) to find instances of these models in two-dimensional image data. The VAPOR model is an idealization of the object; all instances of the model in the image are variations from ideal appearance. The variations are evaluated by the description length of the model, measured in information-theoretic bits. MOSS selects the best model for the given image data by choosing the minimal length description. It was demonstrated how the system performs in a simple domain of circles and polygons and in the complex domain of finding cloverleaf intersections in aerial images of roads.

93 5 13 040

93-10743

413788

14. SUBJECT TERMS

15. NUMBER OF PAGES 8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

80014 5H p.73

Final Report on Grant AFOSR-91-0239

Prepared by: Azriel Rosenfeld
Principal Investigator
Center for Automation Research
University of Maryland
College Park, Maryland 20742-3275

January, 1993

During the period August 1991-October 1992, five Ph.D. dissertations were completed with the support of the Grant:

1) The Ph.D. dissertation of Mr. John Canning on recognition of two-dimensioned objects with variable appearances was issued as a technical report [1].
2) Portions of the Ph.D. dissertation of Mr. Saibal Banerjee on Maximum A Posteriori estimation in image and signal analysis were issued as technical reports [2,4]; earlier parts appeared as technical reports on a predecessor grant, AFOSR-86-0092.
3) The Ph.D. dissertation of Mr. Sven Dickinson on three-dimensional object recognition using volumetric primitives was issued as a technical report [3].
4) The Ph.D. dissertation of Mr. Doron Mintz on robust estimation techniques for computer vision was issued as a technical report [5].
5) Part of the Ph.D. dissertation of Mr. Noah Friedland on optimization-based dilneation of complex two-dimensional objects, with applications to vehicles in FLIR imagery, was issued as a technical report [6]; the entire dissertation is about to be issued as a technical report on the successor grant F49620-93-1-0039.
6) Two papers on polygonal ribbons in two and three dimensions were issued as technical reports [7,9].
7) A bibliography of nearly 1200 references on computer vision, image analysis, and related topics was issued as a technical report [8].
8) A paper on the use of evidence functions in pattern-matching relaxation, with applications to target detection in simulated SAR imagery, was issued as a technical report [10].
9) The Ph.D. dissertation of Mr. Nathan Netanyahu on computationally efficient algorithms for robust estimation was issued as a technical report [11]; a later report [14] includes additional results on the repeated median estimator.
10) An initial portion of the Ph.D. dissertation of Mr. Zoran Duric, on dynamics-based models for vehicle motion, was issued as a technical report [12].
11) A portion of the Ph.D. dissertation of Mr. Mohamed Abdel-Mottaleb, on the use of Bayesian techniques in motion estimation, was issed as a technical report [13].

Abstracts of the tecnnical reports issued on the Grant are appended to this report; the numbers in brackets refer to these abstracts.

1. John Canning, "Recognizing Objects with Variable APpearances—The VAPOR System." CAR-TR-563, CS-TR-2705, August 1991.

ABSTRACT: Current object recognition systems can only "recognize" a limited class of objects. Objects having variable numbers of parts and only loosely constrained shapes cannot be modeled and recognized by these systems. We propose the use of a data structure called the *VAPOR* (Variable APpearance Object Representation) model to represent objects with these kinds of variable appearances and develop a search procedure called *MOSS* (MOdel Space Search) to find instances of these models in two-dimensional image data. The VAPOR model is an idealization of the object; all instances of the model in the image are variations from the ideal appearance. The variations are evaluated by the ***description length*** of the model, measured in information-theoretic bits. MOSS selects the best model for the given image data by choosing the minimal length description. We demonstrate how the system performs in a simple domain of circles and polygons and in the complex domain of finding cloverleaf intersections in aerial images of roads.

2. Saibal Banerjee and Azriel Rosenfeld, "Model-Based Cluster Analysis." CAR-TR-570, CS-TR-2730, August 1991.

ABSTRACT: In this paper the problem of dot clustering is studied from a model-based viewpoint. A set of scatter processes (in brief: scatters) is chosen, each of which associates a probability with each location in a discrete space; in other words, a scatter is a pmf on the space. Some number of dots is then distributed in accordance with each scatter. A scatter and an associated numerosity define a subpopulation of dots. This model is extremely general; the scatter pmf's are arbitrary.

Given a set of dots generated by such a model, we apply Maximum A Posteriori (MAP) methods to recover the most likely set of scatters and numerosities that could have given rise to the dots. This identification problem is different from the partitioning problem, which asks for the most likely partition of the dot population into subpopulations. We show how and why MAP methods are useful in cluster analysis, especially when the scatters are non-Gaussian. We also show that although the general identification problem is intractable, there is a polynomial time solution if the number of clusters is bounded. We show that a similar result holds for the partitioning problem.

3. Sven Dickinson, "The Recovery and Recognition of Three-Dimensional Objects Using Part-Based Aspect Matching." CAR-TR-572, CS-TR-2732, August 1991.

ABSTRACT: We present an approach to the recovery and recognition of 3-D objects from a single 2-D image. The approach is motivated by the need for more powerful indexing primitives, and shifts the burden of recognition from the model-based verification of simple image features to the bottom-up recovery of complex volumetric primitives. Given a recognition domain consisting of a database of objects, we first select a set of object-centered 3-D volumetric modeling primitives that can be used to construct the objects. Next, using a CAD system, we generate the set of aspects of the primitives. Unlike typical aspect-based recognition systems that use aspects to model entire objects, we use aspects to model the finite set of *parts* from which the objects are constructed. Consequently, the number of aspects is fixed and *independent* of the size of the object database. To accommodate the matching of partial aspects due to primitive occlusion, we introduce a hierarchical aspect representation based on the projected surfaces of the primitives; a set of conditional probabilities captures the ambiguity of mappings between the levels of the hierarchy.

From a region segmentation of the input image, we present a novel formulation of the primitive recovery problem based on grouping the regions into aspects. No domain dependent heuristics are used; we exploit only the probabilities inherent in the aspect hierarchy. Once the aspects are recovered, we use the aspect hierarchy to infer a set of volumetric primitives and their connectivity relations. Subgraphs of the resulting graph, in which nodes represent 3-D primitives and arcs represent primitive connections, are used as indices to the object database. The verification of object hypotheses consists of a topological verification of the recovered graph, rather than a geometrical verification of image features. A system has been built to demonstrate the approach, and it has been successfully applied to both synthetic and real images.

4. Saibal Banerjee and Azriel Rosenfeld, "MAP Estimation of Context-Free Grammars." CAR-TR-575, CS-TR-2735, August 1991.

ABSTRACT: In this paper Bayesian inference is used to recover, from a finite set of candidate grammars. the most probable grammar (and derivation) that generated the non-noisy version of an observed noisy string. It is assumed that the noise process is iid and defined by an arbitrary stochastic matrix. It is shown that if the grammars are context-free or stochastic context-free, this problem is solvable in polynomial time. Generalizations are also discussed.

5. Doron Mintz, "Robustness By Consensus." CAR-TR-576, CS-TR-2745, August 1991.

ABSTRACT: This report presents a paradigm for performing robust model fitting. Model fitting is an essential part of many algorithms that try to interpret data. Such algorithms normally use least squares based estimation for fitting purposes. The least squares procedure is very sensitive to noise and variations in the data. Furthermore, it is frequently the case that a model discontinuity exists in the data and it cannot be reliably detected by conventional methods.

Recently, new robust estimation methods have been introduced that deal with several types of variations and model discontinuities. Some of these methods are analyzed and shown not to be effective in the presence of noise such as is usually found in image and signal data. The new paradigm integrates basic robust estimation methods, a decomposition of the problem and a relative majority analysis to perform robust estimation. The robust estimate is a *consensus* among independent, invariant estimates of sub-problems.

The consensus paradigm is shown to provide effective algorithms for signal and image estimation, edge detection, segmentation, and contour matching. The performance of these algorithms is analyzed and compared to existing methods. The paradigm can also be used in connection with other computer vision problems such as stereo, flow and shape from X, as well as for more general sensor fusion and hypotheses integration algorithms.

6. N. S. Friedland and Azriel Rosenfeld, "Lobed Object Delineation Using a Multipolar Representation." CAR-TR-590, CS-TR-2779, October 1991.

ABSTRACT: This paper describes an iterative method of delineating and representing lobed objects, i.e., objects containing multiple compact parts. The method makes use of a new multi-polar representation (MPR). The process begins by constructing a single-center polar representation somewhere inside the object. This establishes a 1-D, cyclic Markov Random Field (1DCMRF) which optimizes edge sharpness and contour smoothness. The 1DCMRF is segmented into sectors at significant minima of the radius; these sectors define lobes. Next, each lobe is assigned a candidate polar representation centered at the lobe's centroid. This new set of representations is then compared to the previous one using a "radius entropy test" (RET). This test selects the representation with the highest degree of roundness, i.e. the representation in which the radii are most uniform in their sizes. When a new representation supersedes an old one, each new polar center establishes a 1-D Markov Random Field and boundary conditions are determined between neighboring MRFs. This process continues recursively within each of these MRFs. To deal with deep lobes and concavities we define two special classes of MRFs. Tunnel MRFs are used to explore long narrow lobes; these MRFs extend the original lobe center by dividing the lobe's radii into fore and aft groups. External MRFs, whose centers lie outside of the object,

are used to delineate deep concavities. These are detected when significant gaps appear between neighboring MRFs. These MRFs must also meet the RET criterion in their creation. MPR generalizes a method developed by the authors to recognize compact objects. Its strength lies in its ability to operate on raw image data without preprocessing, its insensitivity to initialization, and the fact that the object representation is constructed simultaneously with the delineation process, an important advantage for future recognition applications. The optimization of the MRFs is done using simulated annealing.

7. Prabir Bhattacharya and Azriel Rosenfeld, "Triangulated Ribbons in Two and Three Dimensions." CAR-TR-594, CS-TR-2796, November 1991.

ABSTRACT: A triangulated ribbon is a finite sequence of triangles such that each pair of successive triangles intersect in a common side. The study of such ribbons can be regarded as a first step toward the study of triangulated surfaces. In this paper we investigate a number of geometrical and topological properties of ribbons in two and three dimensions. First, we develop a code to represent a ribbon using a sequence of vertices. We then describe how to derive a number of geometric properties of a ribbon, including nonselfintersection, twist and orientability (for a closed ribbon). For nonselfintersecting (NSI) ribbons, we prove that the ribbon and its complement are both connected and that an open ribbon is simply connected, and we characterize the border(s) of NSI ribbons. If a planar ribbon is regarded as a subset of the plane, its connectivity properties need a different treatment; for example, the complement of a (nondegenerate) NSI closed ribbon is not connected. We also give conditions under which given simple polygons can be the borders of NSI ribbons, and we discuss the reconstructibility of ribbons from projections.

8. Azriel Rosenfeld, "Image Analysis and Computer Vision: 1991." CAR-TR-601, CS-TR-2817, January 1992.

ABSTRACT: This paper presents a bibliography of nearly 1200 references related to computer vision and image analysis, arranged by subject matter. The topics covered include architectures; computational techniques; feature detection and segmentation; image analysis; two-dimensional shape; pattern; color and texture; matching and stereo; three-dimensional recovery and analysis; three-dimensional shape; and motion. A few references are also given on related topics, such as geometry, graphics, image input/output and coding, image processing, optical processing, visual perception, neural nets, pattern recognition, and artificial intelligence, as well as on applications.

9. Prabir Bhattacharya and Azriel Rosenfeld, "Rectangular Ribbons in Two and Three Dimensions." CAR-TR-613, CS-TR-2871, April 1992.

ABSTRACT: A rectangular ribbon is a finite sequence of rectangles such that each pair of successive rectangles intersect in a common side. Such ribbons are an interesting first step toward the study of polyhedral surfaces that have rectangular faces. In this paper we investigate a number of geometrical and topological properties of rectangular ribbons in two and three dimensions. First, we develop compact methods of representing a ribbon in terms of vertex coordinates and/or side lengths and dihedral angles. We then describe how to derive a number of geometric properties of a ribbon, including twist, orientability and nonselfintersection. For nonselfintersecting (NSI) ribbons, we prove that an open ribbon and a degenerate closed ribbon are simply connected, while a nondegenerate closed ribbon is not. We also show that the border of an NSI open ribbon is connected while that of an NSI closed ribbon is connected if and only if the ribbon is nonorientable. Finally, we briefly consider the cases of isothetic ribbons and of ribbons lying on a plane.

10. Peter Cucka and Azriel Rosenfeld, "Evidence-Based Pattern-Matching Relaxation." CAR-TR-623, CS-TR-2896, May 1992.

ABSTRACT: In its original form the point pattern-matching relaxation scheme of Ranade and Rosenfeld did not easily permit the representation of uncertainty, and it did not exhibit the desirable property that confidence in consistent pairings of features should increase from one iteration to the next. Because the process of pooling intrinsic support with contextual support is essentially a process of evidence combination, it was suggested by Faugeras that the evidence theory of Dempster and Shafer might be an appropriate framework for relaxation labeling. Here, we address some of the issues involved in the implementation of such an approach and present results from the domain of object recognition in SAR imagery.

11. Nathan S. Netanyahu, "Computationally Efficient Algorithms for Robust Estimation." UMIACS-TR-92-54, CS-TR-2898, May 1992.

ABSTRACT: Data processing for scientific and industrial tasks often involves accurate extraction of theoretical model parameters from empirical data, and requires automated estimation methods that are robust in the presence of "noisy" (i.e., contaminated) data. *Robust estimation* is thus an important statistical tool that is frequently applied in numerous fields of science and engineering (e.g., automated manufacturing, robotic navigation, image processing, and computer vision).

Since the computational complexity of a robust estimator is one of the most important measures of its practicality, searching for methods that reduce the time (and space) complexity of robust estimators is a desirable research goal. In this dissertation we present several *computationally efficient* algorithms for the *exact* computation of robust statistical estimators. In particular, we consider the design and analysis of such algorithms for various problem domains, including line, curve, and surface fitting.

We begin by providing a survey of current computational approaches and complexity results for various robust estimators. Next, a general underlying methodology is introduced for the efficient computation of the classes of estimators considered. Specifically, the application of *computational geometry* techniques in the derivation of robust estimation algorithms is emphasized. Furthermore, it is demonstrated that the derivation, in particular, of *randomized algorithms* for the above tasks results in algorithms that have the following properties: (1) The algorithms always terminate and return the *correct* computational results, (2) the improved (expected) running times occur with extremely high probability, (3) the algorithms are quite easy to implement; (4) the constants of proportionality (hidden by the asymptotic notation) are small (i.e., the algorithms are *practical*), and (5) the algorithms are space optimal (i.e., they require linear storage).

We discuss implementational issues in great detail and report our practical experience with the algorithms. Finally, various future research topics are proposed.

12. Zoran Durić, Azriel Rosenfeld and Larry S. Davis, "Egomotion Analysis Based on the Frenet-Serret Motion Model." CAR-TR-634, CS-TR-2944, August 1992.

ABSTRACT: In this paper we propose a new model, *Frenet-Serret* motion, for the motion of an observer in a stationary environment. This model relates the motion parameters of the observer to the curvature and torsion of the path along which the observer moves. We derive screw-motion equations for Frenet-Serret motion and use them for geometrical analysis of the motion as well as analysis of the resulting velocity patterns in 3-D and motion field patterns on the surface of the velocity egosphere [Albus, 1991]. We use normal flow to derive constraints on the rotational and translational velocity of the observer and compute egomotion by intersecting these constraints in the manner proposed in [Durić and Aloimonos, 1991]. We analyze the accuracy of egomotion estimation for different combinations of observer motion and feature distance. We propose that depth of field should be controlled in order to make the analysis of egomotion on the basis of normal flow possible, and we derive the constraints on depth which make either rotation or translation dominant. The results of experiments on real image sequences are presented.

13. Mohamed Abdel-Mottaleb, Rama Chellappa and Azriel Rosenfeld, "Passive Navigation using Bayesian Estimation." CAR-TR-636, CS-TR-2946, August 1992.

ABSTRACT: We present an algorithm for calculating the camera motion and the structure of a (rigid) scene for the case of a camera moving along a smooth trajectory. The sequence of images (frames) we use is dense so that the displacement from frame to frame is at most n pixels (where n is typically 2 or less) We calculate instantaneous estimates of the focus of expansion (FOE) and of the scene depth map, and keep updating these estimates through the sequence. Our algorithm calculates the maximum a posteriori (MAP) estimate of the subpixel displacement and a confidence measure in that estimate. Using points for which the confidence is high we calculate a MAP estimate for the FOE and the magnitudes of the displacements at these points, hence their relative depths. After determining the FOE we know the direction of displacement at every point in the image and we can again apply the MAP estimation method to get the displacement magnitude at each point and the associated confidence measure. This information is propagated over a long sequence of images by using the a posteriori distribution calculated from a set of images as a prior for the next set of images.

14. Peter J. Rousseeuw, Nathan S. Netanyahu and David M. Mount, "New Statistical and Computational Results on the Repeated Median Line." CAR-TR-643, CS-TR-2976, October 1992.

ABSTRACT: Existing algorithms for affine equivariant regression estimators with high breakdown point are computer intensive. Heuristically, this appears to be due to combinatorial and geometric reasons. Consequently, non-affine estimators may allow a faster computation. This paper describes an algorithm for the repeated median slope estimator which runs in $O(n \log^2 n)$ time, a substantial improvement over the naive $O(n^2)$ method. The new algorithm allows an empirical study of this estimator for n up to 40,000. It turns out that the finite-sample efficiencies converge extremely slowly although the estimator is asymptotically normal.