0

AD-A282 962

TEC-0047

US Army Corps
of Engineers
Topographic
Engineering Center

Vision-Based Navigation and Recognition

T

Azriel Rosenfeld

Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3275

E

July 1994

Approved for public release; distribution is unlimited.

Prepared for:
Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714

Monitored by:
U.S. Army Corps of Engineers
Topographic Engineering Center
7701 Telegraph Road
Alexandria, Virginia 22315-3864

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
July 1994

3. REPORT TYPE AND DATES COVERED
Technical Report April 1992 - March 1993

4. TITLE AND SUBTITLE
Vision-Based Navigation and Recognition

5. FUNDING NUMBERS
DACA76-92-C-0009

6. AUTHOR(S)
Azriel Rosenfeld

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Computer Vision Laboratory/Center for Automation Research
University of Maryland
College Park, MD 20742-3275

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Advanced Research Projects Agency
3701 North Fairfax Drive, Arlington, VA 22203-1714

U.S. Army Topographic Engineering Center
7701 Telegraph Road., Alexandria, VA 22315-3864

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
TEC-0047

11. SUPPLEMENTARY NOTES

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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The research summarized in this report deals with many aspects of computer vision as applied to both navigation and object recognition. In particular, this research has concentrated on eight areas: parallel algorithms for vision; diffusion processes and their roles in early vision; invariant properties and their roles in object recognition; recovery of three-dimensional scene properties from single images; recovery of observer motion and scene structure from image sequences; direct motion analysis; visual interception; and vision-based navigation. Specific topics include algorithms for image and graph computations, parallel search and stereo matching; the application of diffusion processes to image morphing and face recognition; projective, affine, and deformation invariants of images; reliability of geometric computations on images of three-dimensional scenes; properties of foliage regarded as a three-dimensional texture; monocular and binocular recovery of motion and structure from feature correspondents in an image sequence; a unified treatment of feature-based and flow-based motion estimation; motion analysis based on global properties of the flow field; vision-based target interception; visibility on terrain; and landmark-based localization.

14. SUBJECT TERMS Computer vision, parallel algorithms, SIMD machines, stereo matching, diffusion processes, image morphing, face recognition, invariants, geometric computations, 3D texture, motion and structure recovery

15. NUMBER OF PAGES
28

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	UNLIMITED

Contents

List of Figures

PREFACE

This research is sponsored by the Advanced Research Projects Agency (ARPA) and monitored by the U.S. Army Topographic Engineering Center (TEC) under Contract DACA76-92-C-0009, titled "Vision-Based Navigation and Recognition". The ARPA Program Manager is Mr. Oscar Firschein, and the TEC Contracting Officer's Representative is Ms. Lauretta Williams.

1. Introduction

Image understanding research at the Center for Automation Research of the University of Maryland at College Park deals with many aspects of both navigation and recognition. This report summarizes the research conducted under Contract DACA76-92-C-0009 (ARPA Order 8459) during the period April 1992 – March 1993.

The research conducted under the Contract has been concentrated in eight areas:

(a) Parallel algorithms for vision

(b) Diffusion processes and their roles in early vision

(c) Invariant properties and their roles in object recognition

(d) Recovery of three-dimensional scene properties from single images

(e) Recovery of observer motion and scene structure from image sequences

(f) Direct motion analysis

(g) Visual interception

(h) Vision-based navigation

The work done in these areas is summarized in Sections 2–9 of this report. Further details about this work can be found in 20 technical reports issued on the Contract during the period April 1992 – March 1993. A Bibliography of these reports is given in Section 10 of this report; the numbers in brackets in Sections 2–9 refer to this list.

2. Parallel Algorithms for Vision

2.1. SIMD Machines [12]

Single Instruction Stream, Multiple Data stream (SIMD) processor array machines are popular in practical parallel computing. Such machines differ from one another considerably in the level of autonomy provided to each processing element (PE) of the array. An understanding of the levels of autonomy provided by the architectures is important in the design of efficient algorithms for them. SIMD architectures are classified into six categories differing in key aspects such as the selection of the instructions to be executed, operands for the instructions, and the source/destination of communications.

The data parallel model of computation used in processor arrays exploits the parallelism in the data by processing multiple data elements (pixels, in image analysis) simultaneously by assigning one PE to each data element. This scheme does not make efficient use of the processor array when processing relatively small data structures. A technique of data replication was developed that combines operation parallelism with data parallelism, in order to process small data structures efficiently on large processor arrays. It decomposes the main operation into suboperations that are performed simultaneously on separate copies of the data structure. The autonomy of the individual PEs is critical to this decomposition. Replicated data algorithms were developed for several low level image operations such as histogramming, convolution, and rank order filtering. Additionally, a method was developed of constructing a replicated data algorithm for an operation automatically from an image algebra expression for it, thus demonstrating the generality of this approach. A replicated data algorithm to compute single source shortest paths on general graphs was also devised, thus demonstrating the applicability of the approach beyond image analysis. The speedup performance of the algorithms on various interconnection networks was analyzed in order to determine the conditions under which the technique results in a speedup. Implementations of the algorithms on a Connection Machine CM-2 and a MasPar MP-1 yielded impressive speedups.

A parallel search scheme for the model-based interpretation of aerial images under a focus-of-attention paradigm also developed and was implemented on a CM-2. Candidate objects are generated as connected combinations of the connected components of the image and are matched against the model by checking if the parameters computed from the region satisfy the model constraints. This process is posed as a search in the space of combinations of connected components with the finding of an (optimally) successful region as the goal. The implementation exploits parallelism at multiple levels by parallelizing control tasks such as the management of the open list. The level of processor autonomy and other details of the architecture play important roles in the search scheme.

2.2. An Application: Stereo Matching [4]

The use of dynamic programming for stereo matching has been studied extensively. It has been pointed out that this approach is suitable for parallel processing, but there have not as yet been any attempts to implement a dynamic programming stereo matching algorithm on a parallel machine. A massively parallel implementation of Baker's dynamic programming stereo algorithm was developed; the implementation uses many processors per scanline, compared to a naive approach of one processor per scanline. This is important because typical images contain 256 to 1024 scanlines, while massively parallel machines can have many more processors. A method of handling inter-scanline inconsistencies was introduced that is very well suited for parallel implementation. The method increases the amount of processing needed to solve the stereo matching problem by only a small fraction. On a 16K processor Connection Machine the entire algorithm requires as little as 1 second for simple 512×512 images.

3. Diffusion Processes and their Roles in Early Vision

3.1. Diffusion Processing of Range Data [11]

The use of a multi-stage physical diffusion process in early vision processing of range images was investigated. The input range data is interpreted as occupying a volume in 3-D space. Each diffusion stage simulates the process of diffusing the boundary of the volume into the volume. The results of the diffusion process appear to be useful for both discontinuity detection and segmentation into shape coherent regions. Diffusion processing of an image of a human face (the original and three diffusion stages) is illustrated in Figure 1.

3.2. An Application to Image Morphing [19]

Image interpolation and metamorphosis can be performed by using a scale space created by diffusing the difference function of the source and the goal images. This formulation makes it possible to minimize the need for human intervention in the selection of features in a process such as *image metamorphosis*. The smooth transitions are accompanied by a moderated blurring that is useful in displaying the metamorphosis process. The approach can also be applied to motion image sequences as a method of enhancing animation.

3.3. An Application to Face Recognition [15]

An approach to labeling the components of faces from range images was developed. The components of interest are those which humans usually find significant for recognition. To cope with the non-rigidity of faces, an entirely qualitative approach is used. A preprocessing stage employs a multi-stage diffusion process to identify convexity and concavity points. These points are grouped into components and qualitative reasoning about possible interpretations of the components is performed. Consistency of hypothesized interpretations is verified using context-based reasoning.

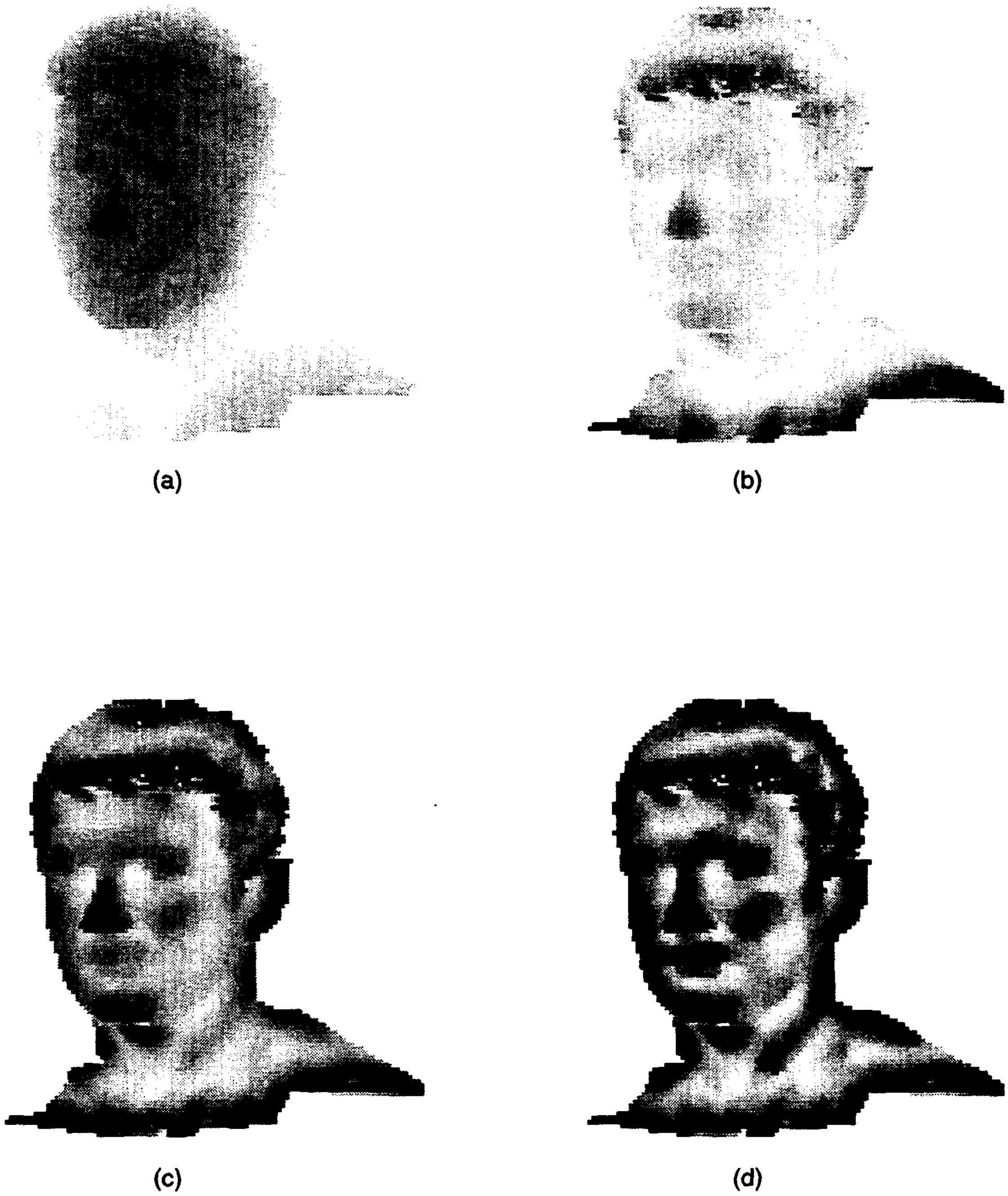


Figure 1: Diffusion processing of a human face. (a) Original. (b–d) Three stages of the diffusion process.

4. Invariant Properties and their Roles in Object Recognition

4.1. Projective and Affine Invariants [1, 10, 16]

Invariants are useful in solving major problems associated with object recognition. For instance, different images of the same object often differ from each other because of the different viewpoints from which they were taken. To match the two images, standard methods thus need to find the correct viewpoint, a difficult problem that can involve search in a large parameter space of all possible points of view and/or finding feature correspondences. Geometric invariants are shape descriptors, computed from the geometry of the shape, that remain unchanged under geometric transformations such as change of viewpoint. Thus, they can be matched without search.

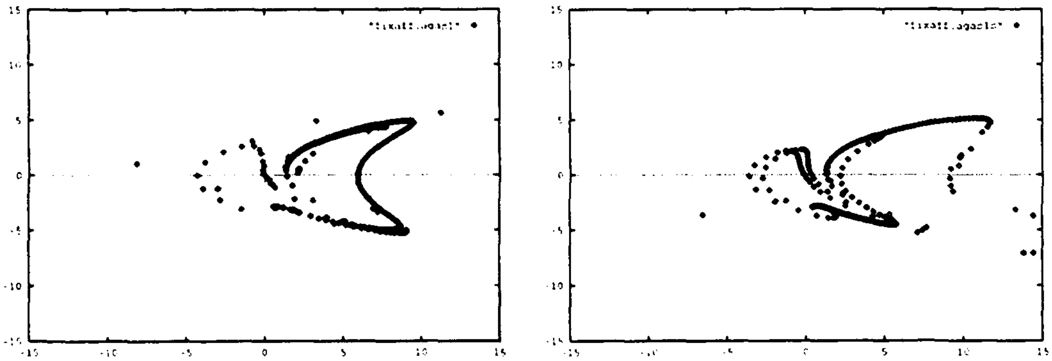
A new and more robust method of obtaining local projective and affine invariants was developed. These shape descriptors are useful for object recognition because they eliminate the search for the unknown viewpoint. Being local, the invariants are much less sensitive to occlusion than the global ones used by others. The basic ideas underlying this method are: (a) employing an implicit curve representation without a curve parameter, thus increasing robustness; (b) using a canonical coordinate system which is defined by intrinsic properties of the shape, independently of any given coordinate system, and is thus invariant. Several shape configurations have been treated using this approach: a general curve without any correspondence, and curves with known correspondences of one or two feature points or lines. The method is applied by fitting an implicit polynomial in a neighborhood of each object contour point. It has been successfully implemented for real images of various two-dimensional objects in three-dimensional space.

4.2. Deformation Invariants [20]

Object recognition means not only recognizing a particular shape but recognizing a class of shapes that are related to each other in some way. For example, two shapes can be regarded as related if one of them can be deformed into the other. The deformation must belong to some predefined set of deformations; it should not be too general. A method of dealing with quasi-affine deformations, i.e. transformations which are approximately linear but also have small non-linear components, was developed. Shape descriptors that are "quasi-invariant" to these deformations were defined and were used to recognize classes of real objects. As an illustration, Figure 2 shows two views of a pear; Figure 3 shows their local affine signatures; Figure 4 shows an image of a banana; Figure 5a shows the two pear signatures, superimposed on one another, and Figure 5b shows the signature of one of the pears superimposed on the signature of the banana. The two pear signatures are very similar, even though the two pear images do not differ by a simple rigid motion; whereas the pear and banana signatures are very different.

(a) (b)

Figure 2: Two views of a pear.



(a) (b)

Figure 3: Affine signatures for the pears in Figure 2.

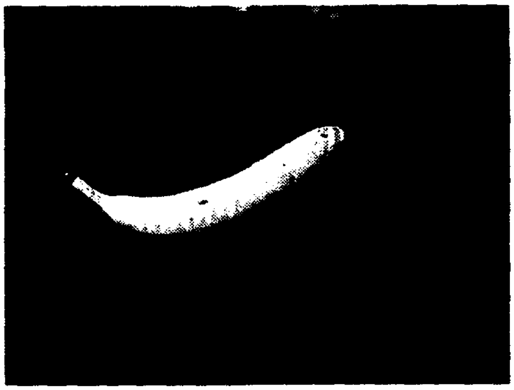
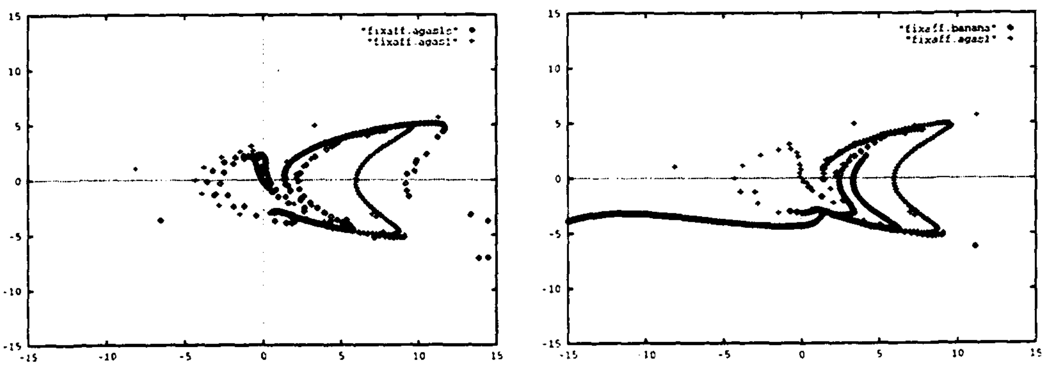


Figure 4: A banana.



(a) (b)

Figure 5: (a) The local affine signatures for the pears in Figure 2 superimposed on one another. (b) The signature of the pear in Figure 2a superimposed on the signature of the banana.

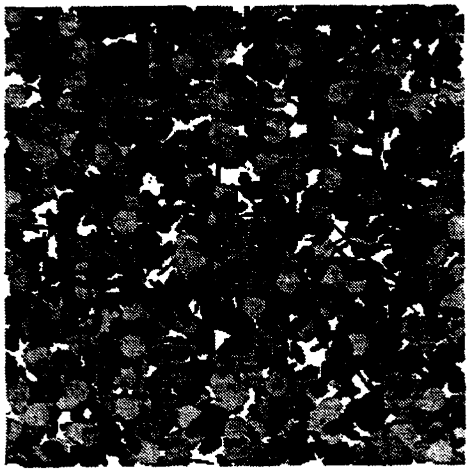
5. Recovery of Three-Dimensional Scene Properties from Single Images

5.1. Reliability of Geometric Computations [2]

The reliability of 3-D interpretations computed from images can be analyzed in statistical terms by employing a realistic model of image noise. First, the reliability of edge fitting can be evaluated in terms of image noise characteristics. Then, the reliability of vanishing point estimation can be deduced from the reliability of edge fitting. The result can then be applied to focal length calibration, and an optimal scheme derived in such a way that the reliability of the computed estimate is maximized. The confidence interval of the optimal estimate can also be computed. The reliability of fitting an orthogonal frame to three orientations obtained by sensing can also be evaluated. Finally, statistical criteria for testing edge groupings, vanishing points, focuses of expansion, and vanishing lines can be derived.

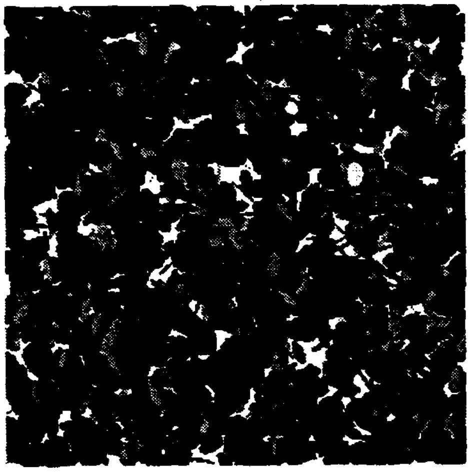
5.2. Three-Dimensional Texture: Foliage [13]

The *distribution* of leaves in a tree crown can be modeled by a random geometric process. For example, one can assume that the leaves are randomly distributed in space, have random spatial orientations, "droop", or face toward the sun. Statistical properties of such distributions can then be derived, including the probability of seeing through the leaves, and the distribution of leaf gray levels under various illumination, reflectivity, and transmissivity models. Figure 6 shows a set of synthetic images of a section of a tree crown, generated using a random spatial distribution of Lambertian disc-shaped leaves, with random or drooping orientations (in the left and right columns, respectively), and frontally illuminated, sidelighted, or backlighted.

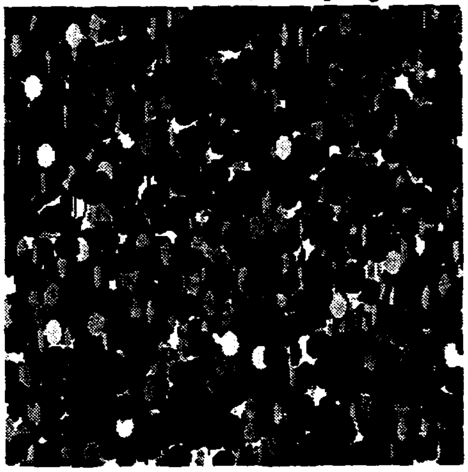


(a) Frontal illumination, random orientation

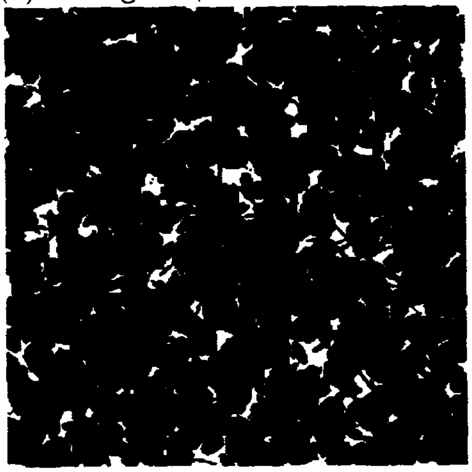
(b) Frontal illumination, drooping orientation



(c) Sidelighted, random orientation



(d) Sidelighted, drooping orientation



(e) Backlighted, random orientation

(f) Backlighted, drooping orientation

Figure 6: Synthetic tree-leaf textures.

6. Recovery of Observer Motion and Scene Structure from Image Sequences

6.1. Monocular and Binocular Recovery of Motion and Structure from Image Features [7, 14, 17]

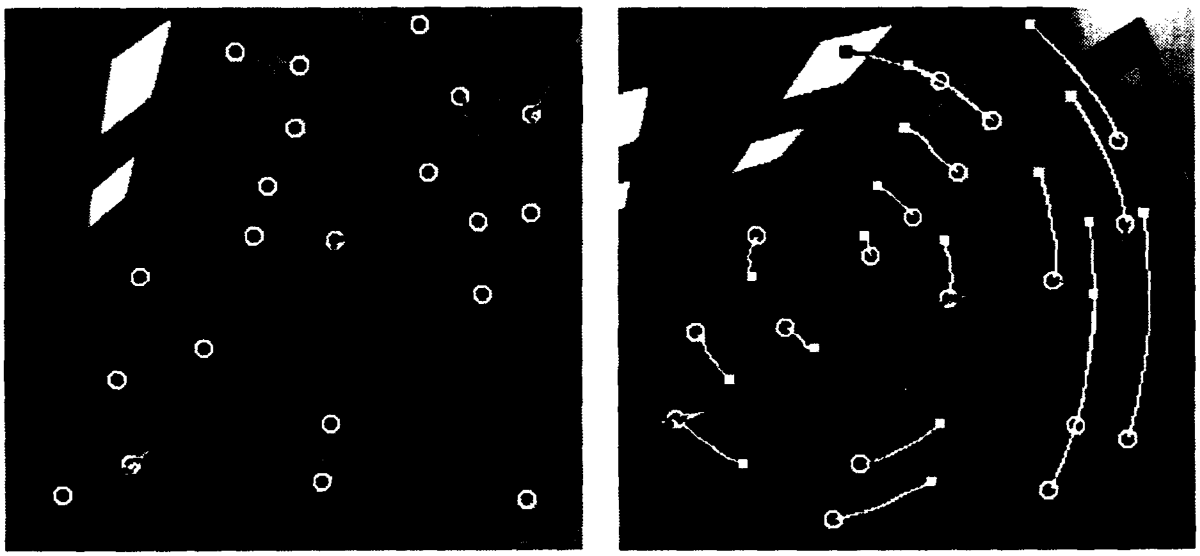
A central problem in vision-based navigation is to use 2-D information from a sequence of images to infer 3-D motion and structure information. By its very nature this problem is ill-posed and most of the algorithms discussed in the literature have proven to be very sensitive to even moderate levels of noise in the images and in the calibration of the camera(s).

Over the last few years, the use of feature-based algorithms and long sequences of images have been advocated for estimating the motion of the observer, the motions of objects, and the spatial structure of feature points. These efforts have resulted in several robust algorithms which have been successfully used for both monocular and binocular real image sequences.

In particular, the problem of estimating the kinematics of the moving camera and the spatial structure of the objects in a stationary environment have been treated. Two estimation techniques, batch and recursive, have been used. The batch technique applies a non-linear least squares method to the stack of images, while the recursive technique uses an iterative extended Kalman filter and analyzes one frame at a time. The approach is based on modeling the motion of the camera using nine parameters, the 3-D coordinates of the rotation center and the linear and angular velocity components, using a perspective camera model. The structure parameters are the 3-D coordinates of the feature points in the inertial coordinate system. These choices of parameters give rise to linear plant models, leading to closed form solutions for the state and covariance transition differential equations. Time consuming numerical integration steps are not needed.

The inputs to the algorithm are feature point correspondences over the image sequence. The task of automatically detecting and tracking features over a long sequence of consecutive frames is a challenging problem when the camera motion is significant. In general, feature displacement over consecutive frames can approximately be decomposed into two components: (a) the displacement due to camera motion, which can be compensated by image rotation, scaling, and translation; (b) the displacement due to object motion and/or perspective projection. The displacement due to camera motion is usually much larger and more irregular than the displacement caused by object motion and perspective deformation.

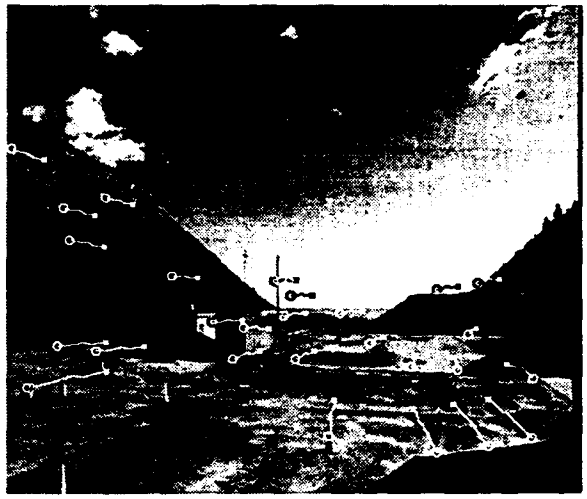
A two-step approach has been developed: First, the motion of the camera is compensated using a recently developed image registration algorithm; then consecutive frames are transformed to the same coordinate system and the feature correspondence problem is solved as one of tracking moving objects using a still camera. Methods of subpixel accuracy feature matching and tracking are employed. The approach results in a robust and efficient algorithm. Results on several real image sequences have been obtained. Figure 7 shows feature points that were automatically detected in the first frame of seven image sequences: a robot arm sequence, a rocket sequence, a traffic cone sequence, and an outdoor sequence, all obtained from the University of Massachusetts; and a coke can and two helicopter sequences, obtained from NASA Ames Research Center. The figure also shows the results of tracking these points for seven frames (in the third and fourth cases, six frames, and in the fifth case, ten frames).



(a) (b)

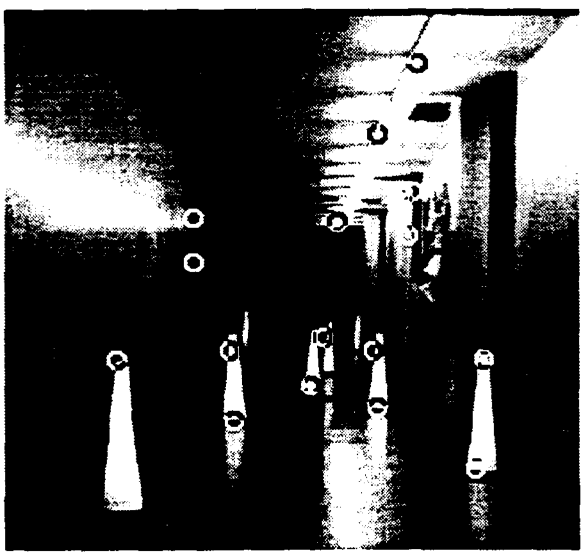
Robot arm sequence. (a) Feature points found in the first frame. (b) Results of tracking the feature points for seven frames.

(c)



(d)

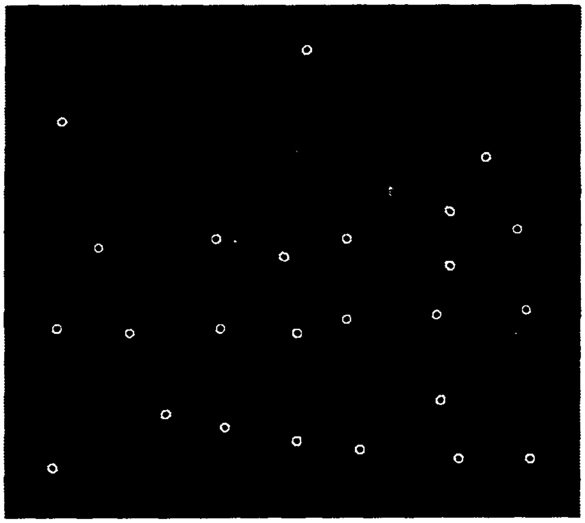
Rocket sequence. (c) Feature points found in the first frame. (d) Results of tracking the feature points for seven frames.

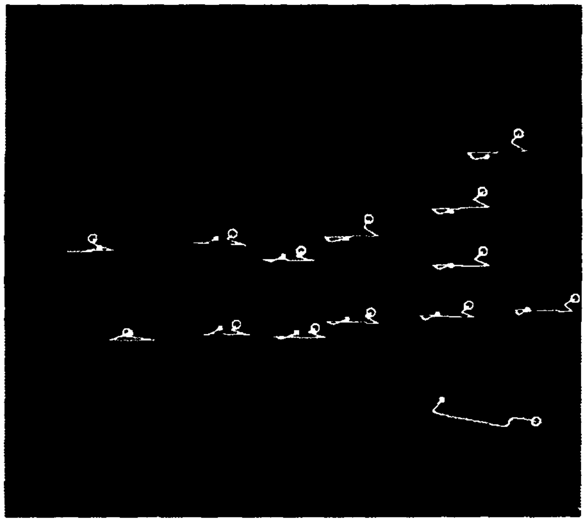


(e)

(f)

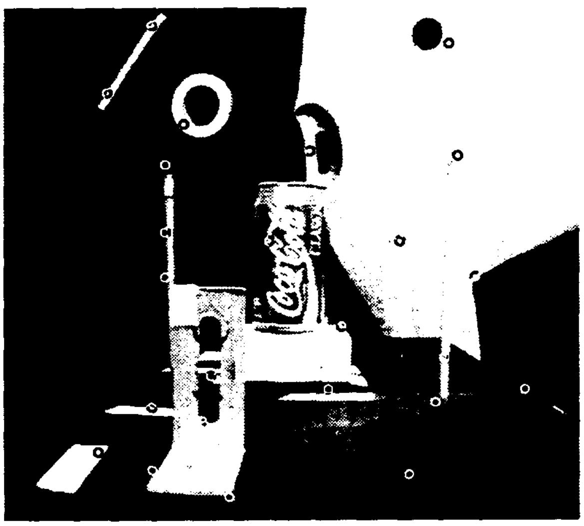
Traffic cone sequence. (e) Feature points found in the first frame. (f) Results of tracking the feature points for seven frames.

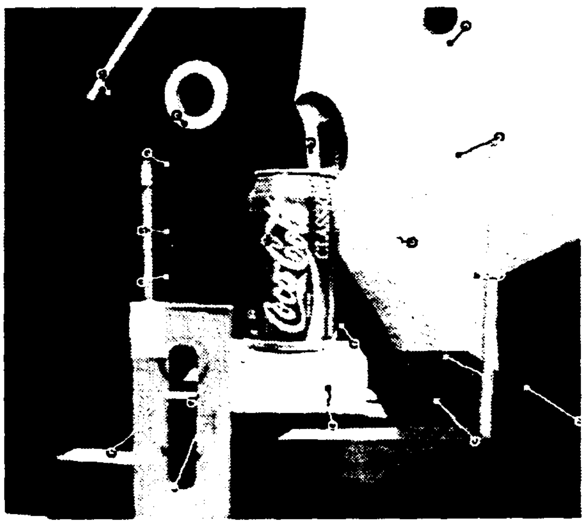




(g) (h)

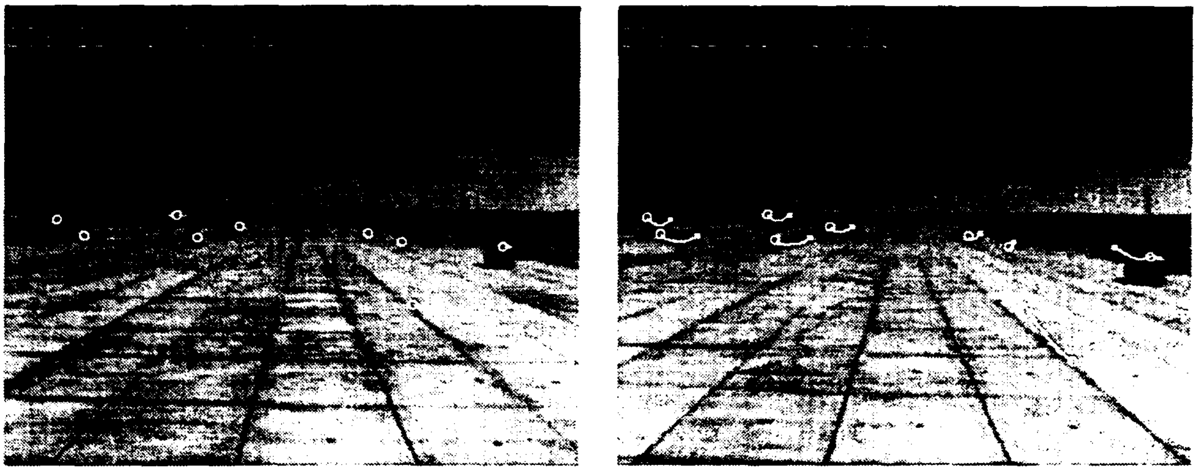
Outdoor sequence. (g) Feature points found in the first frame. (h) Results of tracking the feature points for seven frames.





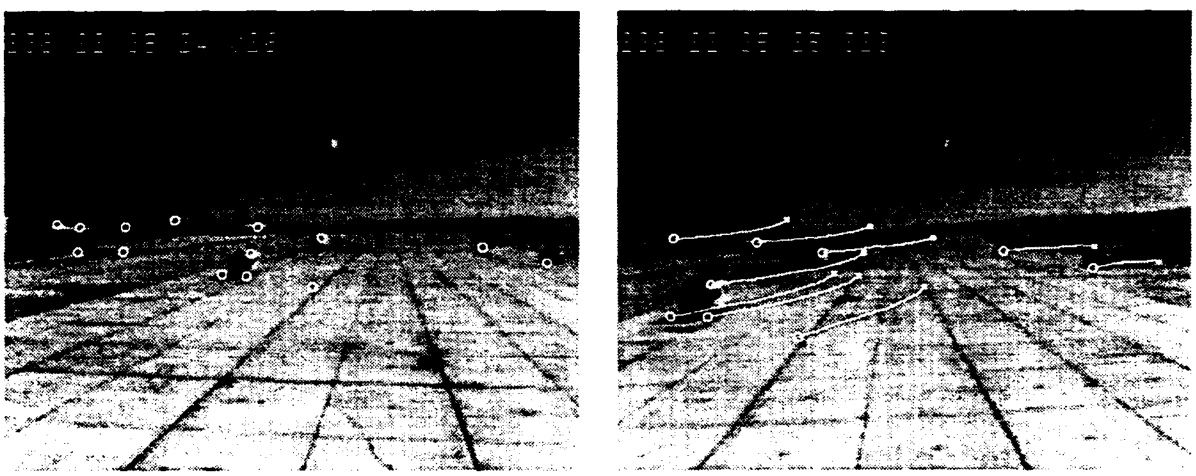
(i) (j)

Coke can sequence. (i) Feature points found in the first frame. (j) Results of tracking the feature points for seven frames.



(k) (l)

Helicopter sequence (line). (k) Feature points found in the first frame. (l) Results of tracking the feature points for seven frames.



(m) (n)

Helicopter sequence (arc). (m) Feature points found in the first frame. (n) Results of tracking the feature points for seven frames.

Figure 7: Feature point trajectories.

The monocular algorithm has also been extended to the case of a binocular moving camera. For binocular imagery, the traditional stereo triangulation method fails when the images are not taken by the two cameras at the same time; but for our algorithm, since asynchronism is allowed, the two cameras can function independently.

6.2. Feature-Based and Flow-Based Motion Estimation: A Unified View [3]

State-of-the-art algorithms for computing 3-D motion from images can make use of either feature correspondences or optical flow. In particular, noise-robust algorithms can be formulated for the feature-based two-view problem—computing the depths of the feature points and the camera motion from correspondences of feature points between two images. For such algorithms, conditions for decomposability and for uniqueness of the solution, as well as direct optimization solutions and "critical surface" conditions, can be formulated. Similarly, noise robust algorithms can be formulated that make use of optical flow; here too, decomposability, uniqueness, direct optimization, and the "critical surface" can be treated, and relationships to the algorithms for finite motion can be analyzed. In both the feature-based and flow-based cases, a simpler treatment can be given for the case of motion on a planar surface.

7. **Direct Motion Analysis** [8,18]

Estimation of 3-D motion directly, without going through the intermediate stage of optical flow or correspondence estimation, has also been studied. The inputs that have been utilized in this approach are the spatiotemporal derivatives of the image intensity function (the normal flow).

From measurements on the image only the relative motion between the observer and any point in the 3-D scene can be computed. The model that has usually been employed in previous research to relate 2-D image measurements to 3-D motion and structure is that of rigid motion. Consequently, egomotion recovery for an observer moving in a static world has been treated in the same way as the estimation of an object's 3-D motion relative to an observer. The rigid motion model is appropriate if only the observer is moving, but it holds only for a restricted subset of moving objects, mainly man-made ones. Indeed, virtually all objects in the natural world move non-rigidly. However, if we consider only a small patch in the image of a moving object, a rigid motion approximation is legitimate. For the case of egomotion, data from all parts of the image plane can be used, whereas for object motion, only local information can be employed. Therefore, conceptually different techniques were developed for explaining the mechanisms underlying the perceptual processes of egomotion recovery and 3-D object motion recovery.

Specifically, solutions to the following problems were developed: (a) *Given an active observer viewing an object moving in a rigid manner (translation + rotation), recover the direction of the 3-D translation and the time to collision by using only the spatiotemporal derivatives of the image intensity function.* Although this problem is not equivalent to "structure from motion" because it does not fully recover the 3-D motion, it is of importance in a variety of situations. If an object is rotating around itself and also translating in some direction, we are usually interested in its translation—for example, in problems related to tracking, prey catching, interception, obstacle avoidance, etc. The basic idea of this motion parameter estimation strategy lies in the employment of fixation and tracking. Fixation simplifies much of the computation by placing the object at the center of the visual field, and the main advantage of tracking is the accumulation of information over time. Methods of tracking using normal flow measurements have been demonstrated, and have been used for two different tasks in the solution process: First, as a tool to compensate for the lack of existence of an optical flow field, and to estimate the translation parallel to the image plane; and second, to gather information about the motion component perpendicular to the image plane. (b) *Given an active observer moving rigidly in a static environment, recover the direction of its translation and its rotation.* This is the task of passive navigation, a term used to describe the set of processes by which a system can estimate its motion with respect to the environment.

The approach to egomotion estimation is based on a geometric analysis of the properties of the normal flow field. The fact that the motion is rigid defines geometric relations between certain values of the spatiotemporal derivatives of the image intensity function. It can be shown that the normal flow gives rise to global patterns in the image plane. The geometry of these patterns is related to the three dimensional motion parameters. By locating some of these patterns, which depend only on subsets of the motion parameters, using a simple search technique, the 3-D motion parameters can be found. The algorithmic procedure developed for doing this is provably robust, since it is not affected by small perturbations in the local image motion measurements. In fact, since only the signs of the normal flow measurements are employed, the direction of translation and the axis of rotation can be estimated in the presence of up to 100% error in the image measurements.

8. Visual Interception [5]

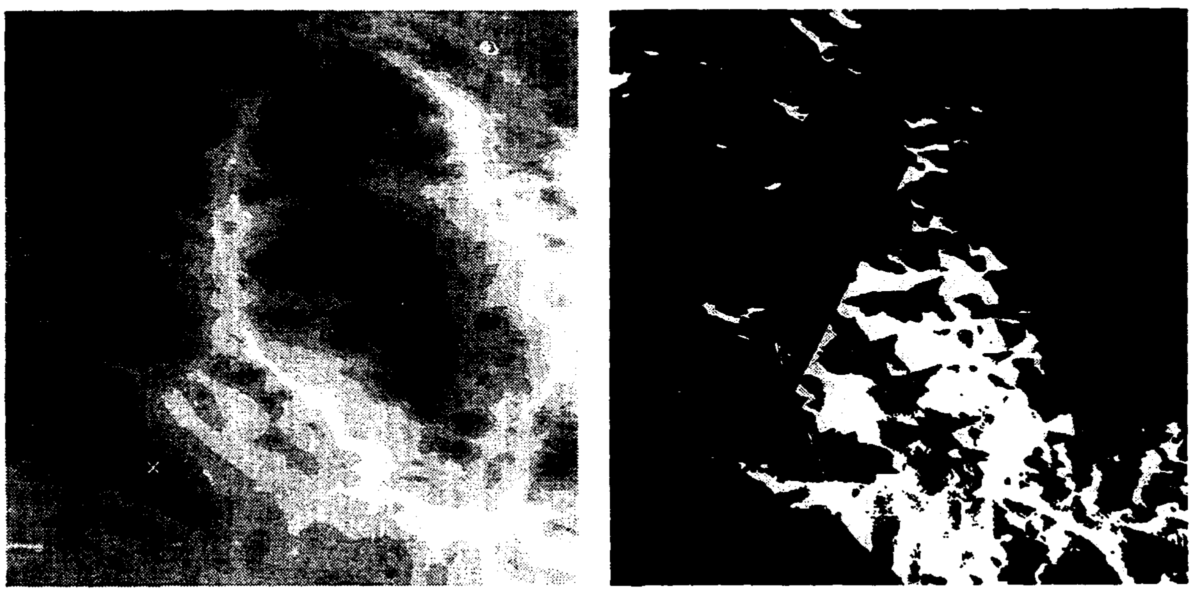
A visual interception system consists of a camera(s), an agent, a target and a mind. The mind uses information from the camera in order to generate the control of the agent so that it will intercept the target. Under the traditional paradigm of considering vision as a recovery problem, visual interception is just another application of the structure from motion module: The module reconstructs the three dimensional positions and velocities of the camera, the agent and the target and then the information is utilized by a planning module to generate correct control of the agent. However, even if such three dimensional reconstruction problems are possible, they are expensive. The inherent difficulties associated with the structure from motion problem have delayed any real time applications, and no general visual interception system is known to exist to date.

Robust solutions to the problem of visual interception under the active qualitative vision paradigm have been developed. The geometry of visual interception does not have to rely on depth. From the image intensity function, the *locomotive intrinsics* of the agent and the target are obtained. Based on this relative information, a control strategy is defined that decides in real time and on the basis of the image intensity function whether the velocity of the agent should be increased or decreased at any time instant, thus guiding the agent to intercept the target. The problem of visual interception can thus be solved using only the spatiotemporal derivatives of the image intensity function, and no correspondence is necessary. The computation is simple and can be performed in real time.

9. Vision-Based Navigation

9.1. Visibility on Terrain [6]

Two classes of parallel algorithms have been investigated for point-to-region visibility analysis on terrain: ray-structure-based methods and propagation-based methods. A new propagation-based algorithm has been developed which avoids problems commonly occurring with such algorithms. The performance and characteristics of the two kinds of algorithms have been compared. The sources of uncertainty in visibility computation and the importance of taking uncertainty into consideration have been analyzed. Different methods for representing the uncertainty have been studied, including Monte Carlo simulation, analytic estimation, and some simple heuristic indicators. Experiments show that these indicators can be used for efficient coarse classification of the likelihood of point intervisibility. Figure 8 shows a digital terrain model (left; darker pixels are higher), with the viewpoint marked $\times$, and a plot of the pixels visible from that viewpoint (right; white pixels are visible).



(a) (b)

Figure 8: (a) Digital terrain. (b) Points visible from "$\times$".

9.2. Landmark-Based Localization [9]

A method of landmark-based localization and positioning has been developed. Localization is defined as the act of recognizing the environment, and positioning as the act of computing the exact coordinates of a robot in the environment. The method is based on representing the scene as a set of 2-D views and predicting the appearances of novel views by linear combinations of the model views. The method accurately approximates the appearance of scenes under weak perspective projection. Analysis of this projection as well as experimental results demonstrate that in many cases this approximation is sufficient to accurately describe the scene. When the weak perspective approximation is invalid, either a larger number of models can be acquired or an iterative solution to account for the perspective distortions can be employed. The method has several advantages over other approaches. It uses relatively rich representations; the representations are 2-D rather than 3-D; and localization can be done from only a single 2-D view. The same general method is applied to both the localization and positioning problems. A simple algorithm for the task of returning to a previously visited position defined by a single view can also be derived from this method.

10. Bibliography of Reports Under this Contract

1. Isaac Weiss, "Local Projective and Affine Invariants." CAR-TR-612, CS-TR-2870, April 1992.

2. Kenichi Kanatani, "Statistical Reliability of 3-D Interpretation from Images." CAR-TR-615, CS-TR-2873, April 1992.

3. Kenichi Kanatani, "Computation of 3-D Motion from Images." CAR-TR-617, CS-TR-2879, April 1992.

4. Ling Tony Chen and Larry S. Davis, "Parallel Stereo Matching Using Dynamic Programming." CAR-TR-620, CS-TR-2884, April 1992.

5. Liuqing Huang and Yiannis Aloimonos, "The Geometry of Visual Interception." CAR-TR-622, CS-TR-2893, April 1992.

6. Y. Ansel Teng and Larry S. Davis, "Visibility Analysis on Digital Terrain Models and Its Parallel Implementation." CAR-TR-625, CS-TR-2900, May 1992.

7. Qinfen Zheng and Rama Chellappa, "Automatic Feature Point Extraction and Tracking in Image Sequences for Arbitrary Camera Motion" CAR-TR-628, CS-TR-2911, June 1992.

8. Cornelia Fermüller and Yiannis Aloimonos, "Qualitative Egomotion." CAR-TR-629, CS-TR-2915, June 1992.

9. Ehud Rivlin and Ronen Basri, "Localization and Positioning using Combinations of Model Views." CAR-TR-631, CS-TR-2926, July 1992.

10. Isaac Weiss, "Geometric Invariants and Object Recognition." CAR-TR-632, CS-TR-2942, August 1992.

11. Yaser Yacoob and Larry S. Davis, "Early Vision Processing using a Multi-Stage Diffusion Process." CAR-TR-633, CS-TR-2943, August 1992.

12. P. J. Narayanan, "Effective Use of SIMD Machines for Image Analysis." CAR-TR-635, CS-TR-2945, August 1992.

13. Adlai Waksman and Azriel Rosenfeld, "Sparse, Opaque Three-Dimensional Texture, 2: Foliate Patterns." CAR-TR-637, CS-TR-2951, September 1992.

14. Ting-Hu Wu and Rama Chellappa, "Experiments on Estimating Motion and Structure Parameters Using Long Monocular Image Sequences." CAR-TR-640, CS-TR-2969, October 1992.

15. Yaser Yacoob and Larry S. Davis, "Qualitative Labeling of Human Face Components from Range Data." CAR-TR-642, CS-TR-2971, October 1992.

16. Ehud Rivlin and Isaac Weiss, "Local Invariants for Recognition." CAR-TR-644, CS-TR-2977, October 1992.

17. Ting-Hu Wu and Rama Chellappa, "Stereoscopic Recovery of Egomotion and Environmental Structure: Models, Uniqueness and Experiments." CAR-TR-646, CS-TR-2988, November 1992.

18. Cornelia Fermüller and Yiannis Aloimonos, "The Role of Fixation in Visual Motion Analysis." CAR-TR-647, CS-TR-2989, November 1992.

19. Yaser Yacoob, Larry S. Davis and Hanan Samet, "Scale Space Interpolation and Metamorphosis." CAR-TR-654, CS-TR-3016, January 1993.

20. Ehud Rivlin and Isaac Weiss, "Recognizing Objects Using Deformation Invariants." CAR-TR-660, CS-TR-3041, March 1993.