(2)

ETL - 0432

AD-A171 618

Vision - based navigation for autonomous ground vehicles - 1986 annual report

Larry Davis

Center for Automation Research
University of Maryland
College Park, Maryland 20742

August 1986

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

Prepared for
U.S. ARMY CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060 - 5546

and
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BOULEVARD
ARLINGTON, VIRGINIA 22209 - 2308

E T L

U.S. ARMY ENGINEER TOPOGRAPHIC LABORATORIES

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A171 618

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp. Date Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) ETL-0432
6a. NAME OF PERFORMING ORGANIZATION University of Maryland 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U.S. Army Engineer Topographic Labs
6c. ADDRESS (City, State, and ZIP Code) Computer Vision Lab, Center for Automation Research, University of Maryland College Park, MD 20742	7b. ADDRESS (City, State, and ZIP Code) Research Institute Telegraph & Leaf Roads Fort Belvoir, VA 22060-5546
8a. NAME OF FUNDING / SPONSORING ORGANIZATION DARPA 8b. OFFICE SYMBOL (If applicable) ISTO	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DACA76-84-C-0004
8c. ADDRESS (City, State, and ZIP Code) 1400 Wilson Boulevard Arlington, VA 22209-2308	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: 62301E PROJECT NO.: TASK NO.: WORK UNIT ACCESSION NO.:

11. TITLE (Include Security Classification)

Vision-Based Navigation for Autonomous Ground Vehicles - 1986 Annual Report

12. PERSONAL AUTHOR(S)

Davis, Larry

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM 85/7/1 TO 86/6/30	86 August	10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	
17	07		Autonomous navigation Road following Computer vision
17	08		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This is an annual report for the project "Vision-based navigation for autonomous ground vehicles" being conducted under Contract DACA76-84-C-0004 (DARPA Order 5096) for the period 1 July 1985 through 30 June 1986. The project to date has focused on three tasks: (1) Support of Martin Marietta as the Autonomous Land Vehicle (ALV) integrating contractor, (2) development of a vision system for autonomous navigation of roads and road networks, and (3) experiments using this vision system on the ALV. Progress on each of these topics is described in the following subsections.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Rosalene M. Holecheck	22b. TELEPHONE (Include Area Code) (202) 355-2769 22c. OFFICE SYMBOL ETL-RI

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

A

ANNUAL REPORT
for the project
Vision based navigation for autonomous ground vehicles

Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 207~~74~~42

This is the annual report for the project **"Vision-based navigation for autonomous ground vehicles"** being conducted under Contract DACA76-84-C-0004 (DARPA Order 5096) for the period 1 July 1985 through 30 June 1986.

Our project to date has focused on three tasks:

1) Support of Martin Marietta
2) Development of a vision system for autonomous navigation of roads and road networks.
3) Experiments using this vision system on the ALV.

We describe progress on each of these topics in the following subsections.

Support of Martin Marietta

The University of Maryland has worked closely with Martin Marietta since the inception of the ALV program. During the first year, Maryland acquired a VICOM image processor identical in configuration to the systems in the laboratory and on the ALV at Martin Marietta. This past year, Maryland received a Butterfly parallel processor, one of the Strategic Computing architectures scheduled for installation on the ALV this summer; and Maryland will be one of the first sites to receive a WARP array processor from General Electric. Maryland has done extensive experimentation with its Butterfly system in order to provide substantial support to Martin programmers concerned with the Butterfly during the coming year.

Maryland has hosted scientists and engineers from Martin Marietta for various periods of time during the first two years of the project. These visitors have participated in the research program at Maryland and have brought large segments of the Maryland Vision system to Denver. For example, the first demonstration of the ALV in May 1985 used an inverse perspective model developed by

Allen Waxman for constructing a three dimensional model of the road from a single monocular image of the road; such a model was needed to generate the piloting commands to the vehicle.

Maryland has also assisted Martin Marietta in the collection of data sets for distribution to the Strategic Computing vision community. During January–March 1986, Dr. Todd Kushner of our laboratory spent several weeks at Martin Marietta developing software on the VICOM that enabled us to simultaneously collect imagery from the ALV's sensors as well as position information from its Land Navigation Systems. Such data sets are critical for testing any vision system that integrates information across several images, as does the Maryland vision system.

Finally, Maryland and Martin Marietta were linked by a video teleconferencing facility that was installed under DARPA support at both sites. This will support even more frequent consultation between Maryland and Martin Marietta personnel concerning vision problems and potential solutions.

Maryland vision system

The Computer Vision Laboratory at the University of Maryland under the direction of Prof. Larry Davis and Dr. Allen Waxman (now at Boston University) has developed a vision system for navigation of roads and road networks. It is briefly described below.

The vision executive module (see Fig. 1) is responsible for the overall coordination of the system. It represents a centralized source of control that is responsible for scheduling the activities of all of the vision and reasoning processes in the system.

The visual knowledge base module implements rule-based reasoning for two separate vision tasks: seeking significant groupings of symbols derived from an image and checking consistency of 3-D shape recovery with models of objects (e.g., roads). Many types of symbolic groupings can be considered in general, although for purposes of road following we concentrated on the groupings of linear features into "pencils" (i.e., concurrent sets) of lines.

The geometry module converts the grouped symbolics in the image domain into a viewer-centered 3-D description of objects in the scene. Our system utilizes a single method for recovering shape from monocular imagery. It is a model-driven "shape from contour" method that inverts the perspective transformation of the imaging process on the basis of the following three (road) model-based assumptions: (1) Pencils in the image domain correspond to planar parallels in the world. (2) Continuity in the image domain implies continuity in the world. (3) The camera sits above the first visible ground plane (at the bottom of the image). Our 3-D reconstruction will then characterize a road of constant width, with turns and changes in slope and bank. The reconstruction process amounts to an integration out from beneath the vehicle into the distance, along local parallels, over topography modeled as a sequence of planar surface patches.

The 3-D representation module converts the 3-D viewer-centered representation of the road scene into an object-centered description consisting of a list of road attributes at each "roadmark" set down. Roadmarks are placed (by this module) along the centerline of the reconstructed road model at the beginning of each new planar patch.

The scene predictor module is used to focus the attention of the system onto a small part of the visual field. Windows in the field of view are determined, inside of which the near parts of the road boundaries are located, following the vehicle's blind travel through the 3-D representation created from an earlier image. This is accomplished by essentially transforming the 3-D data used to create the most recent representation according to a rigid-body translation and rotation associated with the vehicle's motion (as is familiar in computer graphics). The components of travel used in constructing this transformation are obtained from the dead reckoning (and inertial navigation) system aboard the vehicle.

The navigator module provides computational capabilities in support of generalized path planning. Our current navigator module is specific to following obstacle-free roads (corridors), so path planning consists of smoothly changing the heading of the vehicle in accordance with the 3-D representation derived from the visual process. This is accomplished by computing cubic arcs as asymptotic paths.

The pilot module converts the cubic arcs obtained from the navigator module into a sequence of conventional steering commands used over the next several meters. They decompose a curved path into a set of short, straight line segments. These motion commands must then interface to the motor controls of the vehicle. (In our case, the interface is to the motion-control software of a robot arm.) The pilot module is also responsible for sending dead reckoning data from the vehicle to the navigator module. The pilot module converts the raw data into measured travel and returns this information to the navigator module several times per second.

Our current planner module merely specifies a "distance goal," e.g., "move to the point 60 meters farther down the road." A mature planning module would have greater complexity, specifying high level navigation goals, assigning priorities to these goals, monitoring progress as a function of time, and constructing contingency plans.

The image processing module transforms an input image into a symbolic representation of the boundaries of the roads in the field of view. It runs in one of two modes: a bootstrap mode or a feed-forward mode. The bootstrap mode is used to develop an initial representation of the road on which the vehicle is to travel. Since we assume, at this point, that aside from map information the vehicle has no preconceptions about where the road will be in its field of view or what the detailed structure of that road is (e.g., single lane, with or without shoulders, lane markings, etc.), the bootstrap image processing performs a global analysis of the image to identify significant global linear features. These linear features are grouped into elements called pencils (concurrent lines in the image plane), which are the units reasoned about by the visual knowledge base and geometry modules.

During continuous operation, of course, the system has fairly specific expectations concerning the position and appearance of the road. These expectations are generated by the scene predictor module, which, based on a 3-D model of the road constructed by the geometry and visual knowledge base modules and an estimate of the travel between consecutive frames obtained from an inertial navigation system (INS), generates a prediction of where the boundaries of the road will appear near the bottom of the current frame. This prediction is used to constrain the analysis of the image processing operators in the so-called feed-forward mode of operation. Here, based on the prediction of where the road boundaries will appear, the vision executive module identifies small windows in the image that will contain pieces of the left and right road boundaries and, using a tightly constrained analysis (since both the geometric and photometric properties of large pieces of the road can be carried forward from the analysis of previous frames), identifies the projections of the road boundaries through those windows. One the basis of the computed locations of the road boundaries, subsequent windows are placed, and the road is tracked through the image.

Experiments on the ALV

A large subset of the Maryland vision system was reimplemented to run on the VICOM image processor and brought to Martin Marietta by Todd Kushner during the Fall of 1985. Dr. Kushner spent almost two months in Denver, first integrating this software into the ALV laboratory environment, and finally operating the ALV itself under program control. The programs were able to drive the ALV over a portion of the test track at an average speed of 3 kilometers per hour. It is important to note that the subset of the system brought to Denver did not include the prediction mechanism from the Maryland vision system. Thus, the initial windows placed by the programs were unnecessarily large, and reduced the speed at which the programs could drive the ALV. We are currently planning for Dr. Kushner to spend several weeks in Denver during August 1986 to test an enhanced version of the VICOM system. This version of the system will not only include the prediction mechanism (which will allow us to focus the attention of the system on smaller initial windows, and therefore increase the overall speed of the system) but will also include enhancements that will improve both the speed and reliability of the system.

1) We have added an additional image processing module for road boundary detection based on thresholding the pixels within the windows identified by the prediction module. The motivation for considering thresholding as an alternative to edge detection was the success that Martin Marietta had in identifying the road using thresholding techniques. However, unlike the Martin Marietta algorithms which threshold the entire image using a single threshold, the Maryland algorithms choose a threshold adaptively in each window based on the statistics of the pixels in both that window and the previous window on the appropriate side of the road. Our experience in the laboratory indicates that this algorithm is often more accurate and always more efficient than the edge detection based algorithm.

2) Currently, our VICOM algorithms operate on a 256×256 sampled version of the 512×512 television frame. If we were to further decrease the size of the image processed, we would further reduce the computation time of the system. Of course, as the image size is reduced, so is the accuracy with which the road can be identified in three dimensions, since pixels in the reduced resolution image correspond to larger and larger patches of road. Dr. Kushner has modified his programs so that we can now specify the image spatial sampling rate; all other parameters on which the program is based (i.e., window size, angular quantization in various search procedures) are represented in the program as functions of this sampling rate. We will therefore be able to perform experiments on the ALV that will allow us to measure the trade-off between image spatial resolution and the reliability/accuracy of the resulting visual analysis. Eery time the spatial resolution of the image is reduced by a factor of 2 in each dimension the speed of the program is increased by a factor of 4. Thus, it is important to identify the minimal spatial sampling rate that will result in a sufficiently accurate extraction of the road for navigation.

3) The current version of the system does not construct a three dimensional model of the road until the entire image has been processed. This has the unfortunate consequence that if the tracking process ever loses either boundary of the road there is really little hope that it will ever recover. Therefore, much time is lost processing windows of the image that may not even include the road boundaries. This problem can be avoided to a large extent by developing a tighter interaction between image processing and the inverse

perspective models for building the three dimensional model of the road. Quite simply, whenever the road is extended in the image by tracking through a new set of windows, one on the left and the other on the right, a three dimensional model for that extension is developed using one of the inverse perspective models. Three dimensional properties of this model can then be compared against expectations concerning road properties (for example, road width, road orientation, etc.). If the model fails to meet any of these expectations, then we can either consider alternative extensions of the road through the image, or terminate processing of this image and acquire a new image in which we can, hopefully, track the road further. It is this latter alternative that we will be experimenting with this summer.

Finally, it should be noted that most of the processing time in our VICOM system is spent in the image processing of the windows identified by the feedforward focus of attention mechanism. While some of these operations are supported directly by the VICOM's image processing hardware, many of them require using the general purpose host 68000. We have reimplemented many of these algorithms on our Butterfly and have achieved impressive improvements in operating speed. We are currently developing a VICOM/Butterfly system which will have the same functionality as the system that Dr. Kushner will integrate on the ALV this summer, but which should operate at much greater speeds.

Reports

The following reports were produced:

1. Jacqueline Le Moigne, Allen M. Waxman, Babu Srinivasan and Matti Pietikainen, "Image Processing for Visual Navigation of Roadways." CAR-TR-138, CS-TR-1536, DACA76-84-C-0004, July 1985.

ABSTRACT: A system which supports the visual navigation of roadways by an autonomous land vehicle has been developed at the Computer Vision Laboratory. One of the modules of this system is an Image Processing Module which extracts 2-D symbolics from the imagery to be analyzed in the world domain by Reasoning and Geometry Modules. In this report, we describe the Image Processing Module. Different representations can be used in the image domain: boundary-based and region-based are two examples of such representations. We present two kinds of algorithms for extracting roads from imagery, corresponding to these two different representations: linear feature extraction and gray-level or color segmentation. For each kind of processing, two different modes, called "bootstrap" and "feed-forward", may be utilized. The bootstrap mode processes the entire image and assumes no prior information about the location of a road, while the feed-forward mode utilizes a prediction derived from the processing of a prior road segment and the distance traversed in order to focus visual attention. These algorithms are described in detail, and example results are shown. The examples include real road images and "simulator images" obtained in our laboratory with a scale model system comprised of a road network and a robot arm carrying a camera.

2. Allen M. Waxman, Jacqueline Le Moigne, Larry S. Davis, Eli Liang and Tharakesh Siddalingaiah, "A Visual Navigation System for Autonomous Land Vehicles." CAR-TR-139, CS-TR-1537, DACA76-84-C-0004, July 1985.

ABSTRACT: A modular system architecture has been developed to support visual navigation by an autonomous land vehicle. The system consists of vision modules performing image processing, 3-D shape recovery, and rule-based reasoning, as well as modules for planning, navigating and piloting. The system runs in two distinct modes, *bootstrap* and *feed-forward*. The bootstrap mode requires analysis of entire images in order to find and model the objects of interest in the scene (e.g., roads). In the feed-forward mode (while the vehicle is moving), attention is focused on small parts of the visual field as determined by prior views of the scene, in order to continue to track and model the objects of interest. We have decomposed general navigational tasks into three categories, all of which contribute to planning a vehicle path. They are called *long, intermediate, and short range navigation*, reflecting the scale to which they apply. We have implemented the system as a set of concurrent, communicating modules and use it to drive a camera (carried by a robot arm) over a scale model road network on a terrain board.

3. Larry S. Davis and Todd Kushner, "Road Boundary Detection for Autonomous Vehicle Navigation." CAR-TR-140, CS-TR-1538, DACA76-84-C-0004, July 1985.

ABSTRACT: The Computer Vision Laboratory at the University of Maryland has been participating in DARPA's Strategic Computing Program for the past year. Specifically, we have been developing a computer vision system for autonomous ground navigation of roads and road networks. The complete system runs on a VAX 11/785, but certain parts of the system have been reimplemented on a VICOM image processing system for experimentation on an autonomous vehicle built for the Martin Marietta Corp., Aerospace Division in Denver, Colorado. We give a brief overview here of the principal software components of the system, and then describe the VICOM implementation in detail.

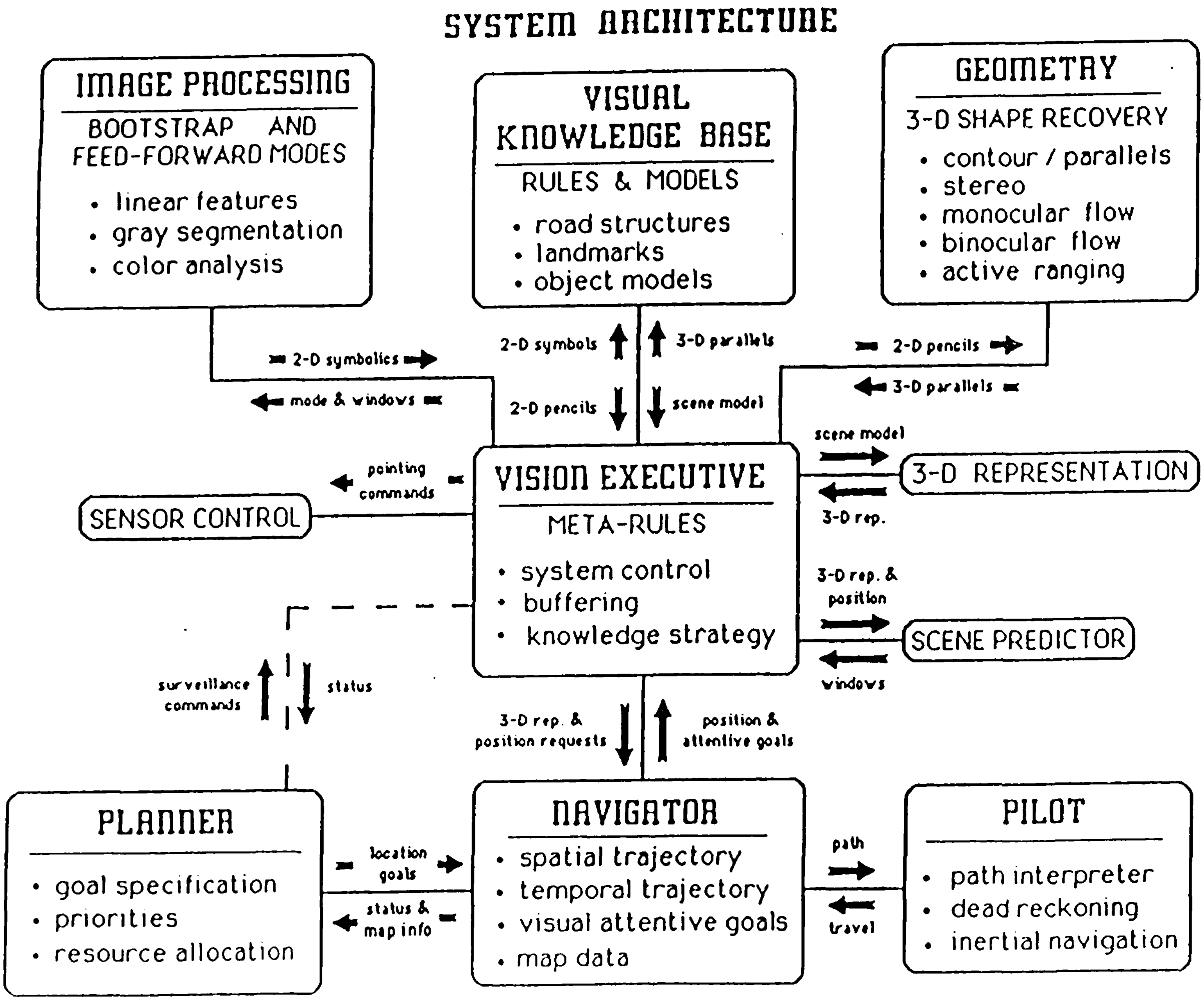

Figure 1. A block diagram of the system