QUALITATIVE LANDMARK-BASED PATH PLANNING AND FOLLOWING

Tod S. Levitt, Daryl T. Lawton, David M. Chelberg, Philip C. Nelson

Advanced Decision Systems, Mountain View, California 94040

ABSTRACT

This paper develops a theory for path planning and following using visual landmark recognition for the representation of environmental locations. It encodes local perceptual knowledge in structures called viewframes and orientation regions. Rigorous representations of places as visual events are developed in a uniform framework that smoothly integrates a qualitative version of path planning with inference over traditional metric representations. Paths in the world are represented as sequences of sets of landmarks, viewframes, orientation boundary crossings, and other distinctive visual events. Approximate headings are computed between viewframes that have lines of sight to common landmarks. Orientation regions are range-free, topological descriptions of place that are rigorously abstracted from viewframes. They yield a coordinate-free model of visual landmark memory that can also be used for path planning and following. With this approach, a robot can opportunistically observe and execute visually cued "shortcuts".

1. INTRODUCTION

The questions that define the problems of path planning and following are: "Where am I?", "Where are other places relative to me?", and "How do I get to other places from here?". A robot that moves about the world must be able to compute answers to these questions. This paper is concerned with the structure and processing for robotic visual memory that yields visual path inference. The input data is assumed to be percepts extracted from imagery, and a database, i.e., memory, of models for visual recognition. A priori model and map data is only relevant insofar as it provides a basis for runtime recognition of observable events. This is distinguished from path traversability planning where the guidance questions concern computing shortest distances between points under constraints of support of the ground or surrounding environment for the robotic vehicle.

Existing robot navigation techniques include triangulation [Matthies and Shafer, 1986], ranging sensors [Hebert and Kanade, 1986], auto-focus [Pentland, 1985], stereo techniques [Lucas and Kanade, 1984], dead reckoning, inertial navigation, geo-satellite location, correspondence of map data with the robot's location, and local obstacle avoidance techniques. These approaches tend to be brittle [Bajcsy et al., 1986], accumulate error [Smith and Cheeseman, 1985], are limited by the range of an active sensor, depend on accurate measurement of distance/direction perceived or traveled, and are non-perceptual, or only utilize weak perceptual models.

Furthermore, these theories are largely concerned with the problem of measurement and do not centrally address issues of map or visual memory and the use of this memory for inference in vision-based path planning or following. Exceptions to this are the work of [Davis, 1986], [McDermott and Davis, 1984], and [Kuipers, 1977]. Davis addressed the problem of representation and assimilation of 2D geometric memory, but assumed an orthographic view of the world and did not consider navigation or guidance. McDermott and Davis developed an ad hoc mixture of vector and topological based route planning, but assumed a map, rather than vision derived world (in their assumptions of knowledge of boundaries, their shapes, and spatial relationships), had no formal theory relating the multiple levels of representation, and consequently did not derive or implement results about path execution. Kuipers developed qualitative techniques for path planning and following that were the inspiration for our approach. He assumed capability of landmark recognition, as we do, but relied on dead-reckoning and constraint to one-dimensional (road) networks to permit path planning and execution.

We develop representation and inference for relative geographic position information that: build a memory of the environment the robot passes through; contains sufficient information to allow the robot to retrace its paths; can be used to construct or update an a posteriori map of the geographic area the robot has passed through; and can utilize all available information, including that from runtime perceptual inferences and a priori map data, to perform path planning and following. The robust, qualitative properties and formal mathematical basis of the representation and inference processes presented herein are suggestive of the path planning and following behavior in animals and humans [Schone, 1984]. However, we make no claims of biological foundations for this approach.

2. TOPOLOGICAL LANDMARK NETWORK REPRESENTATIONS

A viewframe encodes the observable landmark information in a stationary panorama. To generate a viewframe, relative solid angles between distinguished points on landmarks are computed using a sensor-centered spherical coordinate system. We can pan from left to right, recognizing landmarks, L_i, storing the solid angles between landmarks in order, denoting the angle between the i-th and j-th landmarks by Ang_{ij}. The basic viewframe data are these two ordered lists, $(L_1, L_2, \ldots)$ and $(\mathrm{Ang}_{12}, \mathrm{Ang}_{23}, \ldots)$. The relative angular displacement between any two landmarks can be computed from this basic list. The solid angular error is measured by e_{ij} between landmarks i and j. Finally, range estimates for landmarks are recorded as intervals in viewframes. We only require that the true range lie between the bounds specified for the estimate. The range interval associated to landmark L_i is denoted $[r_{i1}, r_{i2}]$. We now explain how it is possible to localize ourselves in space relative to these observed landmarks.

We begin by noting that the set of points in 3-space from which we can observe an angle of θ_{ij} between landmarks L_i and L_j is constrained to a closed torus-like surface; a cut-away of this surface is pictured in Figure 1(a). In terms of the variables pictured in Figure 1(b), we compute:

$$r_1 = s_{12}\cos\phi + s_{12}\cot\theta\sin\phi$$

$$\phi = \cot^{-1}\left[\frac{r_1}{r_2}\csc\theta - \cot\theta\right]$$

The first equation is the polar form for a circle with polar center $\left[\frac{s_{12}}{2}\csc\theta, \frac{\pi}{2}-\theta\right]$ and radius $\frac{s_{12}}{2}\csc\theta$. It has singularities where r_1 or r_2 are equal to zero. Thus we obtain the figure eight like shape pictured in Figure 1(b). Rotating the circular arcs in 3-space about the axis defined by the line segment joining L_1 and L_2, we obtain the figure pictured in Figure 1(a).

Figure 1(b) shows how varying the sensor location along the circular arc is equivalent to varying the absolute ranges to the two landmarks. If we can bound the ranges to the landmarks, then we can localize ourselves along the circular arc accordingly. This is logically equivalent to establishing a local coordinate frame between the landmarks, and bounding our location relative to that frame. If we can register the landmarks with a priori map data, then we can know and use the distance between the landmarks, but this is not necessary.

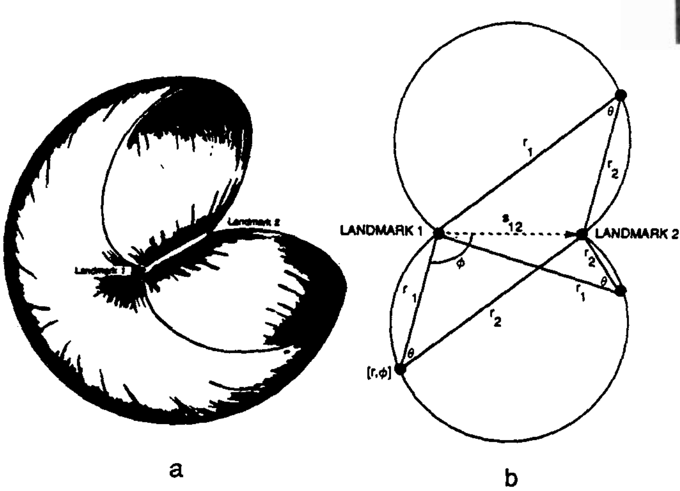


Figure 1: Constant Angle Toroid

Error in angular measurement of the θ_{ij} corresponds to different choices of concentric circular arcs each of which contains L_i, L_j and the sensor focal point. If we union these arcs together, we obtain the localization, Loc_{ij}, of the sensor relative to the two landmarks. Because our localization must be true relative to all observed landmark pairs simultaneously, the intersection of the Loc_{ij} over all $i > j$, give our best localization.

Several simulations of landmark acquisition scenarios and extracted viewframes, orientation regions, viewframe localizations and viewpaths from them have been implemented. Figure 2(a-e) shows the simulated landmark data over which viewframe localizations have been extracted. Figure 2(a) shows an image take at the Martin Marietta Autonomous Land Vehicle development site. Figures 2(b)-(e) show various displays of the 30 meter U.S. Army Engineer Topographic Laboratories terrain grid data gathered over the area in the image.

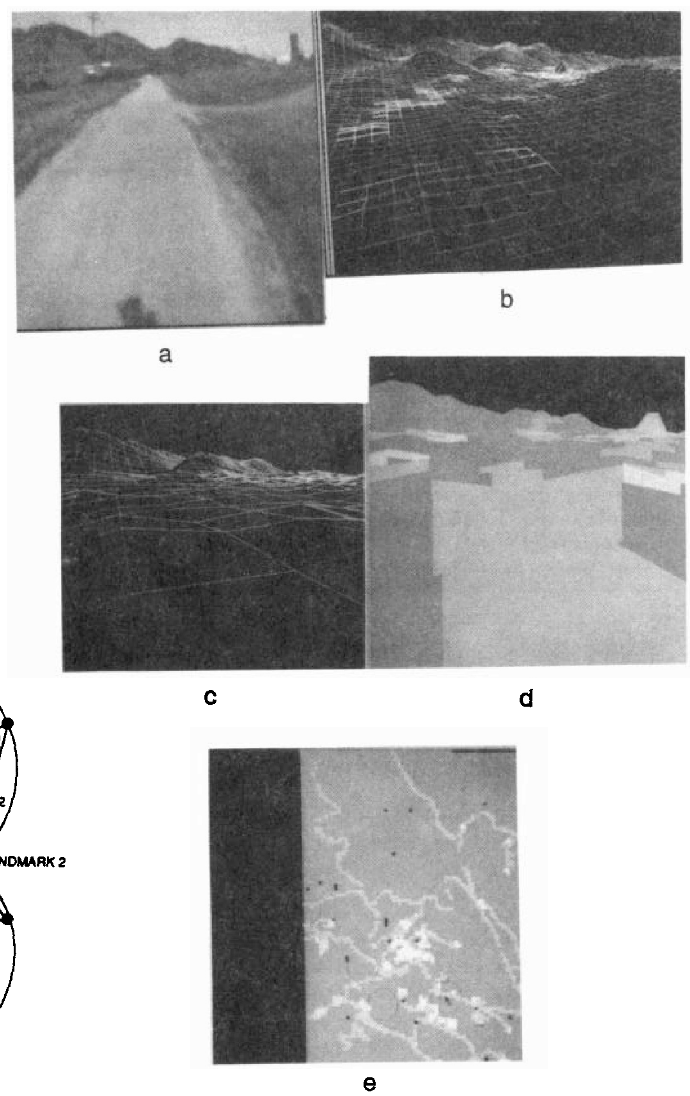


Figure 2: Landmark Data

Figures 2(b) and (e) show two different perspectives on a wire-frame view of the elevation data. The displays include paint for terrain type overlay, and a building present on the site. Figure 2(c) is an orthographic representation of the same grid data, while Figure 2(e) is a painted perspective display. Here the coarse quantization of the road area in the foreground is evident. The circles on Figure 2(c) indicate three manually selected landmark points representing two peaks and the building. The x-ed circle on the right is the location of the simulated sensor.

Figures 3(a-d) show perspective and orthographic views of the viewframe localizations obtained relative to pairs of landmarks. Range intervals of 50% the true range were used. Because the landmarks are approximately 50 pixels from the sensor, this corresponds to range intervals 800 meters long for landmarks 1600 meters away. Angular errors of .1 radian were used. In a 45 degree field of view for an image 512 pixels wide, this is approximately a 65 pixel error. Both errors are far greater than we expect in practice. The intersection of the landmark pair localizations, resulting in the viewframe localization, is shown in Figures 3(c) and (d).

If we drop the range information in viewframes, we are left with purely topological data. If we draw a line between two (point) landmarks, and project that line onto the surface of the ground, then this line divides the earth into two distinct regions. If the landmarks are visible, we can observe which side of this line we are on. The "virtual boundary" created by associating two observable landmarks together thus divides space over the region in which both landmarks are visible. We call these landmark-pair-boundaries (LPB's),

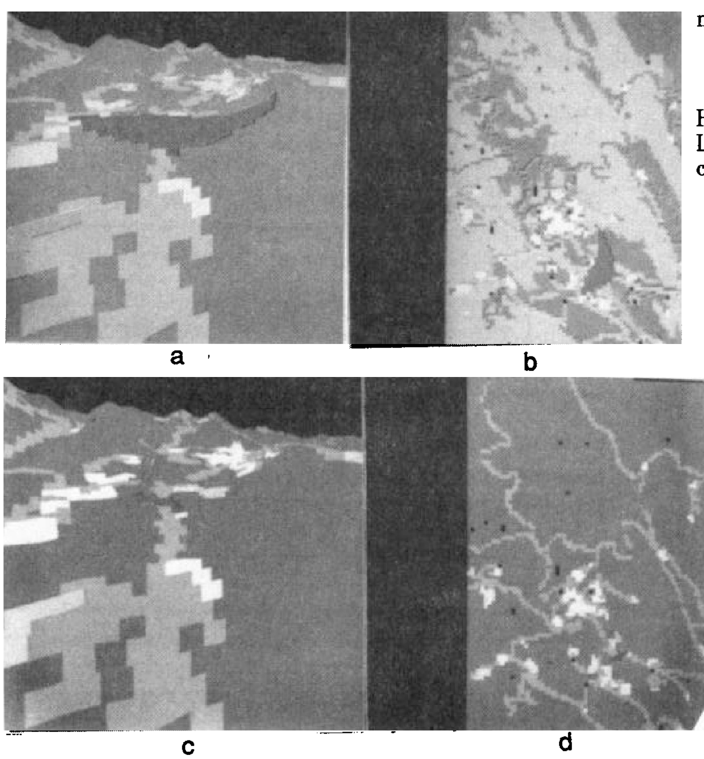


Figure 3: Viewframe Localization

and denote the LPB constructed from the landmarks L_1 and L_2 by LPB(L_1,L_2).

We define:

orientation-of-LPB(L_1,L_2) =

$$\mathrm{sign}(\pi - \Theta_{12}) = \begin{matrix} +1 & \text{if} & \Theta_{12} < \pi \\ 0 & \text{if} & \Theta_{12} = \pi \\ -1 & \text{if} & \Theta_{12} > \pi \end{matrix}$$

where Θ_{12} is the relative azimuth angle between L_1 and L_2 measured in an arbitrary sensor-centered coordinate system. It is a straightforward to show that this definition of LPB orientation does not depend on the choice of sensor-centered coordinate system. LPB's give rise to a topological division of the ground surface into observable regions of localization, called orientation regions. Figure 4(a-b) shows two possible views of the LPB's implicit in a viewframe. Although ranges to the landmarks vary between Figures 4(a) and (b), the topological information of the sensor location and adjacency of LPBs are preserved. If we regard the boundaries of regions as fattened wireframes, then the orientation regions may be thought of as (topological) holes in the surface of the earth. The shape formed by cutting the orientation regions out of the surface of the (hollow) earth is topologically equivalent to a (two) sphere with finitely many holes cut out of it. The number of holes is the number of orientation regions. It can be shown (see [Massey, 1967]) that two shapes induced by orientation regions are topologically equivalent if and only if they have the same number of orientation regions. The total number of regions created by N LPB's is equal to:

number-orientation-regions =

N + (number-of-crossings-with-multiplicites) + 1

Here the multiplicity of a crossing is defined as 1, if two LPB's cross, and, in general, as (number-of-LPBs-crossing - 1) for two or more LPB's.

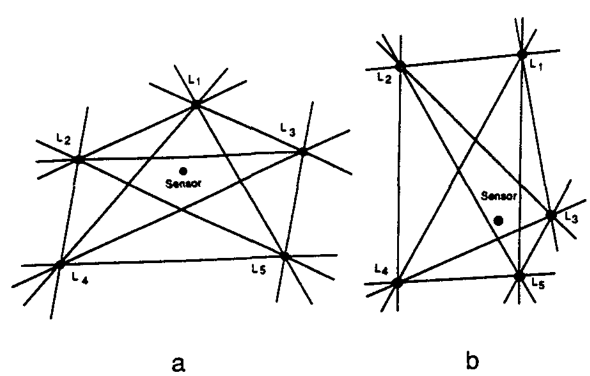
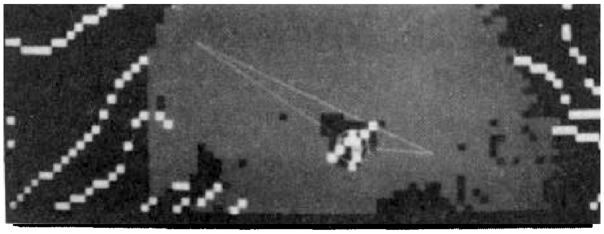
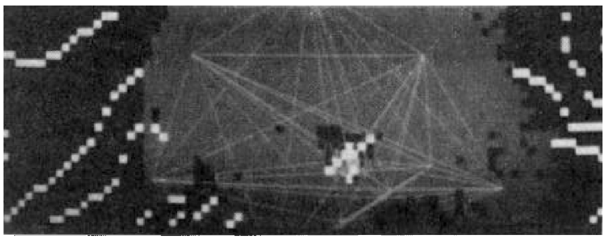


Figure 4: Viewframe Orientation Regions

Figure 5(a) shows the LPB's formed by three landmarks. They are overlaid on a blotch indicating the space in the elevation grid visible from the current location of the simulated sensor. Figure 5(b) shows the LPB's formed by seven landmarks in the same region. Seven landmarks give rise to 162 orientation regions.



a



b

Figure 5: LPB's Formed by Three and Seven Landmarks

3. INFERENCE FOR PATH PLANNING AND FOLLOWING

A heading is defined to consist of: a type; a sensor-centered coordinate system vector; a destination-goal; a direction-function; and a termination-criteria. The type of a heading specifies the coordinate system that the direction conditions are computed in. A metric-heading-type corresponds to an absolute coordinate system, induced by a correspondence of sensor position to a priori map data, from an inertial sensor, geo-satellite location, dead-reckoning from a known initial position, etc. A viewframe-heading-type refers to headings computed between viewframes that share common visible landmarks. This corresponds to reasoning within a local landmark coordinate system. An orientation-heading-type is a coordinate-free heading based on observed relationships with LPB's. Heading types, destination-goals, direction-functions and termination criteria are summarized in Table 1. In the following we focus on qualitative path inference using orientation regions.

Orientation-headings are conjunctions of specifiers for crossing LPB's. These crossings correspond to the visually observable events of L_1 occluding L_2 (or having identical azimuth angle in the sensor centered coordinate system), L_1 and L_2 being separated by 180 degrees, or L_2 occluding L_1. We denote these three possibilities by $l\,[L_1\ L_2]$, $b\,[L_1\ L_2]$ and $r\,[L_1\ L_2]$. We use $a\,[L_1\ L_1]$ to mean "head toward landmark L_1".

Quantizing the LPB into "left-between-right" creates a quantity space in the sense of [Hayes, 1979]. LPB crossings are quantities in the sense of [Forbus, 1984], if we represent them as distance of the robot from the LPB. The derivative part of the Forbus-quantity is then the rate at which the robot is approaching the LPB. The landmarks themselves give rise to partitions of great circles of the earth (modeled as a perfect sphere). These partitions, in turn, give rise to quantities along the robot's path in the sense of [Kuipers, 1986]. That is, the crossing of a LPB creates a

Table 1: Heading Specifications

HEADING TYPE / FIELD	METRIC	VIEWFRAME	ORIENTATION
destination-goal	polygon expressed in absolute coordinate system	viewframe	orientation region
direction function	distance of current estimated point location to destination polygon	if sensor-centered vector-range set then, difference of current location to vector destination relative to point where heading set or maintain visibility of goal landmarks and hill-climb on relative angles	reduce LPB set to unique LPB and track until orientation reversal
termination criteria	estimated error in current location exceeds threshold or goal achieved	if range set then estimated error in relative location exceeds threshold or lose sight of landmarks in goal viewframe or hill climbing to relative angles fails or goal achieved	reduced LPB set is null or lose sight of goal LPB landmarks or must reverse a goal LPB to continue or goal achieved

"landmark value" in the quantity space of the robot's location. (Caution: The term "landmark" is being overloaded.)

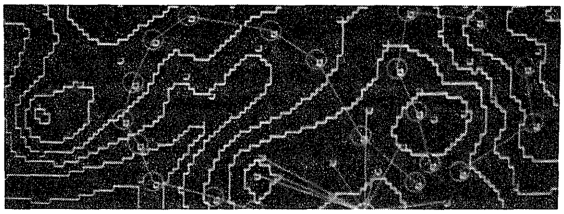
A top level orientation-heading consists of a goal to execute the crossings to get on the correct sides of the LPB's corresponding to visible landmarks that are also listed in the destination-goal. A production system can be used to reduce a conjunction of LPB crossings to a single LPB crossing condition; if relevant, the production system deduces that the specified conjunction of headings is impossible. The production system theory is based on a binary operation between pairs of orientation-headings. Because the operation is associative and commutative, the production system can be executed on a conjunction of orientation-headings, two at a time, in any order. The six possibilites for the productions for combination of headings for two landmark pairs $[A\ \ B]$ and $[C\ \ D]$ are listed in Figure 6.

		VIEWFRAME ORDER OF LANDMARKS					
[A,B] cross specifier	[C,D] cross specifier	ABCD	ACBD	ACDB	CABD	CDAB	CADB
r	r	r [C D]	r [B D]	r [D B]	r [B D]	r [A B]	r [D B]
r	l	b [B C]	--	--	--	--	--
r	b	b [C D]	b [B D]	--	b [B D]	--	--
r	a	a [C C]	--	--	--	--	--
l	r	--	--	--	--	b [D A]	--
l	l	l [A B]	l [A C]	l [A C]	l [C A]	l [C D]	l [C A]
l	b	--	--	--	b [C A]	b [C D]	b [C A]
l	a	--	--	--	--	a [D D]	--
b	r	--	--	b [D B]	--	b [A B]	b [D B]
b	l	b [A B]	b [A C]	b [A C]	--	--	--
b	b	--	b [C B]	b [C D]	b [A B]	--	b [A D]
b	a	--	--	a [C C]	--	--	--
a	r	--	--	--	--	a [A A]	--
a	l	a [A A]	--	--	--	--	--
a	b	--	--	--	a [A A]	--	--
a	a	--	--	--	--	--	--

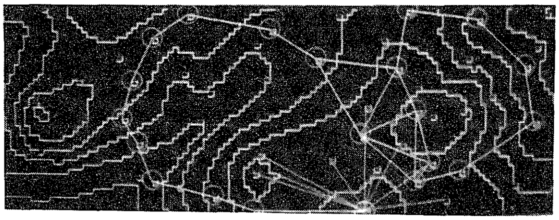
Figure 6: Orientation Heading Binary Productions

A natural environmental representation based on viewframes orientation regions and LPB crossings, recorded while following a path, is given by a list of the ordered sequence of viewframes collected on the path, and another list of the set of landmarks observed on the path. For efficiency, the landmark list can be formed as a database that can be accessed based on spatial and/or visual proximity.

When a new viewpath is added to the database of perceptual knowledge, additional links between viewpaths are constructed based on landmarks seen on both paths. Using coarse range estimates to common landmarks, viewframe headings are computed between viewframes on different viewpaths. This structure is pictured in Figure 7. Figure 7(a) shows two viewpaths, while Figure 7(b) shows the paths augmented with the additional links. It is this augmented visual memory over which path plans are generated prior to path execution.



a



b

Figure 7: Visual Memory Linking

The top level loop for landmark-based path planning and following is to: determine a destination-goal, compute and select a current heading, and execute the heading while building up an environmental representation. The destination-goals implement a recursive goal-decomposition approach to perceptual path planning. The concept underlying the path planning/following strategies encoded in these rules is to mix the following approaches as knowledge is available or can be inferred:

- find landmarks in common between viewframes between point of origin and viewframe-destination and compute vector (i.e., direction and approximate range) headings between viewframes

- locate and get on the correct side of LPB's specified in an orientation-destination, or

- associate visible and goal landmarks with map data and compute a metric heading between current location and goal

Each of these strategies provably reaches its goal, up to the perceptual re-acquisition of landmarks and the traversability of intervening terrain.

```
if viewframe goal landmarks visible
  --> compute viewframe-heading
      if at least one LPB has an incorrect orientation
         relative to our viewframe-destination-goal
      then follow heading for approximate distance
         estimated by the viewframe-heading
      else maintain heading by control-feedback
         path following on relative angles between
         landmarks
      build a new viewpath to destination goal, using
         the existing landmark list where possible

if viewframe goal landmarks not visible and
viewpaths exists
  --> make a viewframe of the currently visible
         region
      chain back through viewpaths until common
         landmarks are located
      chain forward through viewframes setting up
         intermediate destination-goals
      recursively execute viewframe headings to reach
         the destination goals corresponding to
         visible landmarks
if viewframe goal landmarks not visible and no
  viewpath exists
  --> set goal to find a metric heading
```

We have implemented these rules with routines that use A* to plan an initial route to a destination based on data in visual memory. This route is executed using vision, with re-planning based on the currently perceived viewframe at each step. Figure 8(a) shows the plan over visual memory to move between two points. The executed route is shown in Figure 8(b). Notice how much smoother it is. Figure 8(c) shows an original plan, while 8(d) shows a dramatic re-plan based on observing a "short-cut" at runtime.

4. SUMMARY AND FUTURE WORK

A rigorous theory of qualitative, landmark-based path planning and following for a mobile robot has been developed. It is based upon a theory of representation of spatial relationships between visual events that smoothly integrates topological, interval-based, and metric information. The rule-based inference processes opportunistically plan and execute routes using visual memory and whatever data is currently available from visual recognition, range estimates and a priori map or other metric data.

ACKNOWLEDGEMENTS

This document was prepared by Advanced Decision Systems (ADS) of Mountain View, California, under U.S. Government contract number DACA76-85-C-0005 for the U.S. Army Engineer Topographic Laboratories (ETL), Fort Belvoir, Virginia, and the Defense Advanced Research Projects Agency (DARPA), Arlington, Virginia. The authors wish to thank Angela Erickson for providing administration, coordination, and document preparation support.

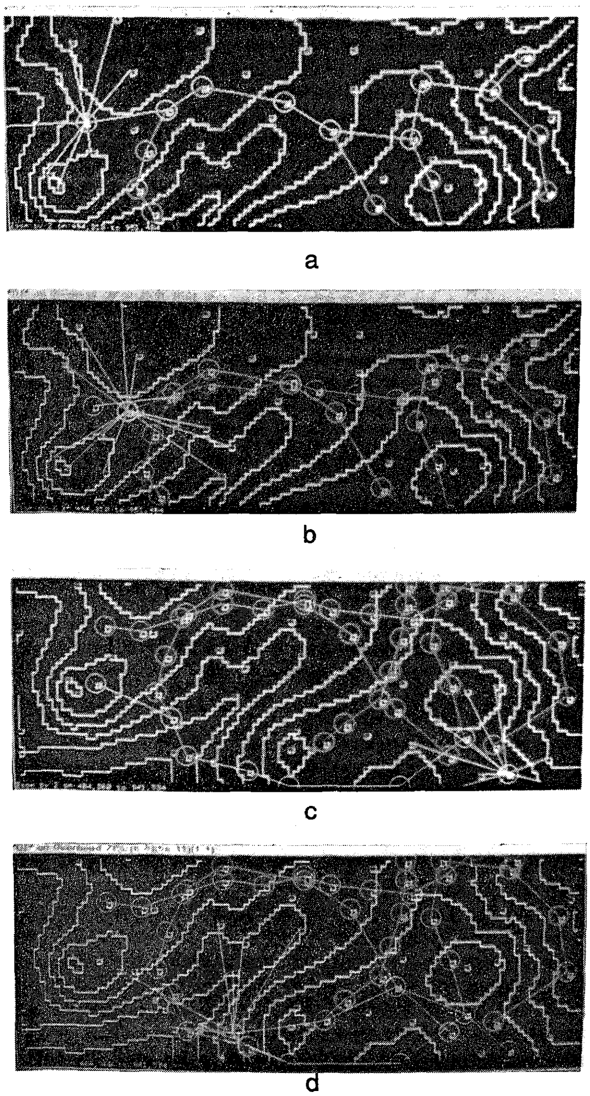


Figure 8: Path Planning and Following Results

REFERENCES

[Bajcsy et al., 1986] - R. Bajcsy, E. Krotkov, and M. Mintz, "Models of Errors and Mistakes in Machine Perception", University of Pennsylvania, Computer and Info. Science Technical Report, MS-CIS-86-26, GRASP LAB 64, 1986.

[Davis, 1986] - E. Davis, "Representing and Acquiring Geographic Knowledge", Courant Institute of Mathematical Sciences, New York University, Morgan Kaufmann Publishers, Inc., 1986.

[Forbus, 1984] - K.D. Forbus, "Qualitative Process Theory", Artificial Intelligence, Vol. 24, December, 1984.

[Hayes, 1979] - P.J. Hayes, "The naive physics manifesto", in Expert Systems in the Micro-Electronic Age, D. Michie (Ed.), Edinburgh University Press, Edinburgh, Scotland, 1979.

[Hebert and Kanade, 1985] - M. Hebert and T. Kanade, "First Results on Outdoor Scene Analysis Using Range Data", Proceedings Image Understanding Workshop, Miami Beach, Florida, December 9-10, 1985, pp. 224-231.

[Kuipers, 1977] - B. Kuipers, "Representing Knowledge of Large-Scale Space", Massachusetts Institute of Technology, Artificial Intelligence Laboratory Technical Report, AI-TR-418, July 1977.

[Kuipers, 1986] - B. Kuipers, "Qualitative Simulation", Artificial Intelligence, Vol. 29 No. 3., September, 1986.

[Lucas and Kanade, 1984] - B. Lucas and T. Kanade, "Optical Navigation by the Method of Differences", Proceedings Image Understanding Workshop, New Orleans, Louisiana, October 3-4, 1984, pp. 272-281.

[Matthies and Shafer, 1986] - L. Matthies and S. Shafer, "Error Modelling in Stereo Navigation", Carnegie-Mellon University, Computer Science Department, Technical Report, CMU-CS-86-140, 1986.

[McDermott and Davis, 1984] - D. McDermott and E. Davis, "Planning Routes through Uncertain Territory", Artificial Intelligence - An International Journal, Vol. 22, No. 2, March 1984, pp. 107-156

[Massey, 1967] - R. Massey, "Introduction to Algebraic Topology", Addison-Wesley, 1967.

[Pentland, 1985] - A. Pentland, "A New Sense for Depth of Field", Proceedings of the Ninth International Joint Conference on Artificial Intelligence, IJCAI-85, Los Angeles, California, August 18-23, 1985, pp. 988-994.

[Schone, 1984] - H. Schone, "Spatial Orientation - The Spatial Control of Behavior in Animals and Man", Princeton Series in Neurobiology and Behavior, R. Capranica, P. Marler, and N. Adler (Eds.), 1984.

[Smith and Cheeseman, 1985] - R. Smith and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty", SRI International Robotics Laboratory Technical Paper, Grant ECS-8200615, September 1985.