Landmark Recognition for Autonomous Land Vehicle Navigation

Hatem Nasr, Bir Bhanu and Stephanie Schaffer

Honeywell Systems and Research Center
3660 Technology Drive, Minneapolis, MN 55418

ABSTRACT

In the Autonomous Land Vehicle (ALV) application scenario, a significant amount of positional error is accumulated in the land navigation system after traversing long distances. Landmark recognition can be used to update the land navigation system by recognizing the observed objects in the scene and associating them with the specific landmarks in the geographic map knowledge-base. In this paper we present a novel landmark recognition technique based on a perception-reasoning-action and expectation paradigm of an intelligent agent. It uses extensive map and domain dependent knowledge in a model-based approach. It performs spatial reasoning by using N-ary relations in combination with negative and positive evidences. Since it can predicts the appearance and disappearance of objects, it reduces the computational complexity and uncertainty in labeling objects. It provides a flexible and modular computational framework for abstracting image information and modeling objects in heterogeneous representations. We present examples using real ALV images.

I. INTRODUCTION

In order to accomplish missions such as surveillance, search and rescue and munitions deployment, an Autonomous Land Vehicle (ALV) has to travel long distances. This results in a significant amount of positional error in the land navigation system. Landmark recognition is used to update the land navigation system, thus guiding the ALV to remain on its proper course. Landmarks of interest include telephone poles, storage tanks, buildings, houses, gates, etc.

Model-based vision has been a popular paradigm in computer vision since it reduces the problem complexity and no learning is involved. Binford [6] has given a summary of model-based vision work. He has described several systems including the work of Brooks [7] on ACRONYM, Riseman and Hanson's [12] work on VISIONS, and Nagao and Matsuyama's [14] work on the analysis of complex aerial photographs. McKeown et al [13] have used map and domain specific knowledge in the SPAM rule-based systems for the interpretation of airport scenes in aerial images. Hwang [10] has also used domain knowledge to guide interpretation of suburban house scenes in aerial imagery. He has used test-hypothesize-act sequence to generate large number of hypotheses which are then integrated into a consistent interpretation. Bhanu [1-4] has used several modeling and relaxation matching techniques for the recognition of 2-D and 3-D nonoccluded and occluded objects. As compared to all the previous related work, as mentioned in the above, the paradigm of an intelligent agent (like the ALV) which we have used here is based on the perception-reasoning-action and expectation cycle. Thus we have an expectation-driven, knowledge-based landmark recognition system called PREACTE (Perception-REasoning-ACTion and Expectation), that utilizes a priori, map and perceptual knowledge, spatial reasoning and knowledge aggregation methods. In contrast to the work of Davis [9], explicit knowledge about the map and landmarks is assumed to be given and it is represented in a relational network. It is used to generate an Expected Site Model (ESM) given the ALV location and its velocity. Landmarks at a particular map site have their 3-D models stored in heterogeneous representations. The vision system generates a 2-D and partial 3-D scene model from the observed scene. The ESM hypothesis is verified by matching it to the image model. The matching problem is solved by using object grouping and spatial reasoning. Positive as well as negative evidences are used to verify the existence of each landmark in the scene. The system also provides feedback control to the low-level processes to permit adaptation of the feature detection algorithms parameters to changing illumination and environmental conditions.

In the following, we present the details of the PREACTE system and examples of landmark recognition using real ALV imagery. It is worth mentioning that PREACTE is a component subsystem of a much larger system. Other parts of the system will be referred to but not described.

II. CONCEPTUAL APPROACH

The task of visual landmark recognition in the autonomous vehicle scenario can be categorized as (a) uninformed and (b) informed. In the uninformed case, given a map representation, the vision system attempts to attach specific landmark labels to segmented image regions of an arbitrary observed scene and infers the location of the vehicle in the map (world). On the other hand, in the informed case, while the task is the same as earlier, there is a priori knowledge (with a certain level of certainty) of the past location of vehicle in the map and its velocity. It is the informed case that is of interest to the discussion of this paper. There are a number of assumptions made in this landmark recognition approach. They include: 1) a forward looking fixed camera model is given, 2) traversal by the vehicle is allowed only on defined routes, and 3) minor range variations from a given site does not lead to major changes in the objects' appearance in the image and their spatial distributions.

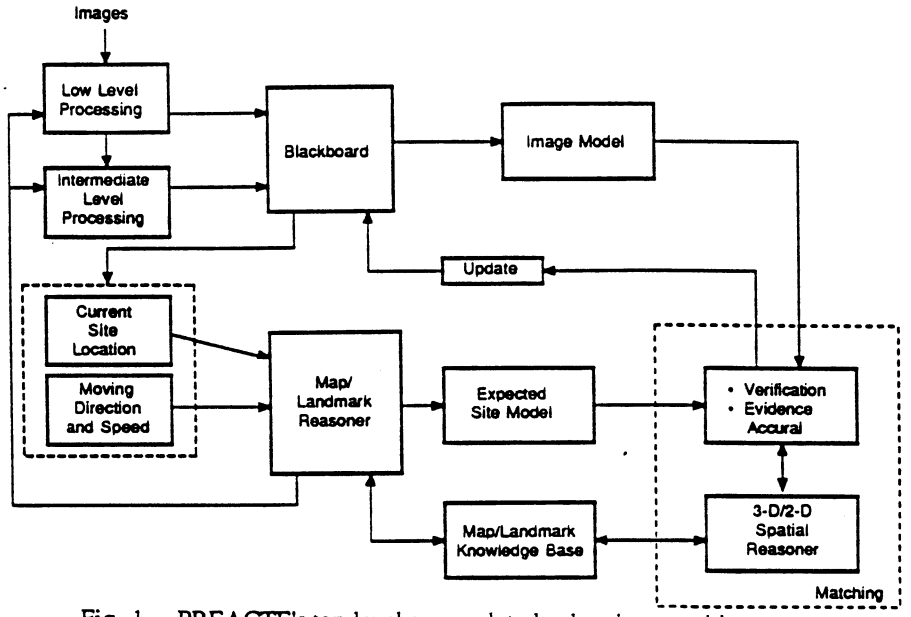

Fig. 1. PREACTE's top-level approach to landmark recognition.

Fig. 1 illustrates the overall approach for PREACTE's landmark recognition task. It is a top-down expectation-driven approach, whereby an Expected Site Model (ESM) of the map is generated based on domain-dependent knowledge of the current (or projected) location of the vehicle in the map and vehicle's velocity. The ESM contains models of the expected map site and its landmarks. This expectation provides the hypotheses to be verified by the content of an image to be acquired after a computed time t, given the velocity of the vehicle and the distance between the current site and the predicted one. While this approach may seem similar to other hypothesis-verification concepts, it is not only unique by its added expectations but also in its extensive and explicit domain specific knowledge which contributes to enhanced performance. Site models introduce spatial constraints on the locations and distributions of landmarks, by using a "road" model as a reference. Spatial constraints greatly reduce the search space while attempting to find a correspondence between the image regions and a model. This mapping is usually many-to-one in complex outdoor scenes, because of imperfect segmentation.

In the segmented image each region-based feature such as size, texture, color, etc. provides an independent evidence for the existence of an expected landmark. Evidence accrual is accomplished by an extension of a heuristic Bayesian formula [8], which will be discussed in Section II.3. The heuristic formula is used to compute the certainty about a map site location based on the certainty of the previous site and the evidences of each landmark existence at the current site. Similar formulation was suggested by Lowe [11] for evidential reasoning for visual recognition.

A. Map/Landmark Knowledge-Base

Extensive map knowledge and landmarks models are fundamental to the recognition task. Our map representation relies heavily on declarative and explicit knowledge instead of procedural methods on relational databases [13]. The map knowledge is represented in a hierarchical relational network, as illustrated in Fig. 2. The entire map is divided into 25 sectors (5 horizontally and 5 vertically). Each sector contains four quadrants which in turn contain a number of surveyed sites (Fig. 3). All map primitives are represented in a schema structure. The map dimensions are characterized by their cartographic coordinates. Schema representation provides an object-oriented computational environment which supports the inheritance of different map primitives properties and allows modular and flexible means for searching and updating the map knowledge base. The map sites between which the vehicle traverses have been surveyed and characterized by site numbers. An aerial photograph with numbered sites is shown in Fig. 3. Knowledge acquired about these sites includes: approximate < latitude, longitude, elevation >, distance between sites, terrain descriptions, landmarks labels contained in a site, etc. Such site information is represented in a SITE schema, with corresponding slots, as illustrated in Fig. 2. Slots names include: HAS_LANDMARKS, NEXT_SITE, LOCATION, SPATIAL_MODEL, etc. A critical slot is NEXT_SITE which has an "active" value. By active, it is meant that it is dependent on a variable (demon) which is the vehicle direction (North, South, etc.). For different traversal directions from the current site, the names of the neighboring sites are explicitly declared in the NEXT-SITE slot, as shown in Fig. 4. The SPATIAL_MODEL defines the "expectation zone" of the landmarks (in the image) with respect to the road and with respect to each others. It also specifies the minimum and maximum distance of each landmark from the road borderline. Each landmark is represented as a schema or a collection of schemas. Each landmark is represented as an instance of a landmark-class which, in turn, is an instance of an object-class. For example, T-POLE-17 is an instance of POLE, which is an instance of MAN_MADE_OBJECTS. Instances in this case inherit some properties and declare others.

This declarative and hierarchical representation of the knowledge allows not only a natural conceptual mapping, but also a flexible means for pattern matching, data access, tracing of the reasoning process and maintenance of the knowledge base. The slots and their values in a LANDMARK schema correspond to the landmark's attributes such as color, texture, shape, geometric model, etc. The landmark attributes are characterized

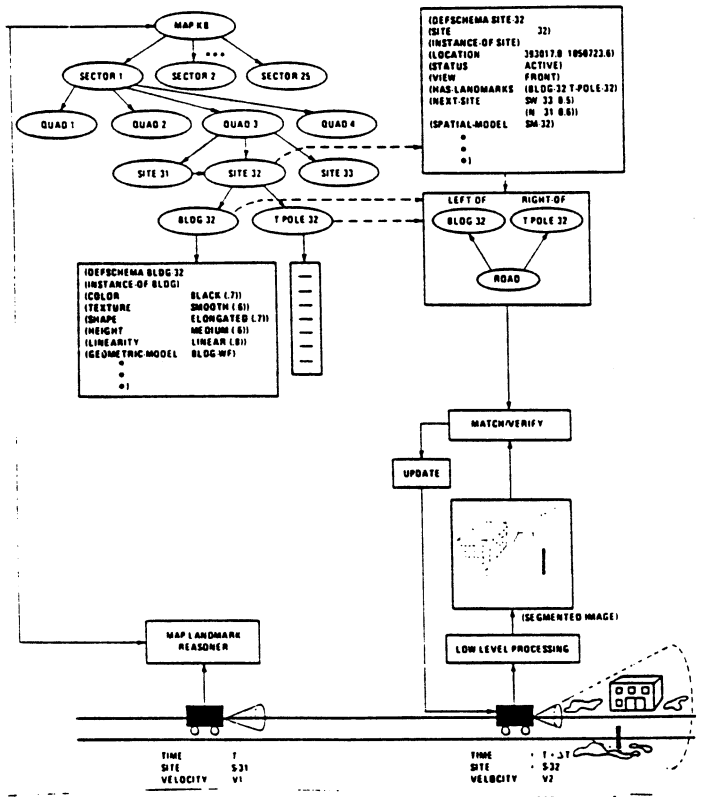

Fig. 2. Map knowledge representation and graphic illustration of the approach based on the perception-reasoning-action and expectation paradigm.

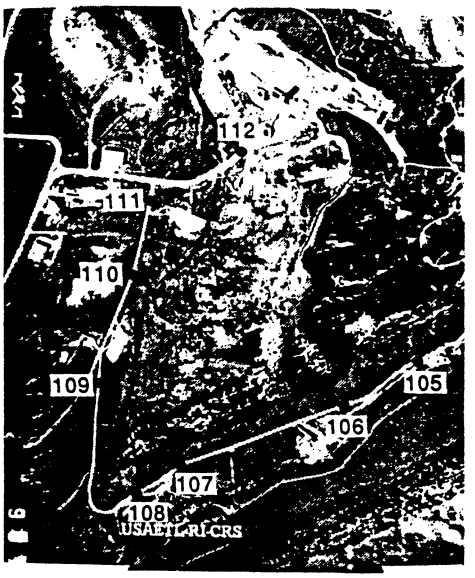

Fig. 3. Aerial photograph of the map.

symbolically, such as color is "black", texture is" smooth", and shape is "elongated". Each attribute's value is assigned a likelihood that characterizes its discriminant strength. For example, the fact that poles are elongated, place a high likelihood value (0.8) on having an elongated shape. The knowledge acquisition for modeling each landmark in the knowledge base is performed by abstracting and characterizing map data through actual observations and measurements, and observations of images taken at individual sites. The groundtruth values of each landmark attribute are obtained from a combination of actual metrics, and approximations of features extracted from hand segmented images. Three dimensional geometric models are represented in the geometric-model slot of a LANDMARK schema. Different modeling techniques are available for different landmarks. For example, buildings are represented as wire-frames, while poles are represented as generalized cylinders. Thus models are allowed to have heterogeneous representations. Image description is obtained by projecting the 3-D model on a 2-D plane using perspective transformations. This hybrid representational framework for object modeling provides a unique ability to combine different types of object descriptions (semantic, geometric and relational). This in turn allows the system to perform more robustly and efficiently, and recover from a single bad representation.

B. Prediction

Given the a priori knowledge of the vehicle's current location in the map space and its velocity, it is possible to predict the upcoming site that will be traversed through the explicit representation of the map knowledge and the proper control procedures. The Map/Landmark Reasoner (MLR) provides such control, by invoking the active values in the NEXT_SITE slot of the current SITE schema, as described earlier. The ESM is a "provision" by the MLR to make the expected site and its corresponding landmark schemas as an "active" hypothesis to be verified. In parallel to predicting the next site, the distance between the current and the expected site along with the vehicle velocity are used to predict the (arrival) time at which the sequence of images should be processed to verify the hypothesized ESM. Evidence accrual and knowledge aggregation is dynamically performed between sites to confirm arrival time at the predicted site [5]. The ESM of SITE-110 shown in Fig. 7 is as follows:

```
(DEFSCHEMA  SITE-110
    (SITE                110)
    (LOCATION            (392961.7        1050742.9))
    (INSTANCE - OF       SITE)
    (STATUS              ACTIVE)
    (VIEW                FRONT)
    (HAS-LANDMARKS       (T-POLE-110 G-TANK-10 BLDG-110))
    (NEXT-SITE           (E 109  0.255)  (W 111  0.153))
    (SPATIAL-MODEL       SM-110)
    (TERRAIN-TYPE        NIL))

(SETQ     SM-110      '( (T-POLE-110    G-TANK-110     BLDG-110)
                         (T-POLE-110                   (LEFT-OF   ROAD)
                                                       (MIN-L-DIST-ROAD 160)
                                                       (MAX-L-DIST-ROAD 200)

                         (G-TANK-110                   (RIGHT-OF ROAD)
                                                       (MIN-R-DIST      200)
                                                       (MAX-R-DIST      250))

                         (BLDG-110                     (ABOVE      G-TANK-110)
                                                       (RIGHT-OF        ROAD )
                                                       (MIN-R-DIST      230)
                                                       (MAX-R-DIST      280) ) ) )
```

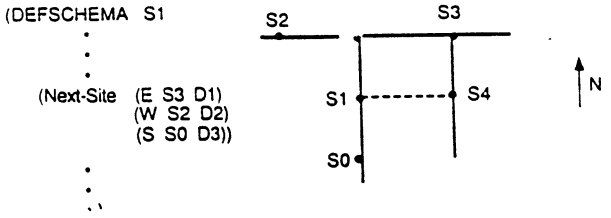

Fig. 4. Next-Site slot representation in a SITE schema provides the expected site and distance to it based on the vehicle direction.

Predictions are also used by the low-level image processing . A priori knowledge of the objects' attributes which will appear in the image and their relative locations guide the segmentation. A rule-based system is invoked to interpret the corresponding information in the ESM, which results in properly adapting the segmentation parameters based on the landmarks' distinguishing attributes, such as color , location, texture,etc.

C. Image Modeling

An image model for the task of landmark recognition is a collection of regions-of-interest extracted by a region-based segmentation method. A region-of-interest for this task is constrained by an upper and lower bound on its size. This means that after performing some region splitting and merging, most of the very small and the very large regions are merged or discarded from the image. In addition, regions-of-interest do not include any regions that represent moving objects in the scene as determined by the motion analysis module. A number of image features are extracted for each region, such as color, length, size, perimeter, texture, Minimum Bounding Rectangle (MBR), etc., as well as some derived features such as elongation, linearity, compactness, etc. All image information is available in the blackboard (Fig. 1), which is a superset model of all the results collected from different image understanding modules. The landmark recognition system operates as a "knowledge source". There are other knowledge sources with other tasks such as object recognition (other than landmarks) and motion analysis. The blackboard plays the role of a central knowledge structure among these different knowledge sources. During the process of extracting regions-of-interest for landmark recognition there is a risk of ignoring regions in the image that are part of the actual landmarks. These regions could have been split to very small regions or merged with very large ones, which is a natural outcome of the inherently weak segmentation methods. Symbolic feature extraction is performed on the region-based features. So, instead of having area = 1500 (pixels) and intensity = 52, we could have area = large and intensity = low. The symbolic characterization of the features using "relative" image information provides a better abstraction of the image and a framework for knowledge-based reasoning. On one hand, this has the advantage of making the feature space smaller, therefore easier to manipulate. On the other hand, it makes it insensitive to feature variations in the image.

Each set of region features is represented in a schema structure instead of a feature vector as in pattern recognition. This schema representation of regions does not have any conceptual justifications, however it provides a compatible data structure with the landmark models in the knowledge-base. Most of the region features have representative attributes in the landmarks models. This allows symbolic pattern matching to be performed easily by the high-level vision knowledge sources. Beyond that, it makes the reasoning process more traceable.

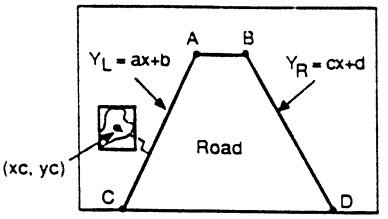

Fig. 5. Road model representation.

A critical region in the image is the road region, which is used as a reference in the image model. Spatial constraints are applied on the regions-of-interest to find which regions in the image fall to the left and to the right of the road. The road is easily segmented out in similar imaging scenarios, using current state-of-the-art road segmentation techniques (currently used in the ALV). This is assuming that it is a "structured" road (i.e., asphalt, concrete, etc.) that provides good contrast (not dirt roads). The road is represented in the model by its vertices and the approximate straight lines of the left and right borders, as shown in Fig. 5. For each region, we determine the position of its centroid and compute the shortest distance from the region to the road border line. This distance is compared to the constraint imposed on each landmark by the site spatial model. Thus we obtain the following top-level structure for the image model:

(<IM-#> <frame-#> <road-region-tag>
<number-of-regions-of-interest>
((<road-vertices>) (<left-border>) (<right-border>))
((<left-regions-list>) (<right-regions-list>))))

D. Hypothesis Verification

Given the expected site model (ESM) and the current image model, the objectives of the matching and verification process are two fold: (a) to label the regions in the image corresponding to the expected landmarks, and (b) to determine the certainty level of the predicted map site location. The process by which the first objective is accomplished is as follows: 1. find the set of regions {R} in the image model (IM) which satisfy the spatial constraints SC_i imposed by landmark l_i in the ESM SPATIAL_MODEL. This constraint application yields to more than one corresponding region r_j. 2. Compute the evidence $E(l_i)$ that each r_j in {R} yields, using the FIND_EVIDENCE algorithm. 3. The r_j that results in $E(l_i)_{max}$ (provided it is a positive evidence) is considered as best match candidate for l_i (there may be more than one given that their values surpass a certain threshold) . The second objective is achieved by aggregating the individual set of evidences ($E(l_i)_{max}$) and the certainty level about the previous map site location and the potential error introduced by range and the view angle of the camera.

The FIND_EVIDENCE algorithm considers that each landmark l_i in the ESM has a set of attributes $\{A_{i1}, ..., A_{iK}, ..., A_{in}\}$, each with a likelihood LH_{ik}, as described earlier. Each region r_j in {R} has a set of features $\{f_{j1}, ..., f_{jk}, ..., f_{jn}\}$. Note that A_{ik} and f_{ik} correspond to the same thing (in the model and the image), such as color, size, texture, etc. Given these features, we want to compute the evidence that l_i is present in the image.

$$P(l_i/f_{j1}, ..., f_{jk}, ..., f_{jn}) = \frac{P(l_i) * P(f_{j1}, ..., f_{jk}, ..., f_{jn}/l_i)}{P(f_{j1}, ..., f_{jk}, ..., f_{jn})} \quad (1)$$

By making the independence assumption among features in a region and among features occurrence in images, the above equation can be rewritten as:

$$P(l_i/f_{j1}, ..., f_{jk}, ..., f_{jn}) = \frac{P(l_i)*P(f_{j1}/l_i)*...*P(f_{jk}/l_i)*...*P(f_{jn}/l_i)}{P(f_{j1})*...*P(f_{jk})*...*P(f_{jn})}$$

$$= P(l_i) * \prod_{k=1}^{n} \frac{P(f_{jk}/l_i)}{P(f_{jk})} \quad (2)$$

where n is the number of features, $P(l_i)$ is the initial probability of a landmark being found in a given site. For now this is set to 1 for all landmarks. However, $P(l_i)$ is actually a function of the certainty level about the previous map site location, navigational error and other variables. $P(f_{jk})$ is the probability of occurrence of a feature in an image, which is equal to 1/(number of possible feature values). For example, if texture can take either of the four values: coarse, smooth, regular or irregular, then P (texture = smooth) = 1/4. Finally,

$$P(f_{jk}/l_i) = \begin{cases} LH_{ik} & \text{if } f_{jk} = A_{ik} \\ \dfrac{1 - LH_{ik}}{d(f_{jk}, A_{ik})} & \text{if } f_{jk} \neq A_{ik} \end{cases} \quad (2.1)$$

which is best explained through the following example:

Given two regions r_1 and r_2 in the image with different sizes (f_{jk}), SIZE (r_1) = SMALL and SIZE (r_2) = LARGE. Given a model of landmark L, with the expected size to be LARGE (A_{ik}), with a likelihood (LH_{ik}) of 0.7. The SIZE feature can take any of the following ordered values: {SMALL, MEDIUM, LARGE}. If r_2 is being matched to L, (2.1) yields to 0.7, because $f_{jk} = A_{ik}$. On the other hand, if r_1 is being matched to L, then (2.1) yields to (1-0.7)/2. The denominator 2 is used because LARGE is two unit distances (denoted by d(.)) from SMALL.We rewrite (2) as:

$$P(l_i/f_{j1}, ..., f_{jk}, ..., f_{jn}) = P(l_i) * \prod_{k=1}^{n} I(f_{jk}/l_i) \quad (3)$$

where I(./.) is the term within the product sign. The value of $I(f_{jk}/l_i)$ can be greater than 1, because the heuristic nature of the formulation does not reflect a probabilistic set of conditional events, as formulated in Bayes theory. Moreover, $P(l_i/f_{ji}...f_{jn})$ can result in a very large number or a very small positive number.

By taking the logarithm of both sides of (3), introducing W_j as a normalization factor for each feature, and dividing by the number of features (n), we have :

$$Log[P(l_i/f_{j1}, ..., f_{jk}, ..., f_{jn})] = Log[P(l_i)] + \frac{\sum_{k=1}^{n} Log[I(f_{jk}/l_i) * W_j]}{n} \quad (4)$$

where W_k is a normalization factor between 0 and 1.

Here we further simplify (4) and introduce the evidence terms E and e to be the logarithm of P and I*W respectively . So, the evidence formula can be written as follows:

$$E(l_i) = \frac{\sum_{k=1}^{n} e(f_{jk}/l_i)}{n} \quad (5)$$

The values of $E(l_i)$ fall between 0 and 1. If $E(l_i) > 0.6$ it means a "positive" set of evidences. On the other hand, if $E(l_i) < 0.3$ it is interpreted as "negative" evidence. Otherwise, $E(l_i)$ is characterized as "neutral".

An important characteristic of the PREACTE system is that it utilizes negative as well as positive evidences to verify its expectations. There are many types of negative evidences that could be encountered during the hypothesis generation and verification process. The one that is of particular interest to us is when there is a negative evidence about a "single" landmark ($E(l_i)_{max} < 0.3$) in conjunction with positive evidences about the other landmarks (average evidence >0.6) and a reasonable level of certainty about the previous site ($U_{s-1} < 5$) (discussed later). This case is interpreted as caused by one or more of the following: (a) error in the dimension of the expectation zone, (b) bad segmentation results, and (c) change in the expected view angle or range.

In such a case, we perform the following steps: 1. enlarge the expectation zone by a fixed margin, and find the evidences introduced by the new set of regions, as shown in Fig. 6. 2. If step 1 fails to produce an admissible set of evidences, then the expectation zone of the image is resegmented using a new set of parameters that are strictly object dependent.

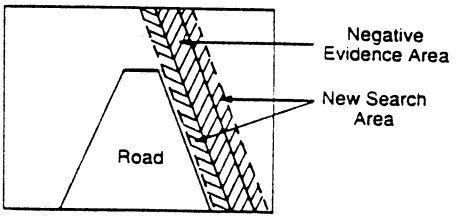

Fig. 6. New search area as a result of negative evidences.

Even though landmark recognition is introduced to assist the autonomous vehicle land navigation system, there is obviously uncertainty attached to the results of the recognition system. We compute the uncertainty U_S at each site location in the following manner:

$$U_s = (U_{s-1} + \alpha) * \prod_{i=1}^{m} \frac{0.5}{E(l_i)_{max}}$$

where U_{S-1} is the uncertainty at the previous site, U_0, the initial uncertainty, is equal to 1, a is the error factor introduced by the navigation system, it is set to a constant of 0.3 (for experimental reasons), and $E(l_i)_{max}$ is the maximum evidence of l_i. If two or more regions return evidences greater than .8 then the average is computed.The value 0.5 is used (as neutral evidence) to stabilize the function, m is the number of landmarks. The multiplicative nature of U_S provides it with the capability of rapidly recovering

its value given a high set of evidences at the current site and a high level of uncertainty at the previous site.

III. RESULTS

We have implemented a prototype system written in Common Lisp and ART (Automated Reasoning Tool) on the Symbolics 3670. The image processing software was implemented in C on the VAX 11/750. The Symbolics hosts all the high-level (symbolic) processing software, including the blackboard. The map and landmarks knowledge-base is implemented as a hierarchical relational network of schemas, the FIND-EVIDENCE algorithm is implemented in Lisp. The Map-Landmark Reasoner is implemented in a rule-based structure. An initial implementation of PREACTE was tested on a video sequence of imagery. Data was collected at 30 frames/second by a camera installed on top of a vehicle and driven on the road connecting the sites shown in Fig. 3.

The system was easily capable of predicting the next site and approximate arrival time at each site. In an experiment where we started at SITE-109 and traveled west at 10 kph, arrival time was predicted to SITE-110, at 245.4 seconds (distance between the two sites is .426 mile). Fig. 7(a) shows the image taken at SITE-110, which contains a pole (T-POLE-110) to the left of the road, a gas tank (G-TANK-110) and a building (BLDG-110) to the right. Initial segmentation of the image is shown in Fig. 7(b). As a result of region splitting and merging and discarding small regions, we obtain the image shown in Fig. 7(c). Rectangular boxes are overlayed over the regions recognized by PREACTE as landmarks, based on the hypotheses generated in the ESM of SITE-110 (shown in Section II.2). The hypothesis verification results are shown in the "match-evidence" column of Table I which contains a listing of the regions yielding the highest evidences and a subset of their features (other features include compactness, intensity variance, etc.), as represented in the image model. The spatial constraint specified by the spatial model (SM-110) yielded to a small number of regions-of-interest for the POLE hypothesis, as a result of the successful post-segmentation effort. More regions-of-interest were considered as candidates for the other landmarks. The road (region 98) in the image is modeled by its approximate left border ($y = -0.8x + 357.0$) and its right border ($y = 1.3x - 126.7$). The T-POLE-110 hypothesis produced two regions with high evidences. Since a threshold of 0.8 was used, both regions 30 and 79 are recognized as T-POLE-110. The lower part of the pole (region 79) is merged with some ground and background regions; nevertheless it still resulted in a higher evidence than the upper part (region 30). Currently effort is underway to implement a region grouping technique based on evidences, proximity, size and other criteria. The lower part of the tank was broken up into six small regions because of the illumination and shape factors. These regions were included in the hypothesis verification and they produced significantly lower evidences.The uncertainty :
$U_{110} = (1+0.3)*((0.5)^3/(0.875+0.62+0.70)) = 0.43$, where 0.875 is the average of 0.92 and 0.83.

IV. CONCLUSIONS

In this paper we have presented concepts and initial results of our perception-reasoning-action and expectation paradigm of our ongoing research for guiding the ALV by recognizing landmarks along the sides of the road. In the future, we will extend to a more general and complex situation where the ALV may be traveling through terrain and it has to determine precisely where it is on the map by using landmark recognition.

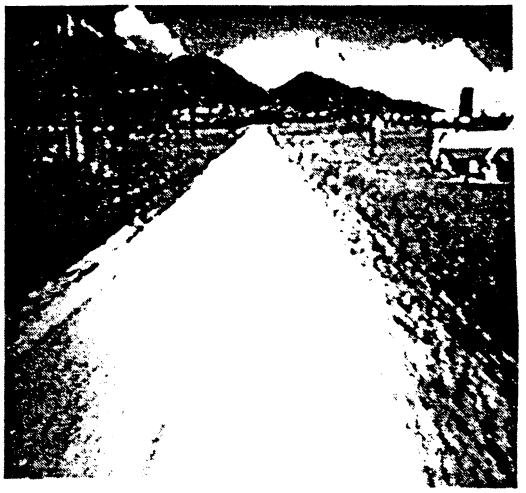

Fig. 7(a) Image obtained at site 110.

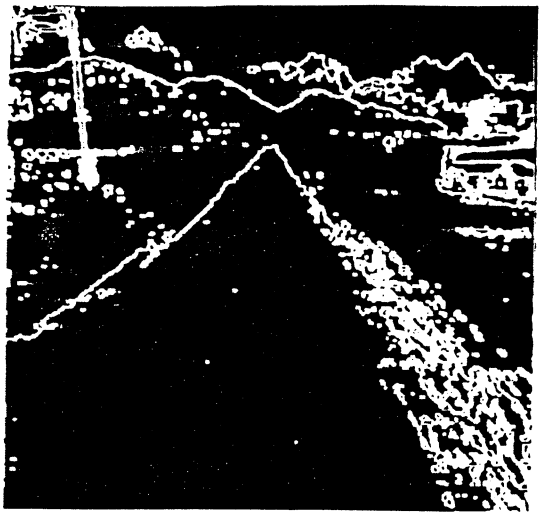

Fig. 7(b) Initial segmentation results.

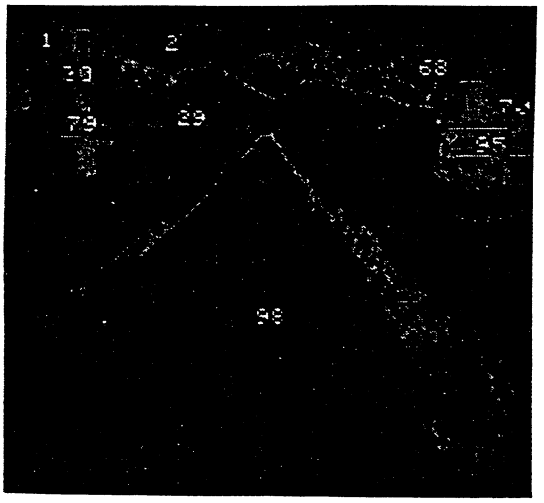

Fig. 7(c) Segmented image after region splitting, merging and discarding small regions. Regions highlighted by rectangles indicate the result of landmark recognition by PREACTE.

	Features							Recognition Results	
Region	Size	Color	MBR	Texture	Elongation	Shape	Location	Match-Evidence	Landmark Hypothesis
30	Small (91)	Black (25.3)	(99, 103, 61, 111)	Smooth	High (50:4)	Long and Linear	(101.2, 82.9)	0.83	T-POLE-110
79	Small (104)	Black (27.3)	(104, 105, 112, 132)	Irregular	High (20:1)	Long and Linear	(104.6, 121.9)	0.92	T-POLE-110
95	Medium (675)	White (224.5)	(445, 510, 140, 155)	Smooth	Low (05:15)	Not convex	(482, 148.3)	0.62	G-TANK-110
70	Medium (672)	Gray (199.7)	(469, 510, 99, 127)	Irregular	Low (41:28)	Linear	(490.2, 115. 3)	0.71	BLDG-110

Table I. Landmark recognition results.

References

[1] B. Bhanu, "Recognition of Occluded Objects," Proc. of the 8th International Joint Conference on Artificial Intelligence, IJCAI-83, Karlsruhe, West Germany, August 8-12, 1983, pp. 1136-1138.

[2] B. Bhanu, "Representation and Shape Matching of 3-D Objects," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, May 1984, pp. 340-351.

[3] B. Bhanu and O.D. Faugeras," Shape Matching of Two-dimensional Objects," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, March 1984, pp. 137-156.

[4] B. Bhanu and T. Henderson," CAGD Based 3-D Vision," IEEE International Conference on Robotics and Automation, March 1985, pp. 411-417.

[5] B. Bhanu and W. Burger, "DRIVE: Dynamic Reasoning Using Integrated Visual Evidences," Proc. DARPA Image Understanding Workshop, University of Southern California, Feb. 1987.

[6] T.O. Binford, "Survey of Model-Based Image Analysis," The International Journal of Robotics Research, Vol. 1, Spring 1982, pp. 18-64.

[7] R. A. Brooks, "Symbolic Reasoning Among 3-D Models and 2-D Images," Artificial Intelligence, Vol. 17, pp. 285-348, 1981.

[8] E. Charniak, "The Bayesian Basis of Common Sense Reasoning in Medical Diagnosis," Proc. American Association of Artificial Intelligence Conference 1983, AAAI-83, pp. 70-73.

[9] E. Davis, Representing and Acquiring Geographic Knowledge, Morgan Kaufman Publishers, Inc. 1986.

[10] S.V. Hwang, "Evidence Accumulation for Spatial Reasoning in Aerial Image Understanding," Ph.D. Thesis, Dept. of Computer Science, University of Maryland, College Park, Maryland, 1984.

[11] D. Lowe, Perceptual Organization and Visual Recognition, Kluwer Publishing Co., 1985.

[12] A.R. Hanson and E.M. Riseman, "VISIONS: A Computer System for Interpreting Scenes," in Computer Vision Systems, A.R. Hanson and E.M. Riseman, (Eds.), New York, Academic Press, 1978, pp. 303-333.

[13] D.M. McKeown, Jr., W.A. Harvey, Jr., and J. McDermott, "Rule-Based Interpretation of Aerial Imagery," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-7, September 1985, pp. 570-585.

[14] M. Nagao and T. Matsuyama, A Structural Analysis of Complex Aerial Photographs, Plenum Press, 1980.