Vision-Based Localization*

William B. Thompson, Thomas C. Henderson,
Thomas L. Colvin, Lisa B. Dick, and **Carolyn M. Valiquette**
Department of Computer Science
University of Utah
Salt Lake City, UT 84112

Abstract

Localization based on visual landmarks requires feature extraction from views and map, matching of features between views and map, and viewpoint hypothesis generation and verification. In this paper, we describe lower-level image and map understanding procedures for extracting features and higher-level problem solving methods for establishing feature correspondences and making inferences about the viewpoint. Each of these processes, including the interaction of high-level and low-level subsystems, is demonstrated on real data.

1 Introduction.

An essential aspect of map-based navigation is the determination of an agent's current location based on sensed data from the environment. Formally, this amounts to specifying the current viewpoint in some world model coordinate system. This *localization* process has two distinct components: one involving the establishment of correspondences between aspects of the sensed data and the map or model and the other involving derivation of constraints on the viewpoint based on the correspondences that have been determined.

Correspondences can be established at the signal or feature level. Signal-level matching correlates sensed data with predictions of how the sensed data should appear. It works best when the uncertainty in the viewpoint is small and when it is relatively easy to accurately generate expected sensor data. For example, in the TERCOM and SITAN cruise missile guidance systems, a digital elevation model is matched against a downward looking, radar sensed elevation profile [Andreas *et al.*, 1978, Baird and Abramson, 1984]. Several researchers have addressed the more difficult problem of signal-based localization at or near ground level using horizontally oriented imaging systems and passive sensing. In [Ernst and Flinchbaugh, 1989], deviations between expected and observered views are determined using curve matching algorithms and a weak perspective model is used to update the position estimate. [Yacoob and Davis, 1991] and [Talluri and Aggarwal, 1992] determine viewpoint under the assumption that viewpoint elevation is known with high precision in the reference frame of the map, a situation which dramatically reduces complexity but is unfortunately not likely to hold in practice. [Stein and Medioni, 1992] proposed an alternate method for determining viewpoint based on the observed horizon line which is similar to the characteristic view approach in object recognition.

Vision-based navigation in unstructured terrain can violate many of the assumptions used in the approaches described above. Often there is limited a priori knowledge about the viewpoint due to travel through indistinct terrain, temporary occlusion of landmark features, or errors in position updating processes. The view of the world at or near ground level is difficult to generate from map data with sufficient fidelity to allow signal-level matching. Furthermore, available digital cartographic data sets often contain inaccuracies that can cause serious problems for correlation-based analysis. For example, in one of the USGS DEMs that make up our test data, the location of the high point of a significant peak is off by over 200m. It is not surprising that most of the published work on vision-based localization from a ground-level perspective has been demonstrated only on synthetic data, where these problems do not occur.

With signal-based techniques, actual viewpoint determination is done using the same types of methods involved in photogrammetry (which solves the same problem) [Sanso, 1973, Thompson, 1958] or in alignment approaches to object recognition [Huttenlocher and Ullman, 1987, Grimson, 1990]. The principal shortcoming in both these methods is the difficulty of introducing realistic error models or effective representations of the uncertainty in viewpoint estimates.

Feature-based approaches hold the potential for avoiding many of these problems. Features are extracted independently from sensed data and maps and then matched symbolically. As a result, there is no longer a need to be able to synthesize an accurate rendition of expected sensed data. The symbolic nature of matching and viewpoint inference allows the introduction of sophisticated problem solving methods which are able to deal with issues such as ambiguity and complex error models.

In the remainder of this paper, we describe one possible approach to feature-based localization in unstruc-

*This work was supported by National Science Foundation grant IRI-9196146, with partial funding from the Defense Advanced Research Projects Agency.

tured, outdoor terrain. We outline methods for extracting terrain features from maps and image data, show how matching can be performed, and describe a collection of qualitative geometric reasoning procedures for determining viewpoint while maintaining an explicit representation of the uncertainty associated with that determination. The approach is demonstrated on a real example involving imagery obtained with a video camera and map data provided by the USGS.

2 Feature Extraction.

Three classes of entities are central to the localization process: *Terrain* is the physical layout of the land. *Maps* are geometric representations of a particular region of terrain, typically from a downward-looking perspective and possibly augmented with information about culture and/or vegetation. *Views* are visually sensed images of a particular region of terrain.

Each class of entities can be described in terms of *features*. In the case of terrain, features are commonly used geographic properties: hills, valleys, ridges, etc. These features can exist across a range of scales, specified in terms of physical extent. (We never actually deal with terrain features, only with manifestations of such features in the map and view.) In the case of maps and views, we need to distinguish between *data-level* and *terrain-level features*: Data-level features are distinctive patterns in the data (e.g., a configuration of edge fragments in a view or a locally defined topographic structure in a map). Terrain-level features are patterns of data-level features likely to correspond to some particular terrain feature. Terrain features, terrain-level map features, and terrain-level view features are distinct, even though they may have common names.

2.1 View features.

Currently, we are concentrating on those view features associated with occluding contours. Because the imagery is acquired from a horizontal perspective, these typically correspond to ridge lines. Ridge line extraction is a classical segmentation problem. The type of data we are working with, however, causes significant difficulties. Image contours corresponding to actual ridge lines should be long, connected, and relatively smooth. Except in pathological cases, they should never fold back on themselves. While this might suggest an approach which looks for "large scale" image features, things are not so simple. Contrast variations across edges that correspond to actual ridge lines can be small and of limited spatial extent. For portions of many ridge lines, contrast variations can be lacking altogether. As a result, scale-space approaches will not succeed.

Instead, we use an approach similar to [Sha'ashua and Ullman, 1988, Nevatia *et al.*, 1992]. An initial edge map is computed using a zero-crossing edge detector. Edge segments are alternately filtered to remove portions inconsistent with the geometric properties of ridge lines and augmented using properties of good continuation to account for locally indistinct ridge segments. Extraction of longer ridge lines is done using A* search [Martelli,

Figure 1: Original Image

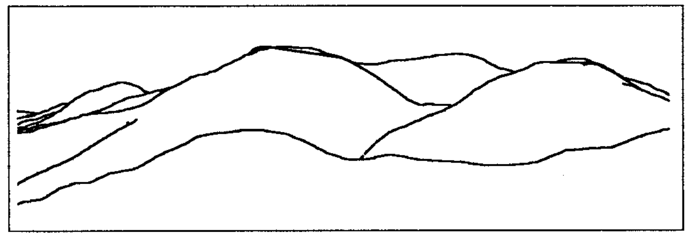


Figure 2: Actual ridge lines.

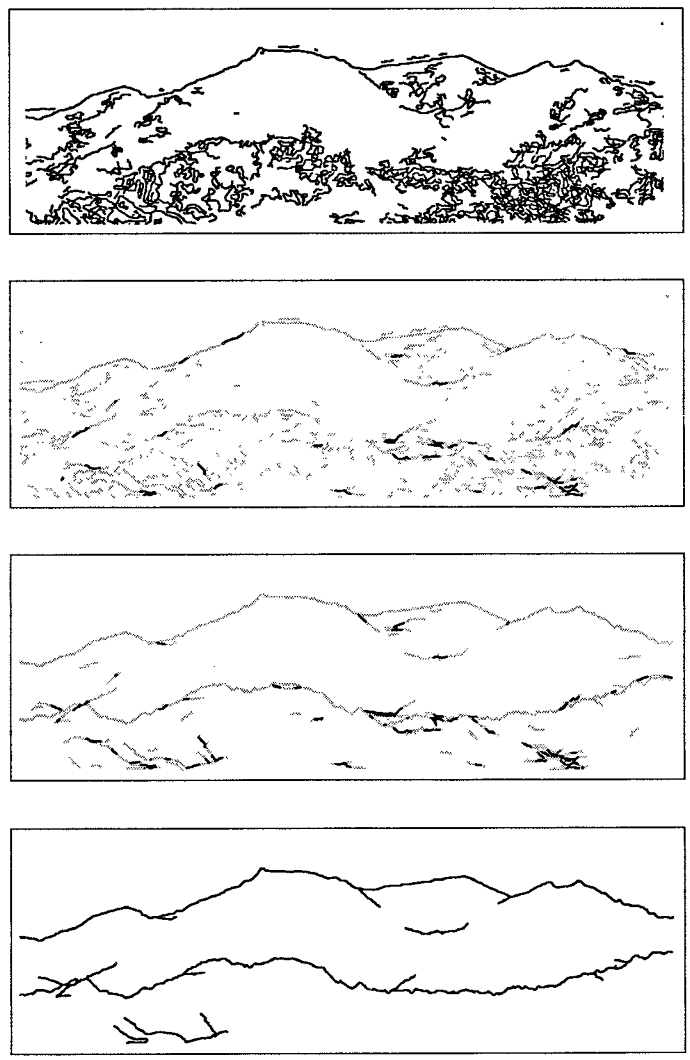
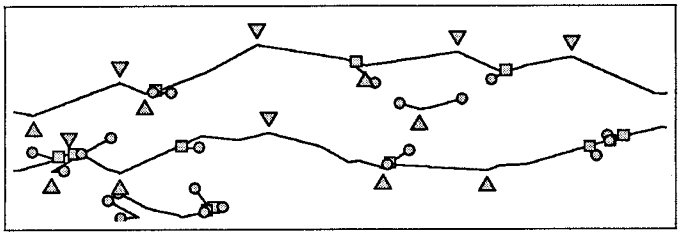


Figure 3: Edge filtering and gap filling



▽ – peaks, △ – saddles, ▣ – junctions, ○ – endpoints

Figure 4: Extracted features

1972] which allows the specification of more global optimization criteria. This is particularly important for the horizon line, which can be difficult to find due to clouds or aerial perspective. Once these operations are completed, junctions, end points, and vertical contour extrema are located, since these often correspond to topographically relevant features.

Figure 1 shows an image of mountainous terrain. Figure 2 shows the actual ridge lines apparent from the viewpoint associated with Figure 1, as determined from DEM data. Figure 3 shows four stages in the edge filtering/gap filling process. The first frame is the output of a hysteresis thresholded zero-crossing edge detector. The next two frames show intermediate results, with filled gaps indicated by darker lines. The last frame shows the final edges. Figure 4 shows extracted line segments, peaks, saddles, T-junctions, and end points.

2.2 Map features.

The extraction of map features involves different problems than those associated with the view, but many of the processing steps are similar. (The cartographic community has done related work, but not specifically in support of localization.) Since we are operating directly on elevation data, we do not need to deal with the ambiguity associated with low-level contrast features in the view. However, we do have to find long ridge contours that may not be immediately apparent at a given scale. The analysis starts with a characterization of local surface shape of the map in terms of ridges, valleys, peaks, and saddles using the method described in [Haralick *et al.*, 1983]. Instead of resampling to produce precise ridge lines, we found it sufficient to impose thresholds when extracting ridge lines and use a thinning algorithm to extract ridge contours.

Navigationally salient ridge lines and peaks cannot be detected from an analysis of features extracted from local differential properties alone. Visually prominent ridge lines often contain broad sections where the spatial derivatives of elevation are low, resulting in a classification as flat ground and so creating breaks in the ridge lines. Local maxima in elevation may or may not correspond to visually identifiable peaks, depending on the nature of surrounding peaks and saddles. Relatively simple gap filling and filtering operations can significantly improve the utility of features extracted using local methods.

The features resulting from this process span a wide

Figure 5: Extracted peaks and ridges.

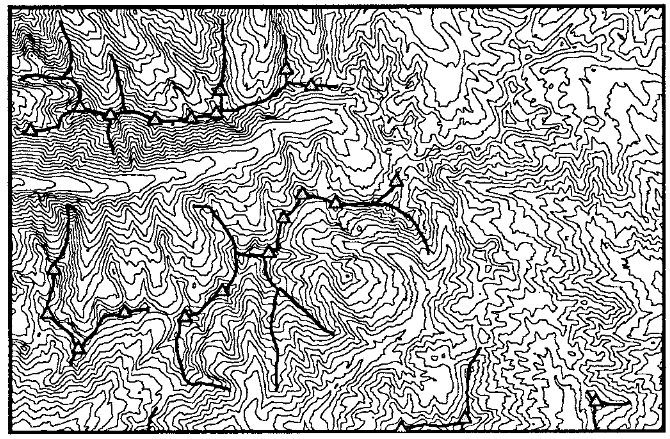


Figure 6: Extracted peaks and ridges at coarser scale.

range of spatial scales. Again, the linear nature of ridge lines limits the value of a straightforward scale-space analysis. A local analysis of ridge line junctions has proven adequate for distinguishing between dominant and subsidiary ridges. This allows the creation of a graph-like description of ridge structure, since spur ridges can in turn contain sub-spurs. Access to this hierarchy can prove significantly beneficial in feature matching. At initial stages of the matching process, only main ridges should be considered. When precise localization hypotheses are being evaluated, however, the detailed structure of the ridge line may become relevant. The hierarchical description makes it easy to avoid this level of detail unless needed.

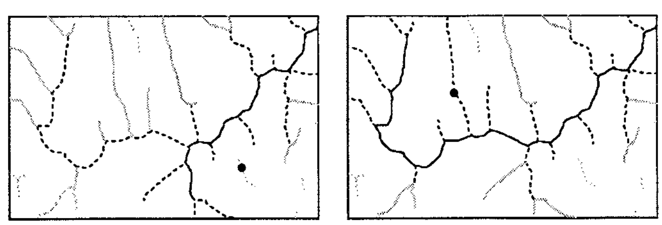


Figure 7: Alternate, viewpoint dependent ridge hierarchies.

Local analysis at times is unable to distinguish with confidence the major and minor ridges coming into a junction. Rather than being a deficiency of the approach, it may provide useful information. These situations are exactly those where there is a viewpoint dependence in the ridges which needs to be attended to. As a result, we can isolate the viewpoint dependent aspects of the representation, use with greater confidence those parts that show no obvious ambiguity, and distinguish between dominant and subsidiary features if and when hypotheses about the viewpoint are available.

Figure 5 shows significant peaks and ridges extracted at one particular spatial scale overlaid onto the corresponding contour map. Figure 6 shows peaks and ridges extracted at a coarser spatial scale. Figure 7 illustrates the hierarchical nature of extracted ridge features. "Dominant" ridge lines are often viewpoint dependent, as shown in the two parts of the figure, each with a view position indicated by a black dot.

3 Matching and Geometric Inference.

Feature-based localization involves problem solving [Heinrichs *et al.*, 1992]. The integration of symbolic problem solving with signal-level image analysis has long been a goal for many in the computer vision community. Few successful examples exist, however. In our case, we are able to effect this integration by restricting ourselves to a specific task and establishing a protocol for the interaction between high and low level analysis routines that is tailored to that task. The problem solving component of the system interacts with the feature extraction modules as if they were databases. Query and response languages were defined that make it possible to easily express relevant information about terrain features. Geometric inference is integrated in a similar manner. The result is a system in which the individual components can be constructed in a nearly independent manner, without a need to understand the details of internal representations and algorithms of other modules. Figure 8 shows the basic organization.

Overall control is determined by the high-level matching and inference system. Both top-down and bottom-up feature extraction is easily accomplished, however. For example, early in the localization process reconnaissance queries can request a general examination of map or view to determine significant features. Later, expectations can be verified in a top-down manner by generating highly constrained queries and examining whether or not any items are returned.

3.1 Matching.

One key observation arising from our study of how expert map users solve difficult localization problems is that they organize map and view features into *configurations* before attempting to match them [Pick *et al.*, in press]. Configurations are small groupings of features (typically two or three) that are close together and often satisfy particular topographic and/or geometric properties that make them distinctive. Matching configurations rather than individual features significantly reduces the combinatorics in two ways: there are fewer candidates for

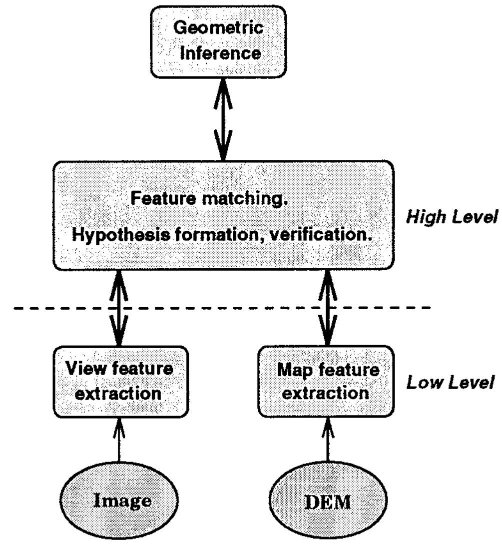


Figure 8: Interaction of high-level and low-level subsystems.

matching and each potential match involves a richer set of properties which can be evaluated for compatibility.

Complexity and possible ambiguity are further reduced by forming target configurations that are likely to be matched. Each map and view feature has a set of associated properties which constitute a geometric description of its shape and position. Particular property values such as high or sharp peaks or near level ridges are an indication that the feature is likely to be easy to find in both map and view. Specific combinations of feature properties are used to compute a set of *prominence* values, represented using a number in the range $[0.0-1.0]$. Prominence alone is a poor criterion by which to select features for forming configurations, however, since it is computed on a per-feature basis and it may turn out for a particular case that there are many features high in some particular property prominence. The *distinctiveness* of a particular prominence type is characterized by a value that is large when the population of features is such that a few have large values for the prominence in question and the rest have small values. The *saliency* of each feature property is computed by multiplying the property prominence by the property distinctiveness. Finally, the overall saliency for features is computed using a simple product rule that favors features with several highly salient prominences:

$$S_{\text{overall}}(f_i) = 1.0 - \prod_j (1.0 - S_j(f_i))$$

where $S_{\text{overall}}(f_i)$ is the overall saliency of feature (f_i) and $S_j(f_i)$ is the individual saliency of the j-th property of feature (f_i).

The formation of configurations is implemented by first sending a *reconnaissance* query to either the map or view feature extraction subsystems, requesting that individual features with high prominence be returned. These are filtered to remove all but the features with the highest overall saliency. Configurations are then formed with a *combination* query to the same low-level subsystem, requesting any sets of features that contain at least one of the salient individual features and satisfy particular geometric and/or topographic properties. Any configurations that result from this process can be searched for using a similar combination query to the other low-level feature extraction subsystem.

3.2 Inference.

Localization involves the determination of viewpoint constraints based on possible correspondences between image and map features. These constraints are used in geometric reasoning operations that either hypothesize a possible viewpoint or evaluate a hypothesis by predicting additional constraints that should be satisfied. There are distinct categories of information about feature position that in turn lead to distinct constraints on viewpoint:

Absolute bearing: This is the "standard" way to solve localization problems. It requires an accurate compass registered to the map coordinate system. Determination of viewpoint is done using straightforward trigonometry.

Relative bearing: Relative bearings between three or more image features with known map positions lead to a classical "pose estimation" problem. Well established numerical techniques exist for solving such problems. [Levitt *et al.*, 1987] describe an alternate method in which only two features are considered at a time. The visual angle between the two features constrains the viewpoint to lie on a particular circle on the map. Using multiple pairs of features usually allows a unique viewpoint to be found by intersection.

Ordinal view: [Levitt *et al.*, 1987] show how ordinal position of two features (e.g., "**A** is *left-of* **B**") can be used to constrain the viewpoint to lie on one side of a line through the positions of **A** and **B** [Levitt *et al.*, 1987, Levitt *et al.*, 1988]. They suggest intersecting this constraint for many different pairs of features.

Exact Alignment: If two features line up along a line of sight, then the viewpoint is constrained to lie on a line connecting the two features. In almost all circumstances encountered in outdoor navigation, it is possible to determine which of the two features is more distant and as a result the viewpoint can be constrained to a half-line.

Approximate Alignment: If two features are much closer laterally (i.e., perpendicular to the line of sight) than in depth (i.e., parallel to the line of sight), then not only is the viewpoint constrained to lie on one side of a line connecting the two features, but it will be "near" this line. This constraint appears to be used with some frequency by expert map users solving real navigation problems.

Viewpoint terrain type: A locomoting agent often has more complete information about its immediate environment than it does about more distant aspects of the terrain. This information relates in a direct manner to the determination of viewpoint. (E.g., "I'm standing on a hill. Therefore, the viewpoint must be located at one of the hills found represented on the map.")

Constraints arising from the information sources listed above can be divided into two categories. All but the last constitute *distant constraints*, since they are based on features distant from the viewpoint. Trigonometric relations are required to relate distant constraints to viewpoint, although qualitative as well as quantitative solutions exist. Viewpoint terrain type leads to *local constraints* which limit the viewpoint to compatible terrain features on the map. Distant and local constraints can be used in three kinds of reasoning about viewpoint (Figure 9):

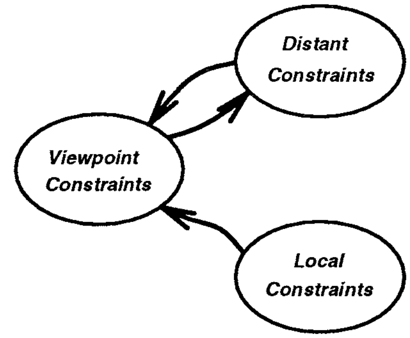


Figure 9: Reasoning about viewpoint involves interacting constraints.

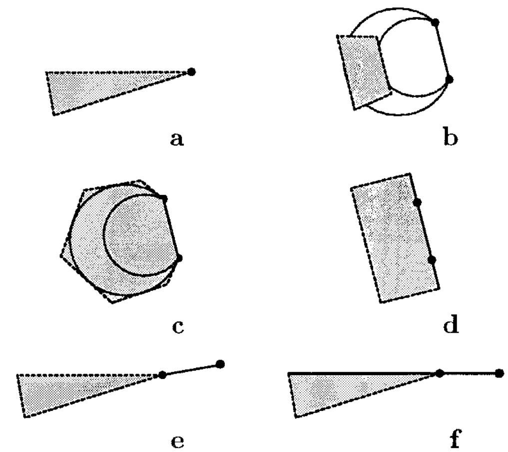


Figure 10: Distant constraints on viewpoint: **a)** absolute bearing, **b)** relative bearing if approximate depth information is available, **c)** relative bearing if no depth information is available, **d)** ordinal position, **e)** exact alignment, **f)** approximate alignment

Distant constraints ⇒ constraints on viewpoint: Map-view feature correspondences for sets of distant features can be used to determine constraints on viewpoint using geometric reasoning methods applied to any combination of the information sources described above.

Local constraints ⇒ constraints on viewpoint: Local constraints allow for the enumeration of possible viewpoints. Such an enumeration can be intersected with the constraint regions that usually arise from consideration of distant features.

Constraints on viewpoint ⇒ expectations about distant constraints: Hypotheses about viewpoint can be evaluated by examining distant features using the geometric

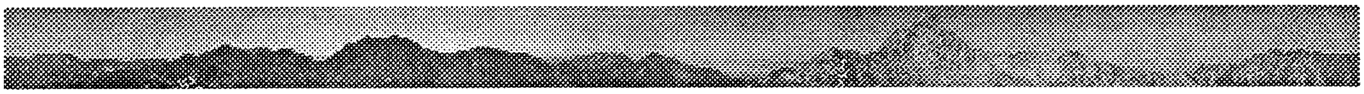


Figure 11: View.

reasoning methods applied to any combination of the information sources described above. Positional and/or orientational constraints on viewpoint can be exploited.

In order to implement the constraint satisfaction shown in Figure 9, geometric representations are needed for *viewpoint regions* (areas in the map corresponding to possible viewpoints), *map search regions* (map regions possibly containing terrain features visible in the view), and *view search regions* (portions of the view in which particular terrain features indicated by the map are expected to be found). A variety of representational formalisms are possible, each with advantages and disadvantages. [Sutherland and Thompson, 1993] use an analytical description of the bounding curve associated with the region in which there is any chance that the viewpoint is located. The localization example shown in section 4 uses a much simpler convex polygon representation. This provides a compact description that is efficient to manipulate. It also fits fairly well with people's intuition about the geometry involved. With a few exceptions, convex polygons have proven to be an adequate basis on which to build the geometric constraint satisfaction algorithms, though they sometimes lead to a very conservative approach such as describing the relative bearing constraint using a circle rather than a crescent (see Figure 10). Figure 10 indicates how distant constraints on viewpoint can be represented using the convex polygon approach. Similar regions have been defined for using viewpoint to develop constraints on view and map positions [Thompson, 1993].

Geometric inference is implemented using a query language similar in principle to that used to interact with the low-level feature extraction subsystems. Viewpoint regions are hypothesized or refined using a query containing a (possibly empty) current viewpoint region hypothesis, a set of corresponding map and view feature locations, and a particular inference method to use. The geometric inference subsystem applies the method to produce a viewpoint region constraint, intersects this with the initial regions supplied, and returns the result. Map search regions result from queries which specify a current viewpoint region hypothesis, a set of view feature locations, and an inference method. View search regions are obtained in an analogous manner.

4 Example.

We have demonstrated the sufficiency of the approach described above by applying it to a real example from the mountains just southeast of Salt Lake City, UT. View features were extracted from the panorama image shown in Figure 11. Note that, rather than the synthesized views commonly used in much of the reported research on outdoor localization, we are using an actual terrain image. Map features were extracted from USGS 30m DEM data covering the equivalent of four 1:24000 7.5′ quadrangles (approximately 21.4 km by 28 km), the upper half of which appears in Figures 5 and 6. The compass orientation of the view was known, but no information about viewpoint was provided other than that it was somewhere within the available map area.

The problem solving subsystem responsible for feature matching and hypothesis generation and evaluation was implemented in Lisp. The geometric inference subsystem was also implemented in Lisp and was interfaced via function calls. Both of the low-level feature extraction subsystems were implemented in C, ran on different machines than the Lisp processes, and were interfaced using simple database-like query/response techniques.

The example used four of the six high-level reasoning strategies described in [Heinrichs *et al.*, 1992]: concentrate on the view first, organize features into configurations, pursue multiple hypotheses, and evaluate hypotheses using a disconfirmation process. Two types of distant constraints on viewpoint (section 3.2) and one type of constraint for determining map search regions based on hypothesized viewpoint were used. Execution proceeded in five stages:

View reconnaissance: The view (image) was searched for significant features. The highest peak in the image stood out well above other features in overall saliency.

Form view configurations: View configurations were formed from the selected peak feature and other nearby features that were prominent. Prominence rather than saliency was used, since the configuration is more distinct than its components. Two dual-feature configurations resulted, one involving the horizontal ridge segment to the left of the peak and one involving the ridge segment to the right. In fact, the second of these was a bad choice. The ridge to the right is actually quite distant from the peak, but this was not detected by the low level image analysis routines since the corresponding T junction was not found. Simultaneous consideration of multiple hypotheses combined with the disconfirmation strategy resulted in a system tolerant of such errors.

Search for configurations in map: Configurations consisting of a high, sharp peak and a nearby horizontal ridge were searched for in the map. Three such configurations were found.

Generate initial hypotheses: Six configuration matching hypotheses were postulated (two view configurations times three map configurations). Each configuration match specified two feature matches which were used to generate a hypothesized viewpoint region using the relative bearing constraint, as shown in Figure 12.

Refine hypotheses and evaluate using disconfirmation strategy: For each hypothesis, highly salient view fea-

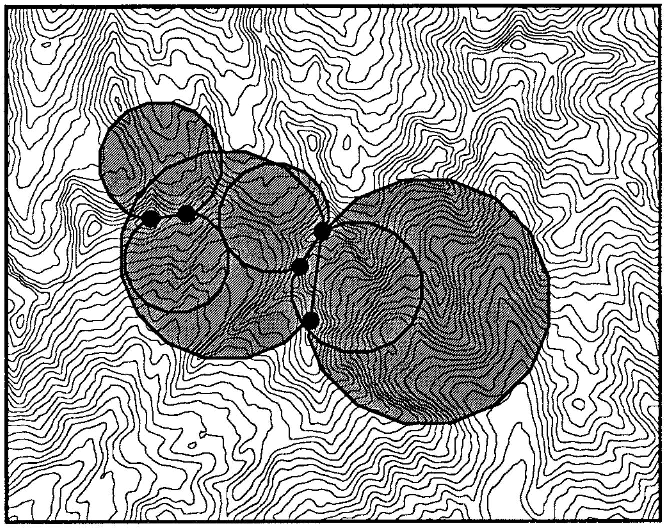


Figure 12: Viewpoint regions corresponding to the six initial hypotheses.

tures were considered in turn and searched for in the map. If exactly one map feature of the correct type was found in the expected location, a match was established and the absolute bearing constraint was used to refine the viewpoint region. If two or more map features were found, no inferences were drawn and the next view feature was processed. If no map feature was found where one was expected, the hypothesis was disconfirmed.

Figure 13 shows the refinement of the hypothesis corresponding to the actual viewpoint. Four view features were searched for in the map: the high peak mentioned previously, two other peaks towards the left of the panorama, and the long ridge line that wraps around from the right edge to the left edge of the panorama image. Three unique matches were found, involving two of the peaks and the long ridge. The remaining peak was ambiguous, with two possible peaks in the map located in positions that could plausibly correspond to the location of the view feature. On the left of the figure is shown the search for corresponding map features. The current viewpoint hypothesis is show together with the search region predicted from the bearing to the chosen view feature. Black dots indicate map features that were found. For the first three view features considered, a unique map feature was found. On the right is shown the current viewpoint, the constraint regions associated with the map feature just found, and the intersection which becomes the refined viewpoint region. The last map search returned an ambiguous result, as can be seen by the two features present within the search region. As a result, no refinement of the viewpoint region was possible. The plot on the lower right of the figure shows the final region, with the actual viewpoint marked. The original viewpoint hypothesis had an area of approximately 1.489 km^2. After the first absolute bearing constraint was imposed, the size of the region was down to 72,800 m^2. The second absolute bearing constraint left this unchanged. The final constraint reduced the area to less than 71,700 m^2, or about .07 km^2.

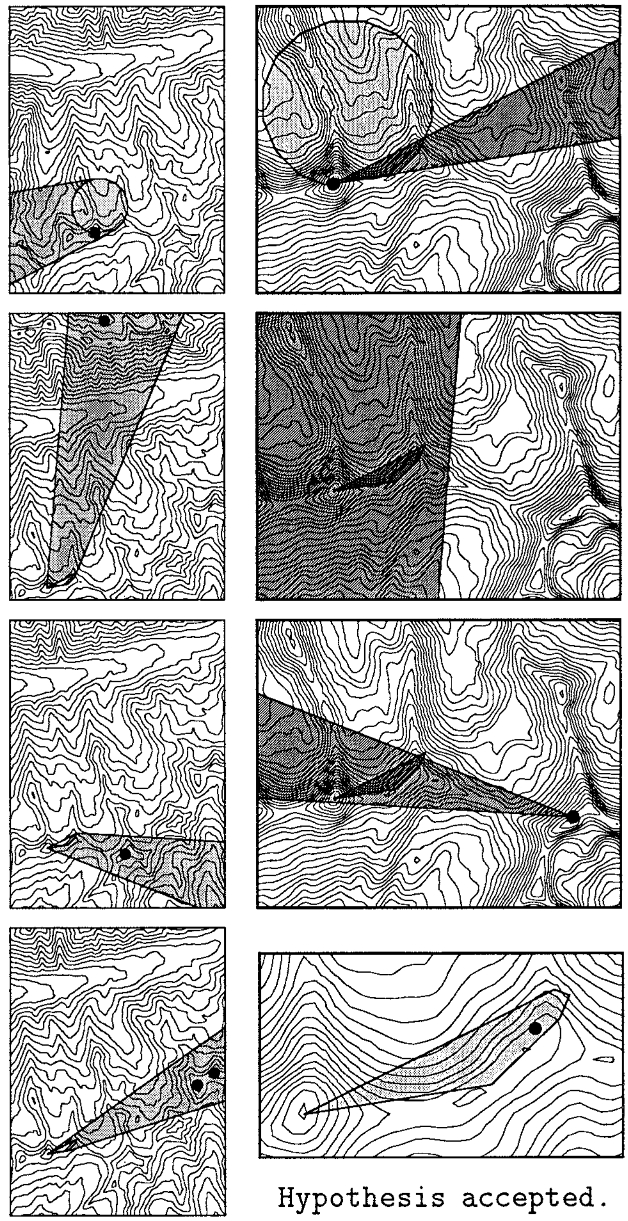


Figure 13: Viewpoint refinement and final region.

Figure 14 shows an attempt to refine one of the incorrect hypotheses. The upper left panel shows the map search region used to look for the most salient view peak. The feature being searched for is shown as an open circle. Due to the fact that the hypothesized viewpoint is wrong, a different feature was found, as indicated by the black dot. The viewpoint region was refined based on this match as shown to the right. The next most salient view feature was the long ridge line. As shown in the lower left panel, this was not found where expected and so the hypothesis was rejected All five hypotheses not including the true viewpoint were rejected in a similar manner.

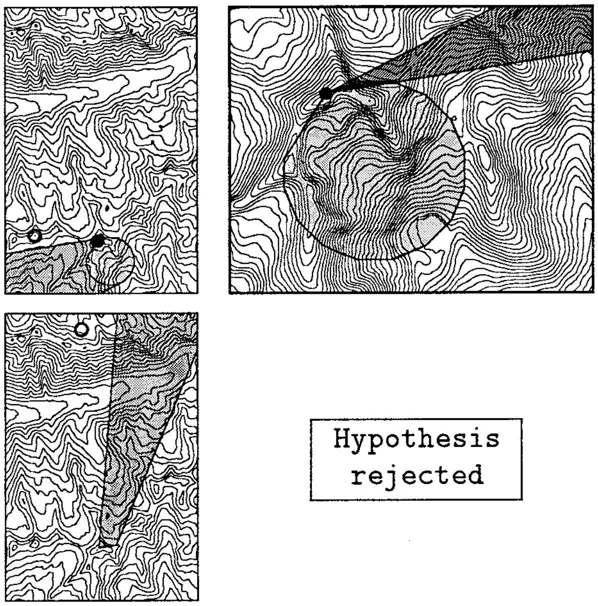


Figure 14: Validation of "incorrect" hypothesis.

Acknowledgements.

Rory Cejka provided many of the support tools used in this effort and assisted in the development of the edge thinning method. Sean Igo implemented the geometric reasoning subsystem.

References

[Andreas *et al.*, 1978] R.D. Andreas, L.D. Hostetler, and R.C. Beckmann. Continuous Kalman updating of an inertial navigation system using terrain measurements. In *Proc. IEEE National Aerospace Electronics Conference*, pages 1263–1270, 1978.

[Baird and Abramson, 1984] C.A. Baird and M.R. Abramson. A comparison of several digital map-aided navigation techniques. In *Proc. IEEE Position Location and Navigation Symposium*, pages 294–300, 1984.

[Ernst and Flinchbaugh, 1989] M.D. Ernst and B.E. Flinchbaugh. Image/map correspondence using curve matching. In *AAAI Symposium on Robot Navigation*, pages 15–18, March 1989.

[Grimson, 1990] W.E.L. Grimson. *Object Recognition by Computer: The Role of Geometric Constraints.* The MIT Press, 1990.

[Haralick *et al.*, 1983] R.M. Haralick, L.T. Watson, and T.J. Laffey. The topographic primal sketch. *IEEE Journal of Robotics and Automation*, 2(1):50–72, 1983.

[Heinrichs *et al.*, 1992] M. R. Heinrichs, K. Smith, H. L. Pick, Jr., B. H. Bennett, and W.B. Thompson. Strategies for localization. In *Proc. DARPA Image Understanding Workshop*, January 1992.

[Huttenlocher and Ullman, 1987] D.P. Huttenlocher and S. Ullman. Object recognition using alignment. In *Proc. First International Conference on Computer Vision*, pages 102–111, 1987.

[Levitt *et al.*, 1987] T.S. Levitt, D.T. Lawton, D.M. Chelberg, and P.C. Nelson. Qualitative navigation. In *Proc. DARPA Image Understanding Workshop*, pages 447–465, February 1987.

[Levitt *et al.*, 1988] T.S. Levitt, D.T. Lawton, D.M. Chelberg, K.V. Koitzsch, and J.W. Dye. Qualitative navigation II. In *Proc. DARPA Image Understanding Workshop*, pages 319–326, April 1988.

[Martelli, 1972] A. Martelli. Edge detection using heuristic search methods. *Computer Graphics and Image Processing*, 1(2):169–182, 1972.

[Nevatia *et al.*, 1992] R. Nevatia, K. Price, and G. Medioni. USC image understanding research: 1990–1991. In *Proc. DARPA Image Understanding Workshop*, pages 3–25, January 1992.

[Pick *et al.*, in press] H.L. Pick, Jr., M.R. Heinrichs, D.R. Montello, K. Smith, C.N. Sullivan, and W.B. Thompson. Topographic map reading. In J. Flach, P.A. Hancock, J.K Caird, and K. Vicente, editors, *Ecology of Human-Machine Systems.* Lawrence Erlbaum Associates, (in press).

[Sanso, 1973] F. Sanso. An exact solution of the roto-translation problem. *Photogrammetria*, 29:203–216, 1973.

[Sha'ashua and Ullman, 1988] A. Sha'ashua and S. Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *Second International Conference on Computer Vision*, pages 321–327, December 1988.

[Stein and Medioni, 1992] F. Stein and G. Medioni. Map-based localization using the panoramic horizon. In *IEEE International Conference on Robotics and Automation*, Nice, France, May 1992.

[Sutherland and Thompson, 1993] K.T. Sutherland and W.B. Thompson. Inexact navigation. In *Proceedings of the IEEE International Conference on Robotics and Automation*, May 1993.

[Talluri and Aggarwal, 1992] R. Talluri and J.K. Aggarwal. Position estimation for an autonomous mobile robot in an outdoor environment. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8:573–584, October 1992.

[Thompson, 1958] E.H. Thompson. An exact linear solution of the problem of absolute orientation. *Photogrammetria*, 13(4):163–178, 1958.

[Thompson, 1993] W.B. Thompson. Geometric constraints for viewpoint determination. University of Utah Computer Vision Group working paper, 1993.

[Yacoob and Davis, 1991] Y. Yacoob and L.S. Davis. Computational ground and airborne localization over rough terrain. Technical Report CAR-TR-591, Computer Vision Laboratory, University of Maryland, 1991.