Proceedings of the 1997 IEEE
International Conference on Robotics and Automation
Albuquerque, New Mexico - April 1997

Automatic Mountain Detection and Pose Estimation for Teleoperation of Lunar Rovers

Fabio Cozman Eric Krotkov
Robotics Institute, Carnegie Mellon University
Pittsburgh, PA, 15213
{ fgcozman, epk }@cs.cmu.edu

Abstract

This paper presents a system aimed at mobile robot operations in space: we discuss an interface which receives and analyzes images sent by a rover operating in a distant environment. We are particularly interested in long-duration space missions, where rovers interact with human operators on Earth. The position estimates are presented to the operator so as to increase situational awareness and prevent loss of orientation. The system detects mountains in images and automatically searches for mountain peaks in a given topographic map. We introduce our mountain detector algorithm and present a large number of illustrative results from images collected on Earth and on the Moon (by the Apollo 17 mission). We present an algorithm for position estimation which uses statistical descriptions of measurements to produce estimates, and discuss results for the Pittsburgh East and Dromedary Peak areas. The implemented system achieves better estimation performance than any competing method due to our quantitative approach and better time performance due to our pre-compilation of relevant data.

1 Introduction

Only with great difficulty can human operators teleoperate rovers in an unfamiliar environment based solely on imagery sent by the rover, even with maps of the rover's environment [1, 6, 9]. Teleoperating a rover presents further challenges when the rover is on a lunar mission [7], because of the 5-second round-trip communication delay [11], coupled with the unfamiliarity of the environment, less gravity, and variable surface properties. For example, astronauts in the Apollo missions had great difficulty determining distances from mountains and craters [5].

This paper presents a system that assists operators driving remote vehicles. The basic idea is to offload navigation functions, permitting the remote driver to concentrate on pilot functions without getting disoriented or lost. Figure 1 summarizes the idea. The operator observes images from the rover and looks at a topographic map of the imaged area. Position of the robot is unknown but constrained to lie in a region the size of the map. The images are analyzed; structures found in the map are marked. In this paper we report on an automatic detector of mountains and a position estimator that operates from the detected peaks. The ultimate goal is to overlay position information on the maps and on rover-acquired images, just as "augmented reality" systems for training and medical applications do [3].

The interface presents three windows to the operator. The first window carries video; we currently use a standard video display in a Silicon Graphics workstation to look at our footage. The second window displays panoramas formed from selected images, and indicates the results of the mountain detector. The third window, depicted in Figure 2, carries the map information. The map can be seen from above, as displayed, or rendered as seen from the ground. The map in Figure 2 shows the topography of the Apollo 17 site on the Moon, generated from the Apollo 17 Landing Area topophotomap [10].

The next section discusses the basic requirements of our system; we then present our vision-based mountain detector and results collected on terrestrial and lunar data. The position estimator is then described, together with results obtained for data in the Pittsburgh East and Dromedary Peak USGS quadrangles. We show the most complete set of test images to date and we report improvements in speed and accuracy relative to previous approaches. The implemented system achieves better estimation performance than any competing method due to our quantitative approach and better time performance due to our precompilation of relevant data.

0-7803-3612-7-4/97 $5.00 ©1997 IEEE

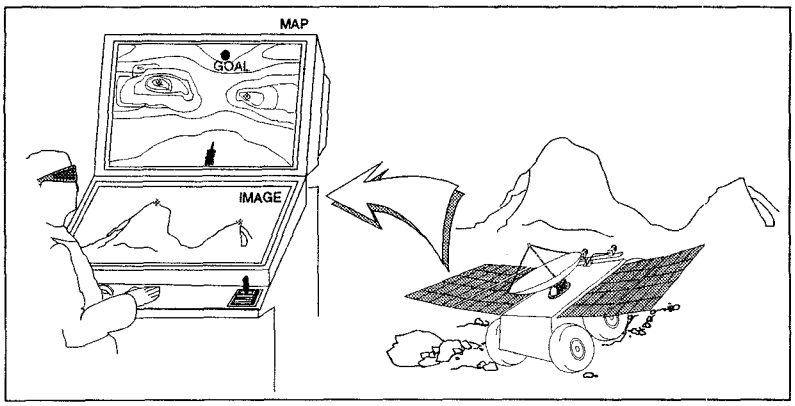

Figure 1: An interface for teleoperation of mobile robots

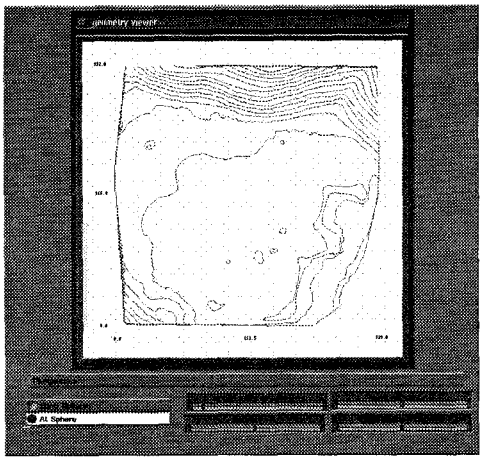

Figure 2: The map window, with a top view of the Apollo 17 topographic map (North and South Massifs appear on top and bottom respectively)

2 Localization from Outdoor Imagery

A variety of methods can be used to determine position, some using internal measurements (dead reckoning) or external measurements of motion. Since dead reckoning degrades over time, external measurements are necessary. A mobile robot performing a long traverse on the Moon or other celestial bodies poses a great challenge for position estimation technologies, due to the absence of the GPS infrastructure. On celestial bodies, attitude measurements can be generated accurately from a star sensor, since the visibility of stars from bodies without atmosphere is excellent, but position accuracy is low, on the order of kilometers [19]. In short, relative position and absolute orientation are available, but absolute position is not.

Our system uses mountains to fix position, as usually done by human sailors in marine navigation. Other researchers have studied the possibility of vision-based localization in outdoor applications [4, 8, 13, 16, 17], but real data analysis has been scarce (the only available data sets have been produced by Thompson and his group [17]), probably due to the complexity of collecting outdoor imagery coupled with reliable ground-truth data. In this paper we present a rich collection of images obtained by a customized platform, which allows us to tag every image with six dimensional ground-truth (position and orientation).

Each image feature can be converted to an angle in the robot's coordinate system; the angle is called a *bearing* [2]. The image contains a set R of m image bearings. The map contains a number n of landmarks $l_j = [X_j, Y_j]$. The pose of the robot in the global coordinate system is $\Gamma = [x, y, \phi]$ (ϕ is orientation). An interpretation I of the image is a set of correspondences between image features and map features.

To be able to construct an interpretation, the system must have information about the position of mountains in a topographic map of the rover's environment. To test our system with Earth images, we use 7.5 minute Digital Elevation Maps (DEMs) provided by the United States Geographical Survey (USGS), covering approximately 140km^2. Topographic maps are pre-processed and mountains in such maps are marked in an off-line stage, which automatically detects local maxima in the 7.5 minute DEM [18]. For example, the map in Figure 3 covers approximately 37 km^2 over the city of Pittsburgh, Pennsylvania. The black dots indicate the points where the system found mountains. The user has the ability to change or correct mountains detected at this stage, but we have never had to do that to obtain our results.

Figure 3: Topographic map of the Pittsburgh area with detected mountains

As the images are received by the system, mountains must be detected and position estimates have to be generated. Currently we do not operate directly on a stream of video; individual images are selected from the video window and manually pasted into a mosaic [15]. The mountain detector works on the mosaics. The mosaic is first segmented in two regions, one corresponding to the sky, another to the ground, and then the boundary between the two regions is searched for mountains. The theory and implementation of this mountain detector was described in a previous publication [2], together with preliminary results. Here we concentrate on a description of results obtained with a more comprehensive set of tests. The mountain detection system takes at most 2 seconds per full panorama in the current implementation on a Silicon Graphics workstation, depending on the complexity of the scenes.

Figure 4 shows a gallery of panoramas processed by our system. Images come from Pennsylvania, Utah, California and the Apollo 17 site on the Moon. Detected mountains are enclosed in rectangles; the detected position of a mountain peak is marked through a vertical line inside the rectangle.

The panorama (1) composes a sequence of images taken by the Apollo 17 Lander Module. The lunar cart can be seen at the extreme right of the panorama. The first large structure is the South Massif, followed by two smaller mountains and then the North Massif. The system detects all of those mountains plus a smaller formation which follows the North Massif. For lunar images, we have found that a very simple thresholding operation is enough to segment the sky from the ground reliably; that was used in panorama (1).

Panoramas (2) and (3) bring images from California; the first is by the Don Pedro Resort and the second is by the East entrance of the Niles Canyon. Mountains are very distinct in the Don Pedro area, but not so in the Niles Canyon. The large structures in this latter panorama are captured, but the peak in the extreme right has a skewed position because of trees on the top of the mountain. These images were acquired under bright sunlight, and the resulting panoramas suffer some effects of the camera's auto-iris operation: the sky shows "waves" of high brightness. This would make it impossible to use a simple thresholding operation as in the lunar images, but our sky segmentation algorithm [2] works without problems.

Panoramas (4) and (5) were acquired by Prof. William Thompson and his group and are publicly available. The detector has no problems, except in the detection of very small sequences of mountains that are quite distant.

Panoramas (6), (7) and (8) were obtained in Pittsburgh, by the Allegheny river. All distinct mountains are detected, except the big mountain in panorama (6), where the system gets confused by the pole in the extreme left. The system has proven to be quite reliable and flexible to adjustments, but some occasional misses or false detections occur as a result of unforeseen events like the pole in panorama (6). We plan to give some latitude to the operator, so that the system's conclusions can be modified when mistakes occur.

3 The Estimation Algorithm

Given the bearings, we must generate estimates of position. Two decisions must be made. First, what is the space in which we conduct the search. Second, how to establish a convenient figure of merit to evaluate the possible estimates of position.

The search for best estimates can be conducted in the space of interpretations [17], but only when using few peaks and image features; otherwise too many interpretations have to be visited. Use of few features compromises accuracy, which requires the use of large *and* small peaks in the horizon. Another search strategy is to look at the space of possible renderings of the map [13]. In order to conduct such a search, speed-up techniques like quantization must be used, again causing a loss of accuracy.

A more promising approach is to search the space of positions. To handle a USGS topographic quadrangle, we need to go through 4×10^4 positions and establish a figure of merit for each. This seemingly daunting

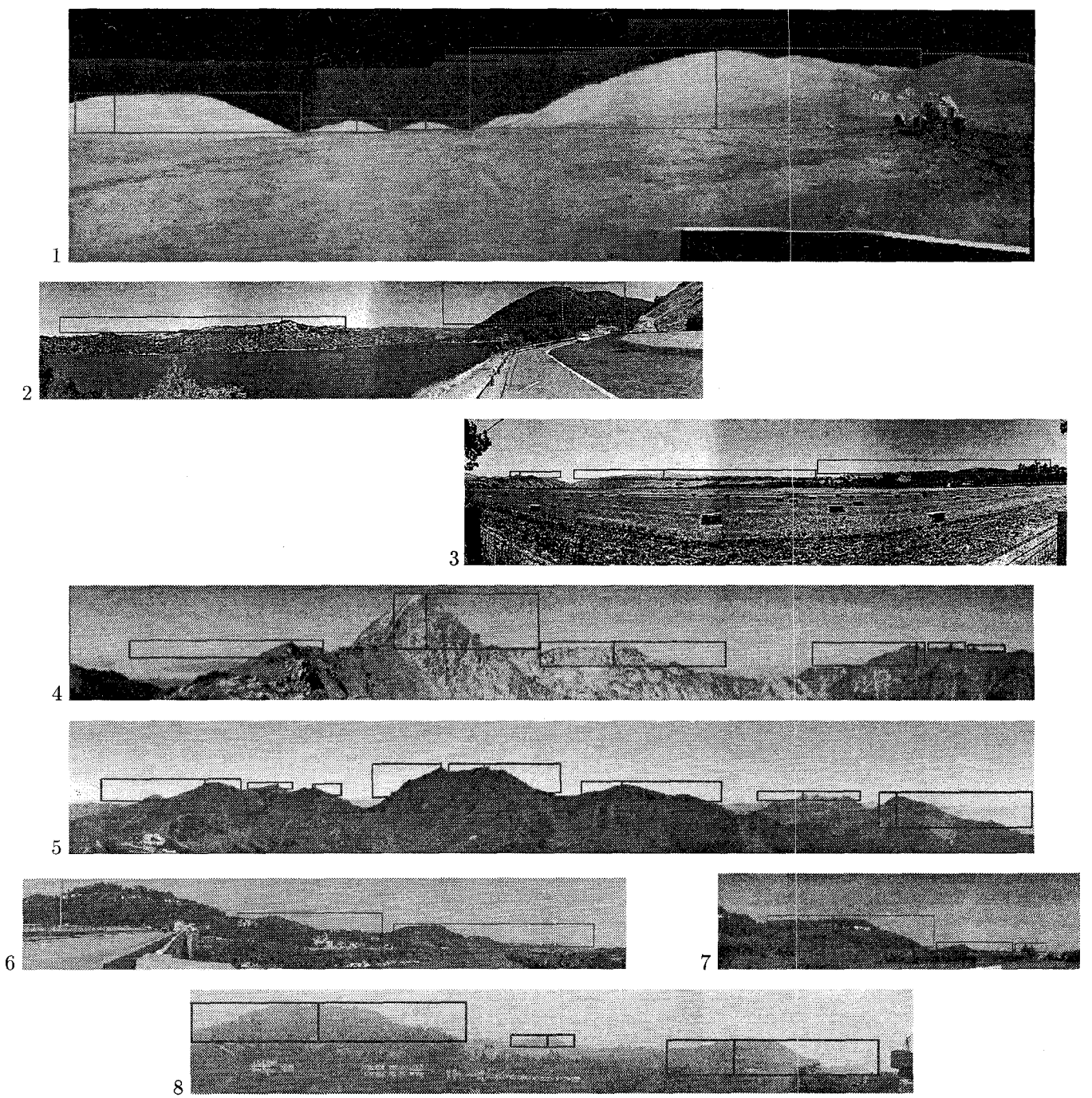

Figure 4: Peak detection results from a wide variety of scenes

task has been pursued by Talluri and Aggarwal [16]; to speed up the search, they have used a single point in the image as a measure of "goodness of fit" between image and map. In real images such a simplified figure of merit is unrealistic, since it does not take into account the noise and artifacts of real images.

We search in the position space, but we construct a figure of merit that reflects the various disturbances in the image and can also include prior knowledge about position. The key idea to speed-up our algorithm is to simplify the search procedure by pre-computing virtually all the calculations that must be performed during search. We automatically create (off-line) a table containing all the peaks that can be found in the map. We also create (off-line) a table containing all the peaks that are visible from every possible position. During search, we need only access the latter table for retrieving the index of visible peaks, access the former table for retrieving characteristics of the peaks, and compute the posterior probability for position.

Our estimation objective is to maximize the posterior probability of position given the bearings, $p(x, y|R)$. Position is discretized in intervals of 30 meters, agreeing with the discretization of Digital Elevation Maps. At first, we must specify a prior density for position, $p(x, y)$. Currently we use a uniform distribution to signify absence of prior knowledge. Secondly, we must specify likelihood of bearings (the measurements) given position, $p(R|x, y)$. This distribution is constructed by assuming independence of bearings given a particular interpretation, and Gaussian distributions for the distribution of bearings [2]. Error that exceed 18 degrees are discarded as mistakes and assumed distributed uniformly on the interval [0, 360] degrees. We construct the posterior distribution by using Bayes rule on these models. Since bearings do not need to be associated to all possible mountains, the posterior distribution can be quickly calculated. As an example, computation of the posterior probability for the more than 3×10^4 positions in the Pittsburgh map of Figure 2 takes 3 seconds running in an Impact SGI workstation.

4 Experiments

In the current version of the system, images are obtained from a customized platform featuring a camera, an electronic compass, and a differential GPS system, all mounted on a tripod. The compass provides absolute orientation measurements; once calibrated, accuracy is ± 0.5 degree. A differential GPS device obtains ground-truth measurements of position, with accuracy of 3 meters. The camera is calibrated [12], and errors of ± 2 degrees are introduced by calibration inaccuracies. The images are stored on tape and played back when testing the system; every image is tagged with six dimensional ground-truth.

We have run the estimation procedure with data obtained in Pittsburgh using sequences of images as in the last two panoramas of Figure 4. Two mountain peaks are detected in the first panorama, three in the second. The bearings, in degrees, are $\{177.6, 131.1, 92.4, 102.3, 65.3\}$. The system estimates position with an error of 87 meters. We have also run the system on data from Dromedary Peak, Utah; the panoramas are shown in Figure 4. The five bearings detected automatically are (in degrees): $\{224.5, 198.8, 163.7, 107.2, 100.8, 96.3\}$. A square area of 6km by 6km was used the Dromedary Peak quadrangle, to make the test similar to the Pittsburgh tests. The system estimated position with accuracy of 95 meters, which can be compared to the 71,700m^2 obtained by Thompson [17]. We benefit greatly from our reliance on all available features, large and small, present in the image.

5 Conclusion

We have presented a quantitative, feature-based approach for pose estimation from outdoor visual information. Our objective is the construction of an interface for rover teleoperation that can intelligently process rover imagery and help human operators. We presented results with real data which demonstrate an improvement over the state of the art on outdoor position estimation.

The superiority of our implementation, as compared to others, stems from two factors. First, we allow several mountain features, large and small, nearby and far away, to be used by the estimator. To manage the complexity created by this rich set of measurements, we have to impose some quantitative structure; our posterior distributions give this structure. Second, we have a fast, efficient implementation of the pre-compilation stage, where all possible visibility relationships are calculated and stored. Both factors are responsible for the accuracy and time performance of our method.

There are several aspects of the work call for further development. A stream of images must be presented to the user and processed by the system, and results of position estimation must be overlayed on the images so they can be assimilated by the operator. Such achievements will make it easier to remotely drive a

rover in a wide variety of environments, with a particular impact on lunar missions.

Acknowledgements

This research is supported in part by NASA under Grant NAGW-1175. Fabio Cozman was supported under a scholarship from CNPq, Brazil.

We thank Prof. W. Thompson for his assistance in the processing of the Dromedary Peak quadrangle data, Yoichi Sato for his help with the mosaicing algorithm, and Andrew Johnston and Jim Zimbelman from the Museum of Air and Space in Washington DC (USA), for their invaluable assistance in providing lunar data, images and maps.

References

[1] W. Aviles, T. Hughes, H. Everett, A. Umeda, S. Martin, A. Koyamatsu, M. Solorzano, R. Laird, and S. McArthur. Issues in mobile robotics: The unmanned ground vehicle program teleoperated vehicle (TOV). *Proc. of Mobile Robots SPIE*, V:587–597, November 1990.

[2] F. G. Cozman and E. Krotkov. Position estimation from outdoor visual landmarks for teleoperation of lunar rovers. *Proc. Third IEEE Workshop on Applications of Computer Vision*, December 1996.

[3] S. Feiner, B. Macintyre, and D. Seligmann. Knowledge-based augmented reality. *Communications of the ACM*, 36(7):53–62, 1995.

[4] M. Fishler and R. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, June 1981.

[5] G. Heiken, D. Vaniman, and B. French, editors. *Lunar Sourcebook*. Cambridge University Press, 1991.

[6] G. Kress and H. Almaula. Sensorimotor requirements for teleoperation. Technical Report R-6279, Corporate Technology Center, FMC Corporation, Santa Clara, CA, December 1988.

[7] E. Krotkov, J. Bares, L. Katragadda, R. Simmons, and R. Whittaker. Lunar rover technology demonstrations with Dante and Ratler. In *Proc. Intl. Symp. Artificial Intelligence, Robotics, and Automation for Space*, Jet Propulsion Laboratory, Pasadena, California, October 1994.

[8] T. Levitt, D. Lawton, D. Chelberg, and P. Nelson. Qualitative navigation. In *Proc. Darpa Image Understanding Workshop*, pages 319–465, 1987.

[9] D. McGovern. Experiences and results in teleoperation of land vehicles. Technical Report SAND90-0299, Sandia National Laboratories, Albuquerque, NM, April 1990.

[10] NASA. Apollo 17 landing area topophotomap. sheet 43D1S1(50), 1972.

[11] R. Newman. Time lag considerations in operator control of lunar vehicles from Earth. In C. Cummings and H. Lawrence, editors, *Technology of Lunar Exploration*. Academic Press, 1962.

[12] L. Robert. Camera calibration without feature extraction. *Proc. Intl. Conf. Pattern Recognition*, 1994.

[13] F. Stein and G. Medioni. Map-based localization using the panoramic horizon. In *Proc. IEEE Intl. Conf. Robotics and Automation*, pages 2631–2637, Nice, France, May 1992.

[14] K. Sutherland and W. Thompson. Inexact navigation. In *Proc. of the IEEE Int. Conference on Robotics and Automation*, pages 1–7, Scottsdale, AZ, May 1993.

[15] R. Szeliski. Image mosaicing for tele-reality applications. Technical Report CRL94/2, DEC Cambridge Research Lab, May 1994.

[16] R. Talluri and J. Aggarwal. Position estimation for an autonomous mobile robot in an outdoor environment. *IEEE Trans. Robotics and Automation*, 8(5):573–584, October 1992.

[17] W. Thompson, T. Henderson, T. Colvin, L. Dick, and C. Valiquette. Vision-based localization. In *DARPA Image Understanding Workshop*, pages 491–498, Maryland, April 1993.

[18] W. Thompson and H. Pick Jr. Vision-based navigation. In *DARPA Image Understanding Workshop*, pages 149–152, January 1992.

[19] J. Wertz. *Spacecraft Attitude Determination and Control*, volume 73 of *Astrophysics and space science library*. Reidel, Boston, 1978.