Short Papers

The Effect of Monocular Target Blur on Simulated Telerobotic Manipulation

Andrew Liu and Lawrence Stark

Abstract—**Proper presentation of visual information enables efficient teleoperation. We simulated three types of telerobotic tasks that require information about the spatial position of objects. Our results indicate that, for trained subjects, monocular target blur can reduce the advantage of using stereo displays. This is similar to the results of psychophysical experiments examining the effect of blur on stereoacuity, likely to be required for the execution of our tasks. This suggests that other psychophysical experimental results could be used to predict operator performance for other telerobotic tasks.**

I. Introduction

Telerobotic manipulation requires an interface between the human operator and the telemanipulator that quickly and accurately provides information about the remote environment [1]. Visual interfaces are most commonly used because of their simplicity and cost effectiveness. Two-dimensional or monoscopic displays rely on depth cues such as perspective, occlusion, and shading to convey the percept of depth to the operator. Stereoscopic displays are based on another powerful depth cue called binocular disparity. Recent studies of telemanipulation have shown that human performance for "typical" telerobotic tasks is better with stereo displays than with monoscopic displays, although on-the-screen enhancements can compensate for the difference [2]-[4]. The use of stereoscopic displays might also reduce mental fatigue and enhance awareness because of better correspondence with other depth cues that would be perceived in the remote environment [5].

The images seen by the human operator might be degraded by factors such as transmission noise, bandwidth limitations, or changes in the scene illumination. Pepper *et al.* have demonstrated the advantage of using stereoscopic displays when image contrast is reduced [4]. This result suggests that binocular disparity plays an important and robust role in the operator's ability to interpret the remote environment. However, other visual degradations may affect binocular displays more adversely than monocular displays, especially if the degradations are present in only one eye. We have found no other papers that investigate the effects of other image degradations such as target blur or image noise on the performance of telerobotic tasks.

Ogle was the first to describe the reduction in stereoacuity when one or both images were blurred with spectacle lenses [6]. It was also found that the blurring of one image reduced stereoacuity more than blurring both images (see Fig. 3(a) given below). The solid plot symbols represent data from binocular blur and the open symbols represent monocular blur data. Subsequent experiments have shown that convex and concave lenses reduce stereoacuity to similar levels [7]-[9]. On the other hand, experiments using ground glass to blur "stereo-vernier" targets did not reveal any reduction in stereoacuity [11]. Stereoacuity was measured using either a Howard-Dolman test [7]-[9] or line stereograms [10], [11]. Blur has also been described in the spatial frequency domain as a loss of the high-frequency components in an image. The effects of spatial frequency filtering on human stereoacuity also have been investigated [10]. As expected, it was found that filtering any components in the range of spatial frequencies of an image resulted in a loss of stereoacuity. Interestingly, high-pass filtering was found to be much more detrimental to stereoacuity than low-pass filtering. This is similar to the results of experiments using random-dot stereograms [12]. To determine if the reduction in contrast due to blur caused the loss in stereoacuity, contrast was enhanced after blurring. Stereoacuity improved only slightly, which suggests that the spatial frequency content of the image is important to the detection of disparity.

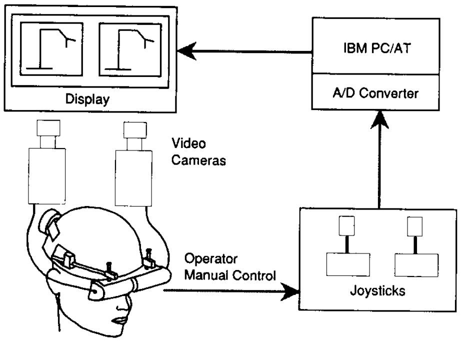

Fig. 1. Overview of the simulation of telerobotic tasks.

This short paper presents the results of our investigation into the effect of target blur on operator performance of simulated telerobotic tasks. Target blurring could occur as a result of a bandwidth limitation on the video signal from the remote site. By using spectacle lens blur to degrade the stereo images, we can correlate task performance with the degree of visual degradation. If stereoacuity is an important cue for effective telemanipulation, then the changes in operator performance should follow the results of psychophysical experiments testing the effect of target blur on stereoacuity, or the ability to discriminate between two points in depth. Thus, these and other psychophysical results might be useful for predicting changes in telerobotic task performance.

II. Methods

Two telerobotic simulators consisting of a helmet-mounted stereoscopic display, a graphic display system, joysticks, and a microcomputer was arranged in the experimental setup described below and elsewhere [4], [5], [13] (Fig. 1).

A. Experimental Apparatus and Setup

In one simulator, a Hewlett-Packard 1345A vector display, was used for the real-time dynamic display. It has 2048 × 2048 resolu-

---

Manuscript received September 12, 1989; revised November 16, 1990. This work was supported in part by the NASA Ames Research Center, Dr. S. Ellis, Technical Monitor, and by the Jet Propulsion Laboratory, California Institute of Technology, Dr. A. Bejczy, Technical Manager. A portion of this work was presented at the 7th University of California at Davis Bioengineering Symposium, Davis, CA, April 27-28, 1989.

The authors are with the Bioengineering Graduate Group and Telerobotics Unit, University of California, Berkeley, CA 94720.

IEEE Log Number 9142995.

1042-296X/91/0600-0372$01.00 © 1991 IEEE

tion and a drawing speed of 8194 cm of vectors at a 60-Hz refresh rate. In the second simulator, a Silicon Graphics IRIS 4D/120GTX with a 1024 × 1280 resolution monitor, was used to render the scene. Two video cameras (Panasonic WV-1410 CCTV or Sanyo VM-10) were focused on the screen of the display, each camera aimed at one of the images of the stereo pair. The output of each camera went to a small helment-mounted viewfinder (Sony VF-208). The two viewfinders could be adjusted independently such that the subject could easily fuse the images and achieve the stereoscopic effect.

For the pick-and-place and three-axis tracking tasks, the operator used the two displacement joysticks to control the position of the end-effector of a three-degrees-of-freedom manipulator. The position in robot base Cartesian coordinates was determined by three axes of the two joysticks. For the axle manipulation task, the joysticks controlled the translation and rotation of the axle. A switch on the joysticks enabled the operator to change between the two modes of control. The joystick outputs were connected to a 12-bit analog-to-digital converter in a PC/AT. The PC/AT performed the necessary computations for the stereoscopic display, manipulator, and/or target motions and measured the task completion time or stored the tracking response data. Because the scene in the axle manipulation task was much more complicated, the PC/AT transmitted the joystick data via a serial connection to the IRIS, which performed the rendering and data collection.

B. Description of Simulated Tasks

Three types of simulated telerobotic tasks were performed: a pick-and-place task, a three-axis tracking task, and an axle manipulation task. The subjects were four graduate students with normal stereo vision who were very familiar with the tasks and apparatus. Familiarity with the tasks and equipment was essential to simulate the fact that trained operators would probably be performing most telerobotic tasks. To eliminate any learning effects, all subjects were trained until their performance reached an asymptotic steady state.

The pick-and-place task required the operator to pick up four randomly placed targets in a certain order and place them in their respective boxes. The targets were marked by four dots and put into boxes located on the right side of the work space floor. The disparity between the two images was the only cue to the position of the targets. When the operator had moved the end effector to the same position as the target, the computer would beep to indicate that the target had been successfully grasped. Then the operator would move the end-effector to the box on the floor plane and the computer would beep again when the target was placed in the box and released. Performance was measured by the time required to pick up each target. One complete trial consisted of five screens of four targets for a total of 20 targets. Five target files were used. The first target file consisted of randomly placed targets in the work space. The remaining four files had all targets constrained to one of four subregions: top, left, right, and middle. These regions were created by subdividing the work space into 64 equally sized cubes and selecting certain cubes for each subregion.

In the three-axis tracking task, the operator was asked to track the path of a randomly moving target as closely as possible. In this case, tracking performance was measured by the rms error, normalized to the rms value of the target trajectory. Hence, if the operator were to track the input exactly, the rms error would be zero, and if the operator response was constant and equal to the mean of the signal (i.e., the operator did not move the cursor), then the rms error would be one. The target was marked by a small diamond at full brightness, and the cursor was marked by a small cross at half brightness. Because of the small size of the target and cursor, neither was displayed in perspective. The target trajectory was precomputed and comprised of three independent single-axis trajectories. Each single axis trajectory was the summation of eight sinusoids of various magnitudes and frequencies. The amplitude spectrum was made to be similar to a low-pass filter with a cutoff frequency of 0.1 Hz. By altering the relative phase of the frequency components, we created two additional input trajectories. Each trial lasted for 1 min, but the first 10 s constituted a "warm-up" period. The position of the joystick was sampled at a frequency of 40.96 Hz.

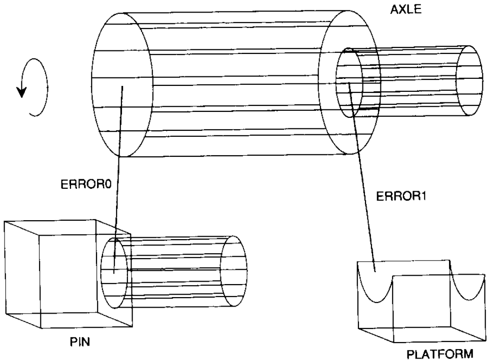

Fig. 2. Example of the objects in the axle manipulation task.

The axle manipulation task was designed to be a more realistic task for testing operator performance. As a result, the objects in the scene were more complicated and presented more evident perspective depth cues, i.e., foreshortening. The subject was required to use the joysticks to maneuver a long axle to a randomly determined position in space, designated by two end pieces (Fig. 2). The smaller end of the axle was supposed to rest in the notch of the platform while the larger end was to be centered over the large pin. Performance was measured with three parameters: the task completion time, the distance error at the platform, and the distance error at the pin. Two subjects performed a total of 20 trials for each viewing or blur condition in two sessions of 10 trials. There was no feedback to indicate to the subject how well he or she had performed during the task.

C. Experimental Procedures

In our experiments, we investigated the effect of monocular target blur on the performance of the three tasks. We presented the simulated scenes in three ways: bioptic (binocular without disparity), ideal stereoscopic, and blurred stereoscopic. The bioptic and ideal stereoscopic conditions were the control cases in our study. The bioptic case only presented the monocular depth cues, so the subject's performance should have been relatively poor, i.e., slow completion time for the pick-and-place task or large error for the tracking task. The ideal stereoscopic case had the disparity depth cue present so that performance should have been better, i.e., faster or more accurate. Monocular blur was used to ensure a large range of reduction in stereoacuity. The blur was induced by placing a lens (from +2D to +9D) approximately 25 mm in front of the camera focused on the right image of the stereo pair.

III. RESULTS

The results of the pick-and-place experiments using random target locations (circle) and constrained locations (hatched square) are plotted as the mean completion time (ordinate) versus blur in diopters (abscissa) [Fig. 3(b)]. The mean completion time is the

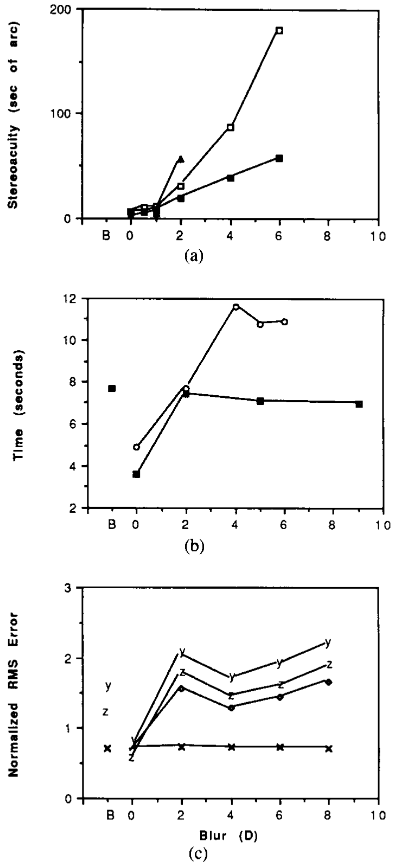

Fig. 3. Results of experiments with monocular target blur. (a) Stereoacuity versus blur for standard psychophysical tasks from Lit [7] (square) and Westheimer and McKee [10] (triangle) experiments. Solid squares indicate data for binocular blur; open squares and triangles indicate monocular blur. (b) Mean completion time versus blur for pick-and-place tasks with random target (circle) and constrained target (hatched square) distributions. (c) Normalized rms error versus blur for tracking tasks (x—horizontal, y—vertical, z—depth, ◇—rms error in three dimensions). Bioptic display (B).

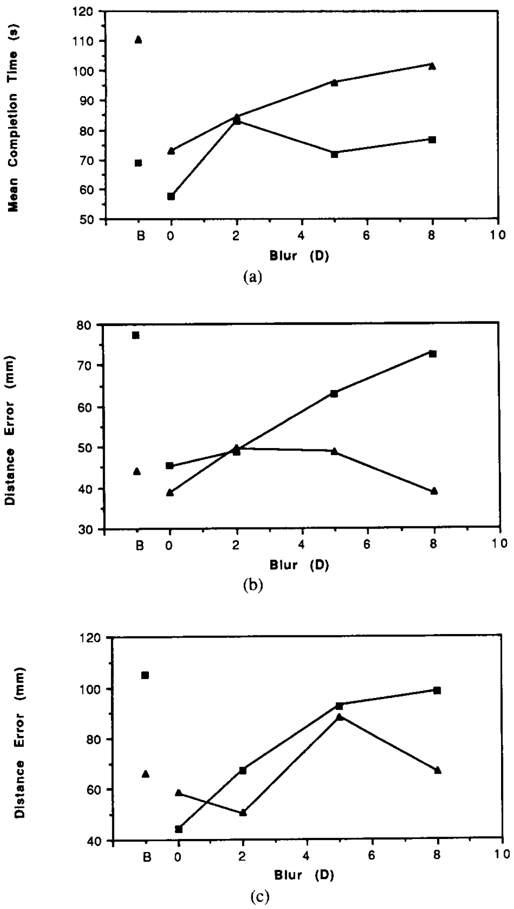

Fig. 4. Results of the axle manipulation task. The solid triangles are data from Subject 1 and the solid squares are the performance of Subject 2. (a) Performance measured by the mean completion time versus blur. (b) Distance error at the pin versus blur. (c) Distance error at the platform versus blur. Bioptic display (B).

average of the 20 pick-and-place completion times in one trial. The data from the constrained location experiments consist of the average of seven trials of 20 targets whereas the random location data are from one trial of 20 targets.

Results of the three-axis tracking experiment are expressed as normalized rms error (ordinate) versus blur in diopters (abscissa) [Fig. 3(c)]. We have plotted the rms error along each axis as well as the overall rms error. The values of the rms error along the three primary axes are shown by the plot symbols x, y, and z. The overall rms error is shown by the solid triangles.

The performance of the axle manipulation was measured by both the mean completion time [Fig. 4(a)] and the mean end position error, err0 [Fig. 4(b)] or err1 [Fig. 4(c)]. The means are calculated from 20 trials at each viewing condition. Subject 1 is shown by the solid triangles, and Subject 2 is shown by the solid squares.

We used analysis-of-variance methods to check if the trends in the data were significantly different from each other [19]. For the pick-and-place task, the mean completion times for each level of blur were found to be statistically different ($p < 0.001$) from the ideal stereoscopic case, whereas they did not differ from the bioptic case. Similar results were found for the rms errors. In the axle manipulation task, the mean completion times for 2D and 8D blur were significantly different from the ideal stereoscopic case ($p < 0.05$). However, the mean distance errors at the pin were not different while the errors at the platform were significantly different at 5D and 8D blur ($p < 0.05$).

IV. DISCUSSION

The results of pick-and-place experiments with randomly placed targets show the expected behavior: performance worsens as blur increases. However, they do not provide conclusive evidence because of the large deviation in completion times. Both of the subjects noted that certain target positions required more time to pick up. This was due to the boxes on the right side of the grid that infringed upon targets along the same line of sight, hindering the subject's depth judgement and task performance. The disparity of objects located in the periphery might have been affected by small misalignments of the viewfinders of the stereo display, which could lead to erroneous depth judgements.

The results of our second pick-and-place experiment with constrained targets also follow the pattern for stereoacuity. Task performance decreased to the performance level of bioptic viewing after 2D of blur, roughly a twofold increase in the completion time. The psychophysical results indicate that stereoacuity becomes 3-15 times worse with the addition of monocular blur [7]-[10].

The axle manipulation experiment was designed to present a more realistic task and scene to the subject. For example, since the objects are not just dots, their sizes change as they move in depth and the foreshortening of lines becomes more evident. Therefore, the contribution of the monocular cues can be more readily evaluated. The results of this experiment suggest that the presence of the monocular cues may improve performance under slightly blurred conditions. In general, the performance under moderate levels of blur (2D) is not as poor as the bioptic performance or performance under high levels of blur (> 5D). The cues do not entirely compensate for the effect of blur since performance with 8D blur is approximately the same as for the bioptic case. There is still another possible explanation for the improved performance at low levels of blur. The increased complexity of the scene resulted in more object edges where disparity could be computed. This could help the subject to see the objects in depth despite the blurring of one eye.

Our result suggests that care should be taken when setting up a stereo display for telerobotics. Even moderate reductions in stereoacuity can be enough to reduce or even eliminate the advantage of using stereo displays. This would explain how early experiment with stereoscopic displays did not show any advantage of stereo displays over monocular displays. These early studies were probably not careful in setting up the stereo displays, so that factors such as vertical misalignment or improper camera configuration degraded stereoacuity enough to reduce task performance as suggested by Pepper *et al.* [4].

The results of the three-axis tracking experiments exhibited similar behavior in overall rms error (rms xyz) and in depth (rms z). As expected, the error in the horizontal direction (rms x) did not change because horizontal motion in the picture plane can still be seen by the unblurred eye. However, the error in the vertical direction (rms y) became larger when the blur was added. By examining data from previous tracking experiments using monocular displays [14], we found that the azimuth and elevation angles of display affect the magnitude of error in the horizontal or vertical directions. In both sets of tracking experiments, the target and cursor did not change size as they moved in depth. Therefore, under the blurred viewing conditions when stereo was lost, the operator interpreted the movement in depth as movements in the display plane. This looming cue has been suggested to be a better cue to motion-in-depth than changing disparity [15]. Also, it has been demonstrated that movements in the display plane are more readily seen than movements in depth [16]. Thus, while disparity is a powerful cue for telemanipulation, monocular cues may still be important, especially under degraded viewing conditions.

The small number of subjects does cast some uncertainty over the generalizability of the results. However, the fact that the trends of worsening performance for increased blur holds for three very different tasks gives evidence to the conclusions we have drawn. Again, the subjects were practiced, as opposed to naive subjects, but this was done to simulate our assumption that the use of well-trained operators would be most common. We are hoping to perform other types of tasks that will further strengthen the generalizability of our results.

The major difference between our simulated telerobotic tasks and the stereoacuity tests is that our simulations contain moving stimuli. Recently, Zinn and Solomon suggested that dynamic stereoacuity could not be predicted from static stereoacuity [17]. They devised a system in which a cart, carrying four stereograms of circles, moved toward a subject. The subjects were asked to determine which circle was closest to them as quickly as possible. They found little correlation between the subjects' tested static stereoacuity and their dynamic stereoacuity performance. Our results seem to disagree with this hypothesis since the degradation of performance was consistent with the results of the static stereoacuity testing. It has also been shown that stereoscopic depth resolution is independent of precise placement of targets on the retina of either eye [18]. To investigate this question further, we hope to perform future tracking experiments comparing performance given the disparity cue versus the looming cue.

V. CONCLUSION

It has been demonstrated that refractive errors in the helmet-mounted stereo display system can affect performance in three types of telerobotic tasks. The results of both sets of experiments indicated that monocular target blur of two diopters or more degraded stereo display performance to the level of monocular displays. This indicates that moderate levels of visual degradation that affect the operator's stereoacuity may eliminate the performance advantage of stereo displays. Our experiments also suggest that the many psychophysical studies of human vision might be a very useful starting point for predicting performance and visual requirements for teleoperation. However, each aspect of vision should also be tested in terms of operator performance carrying out actual tasks.

ACKNOWLEDGMENT

The authors would like to thank Dr. G. Tharp for his technical assistance and helpful discussions.

REFERENCES

[1] A. K. Bejczy, "Man-machine interface issues in space telerobotics: A JPL research and development program" in *Proc. Workshop Space Telerobotics*, vol. 2, July 1987, pp. 361-369.

[2] W. S. Kim, S. R. Ellis, M. E. Tyler, B. Hannaford, and L. W. Stark, "Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-17, pp. 61-72, Jan. 1987.

[3] W. S. Kim, F. Tendick, and L. W. Stark, "Visual enhancements in pick-and-place tasks: Human operators controlling a simulated cylindrical manipulator," *IEEE J. Robotics Automat.*, vol. RA-3, pp. 418-425, Oct. 1987.

[4] R. L. Pepper, D. C. Smith, and R. E. Cole, "Stereo TV improves operator performance under degraded visibility conditions," *Opt. Eng.*, vol. 20, pp. 579-585, July 1981.

[5] P. Tolin, "Maintaining the three-dimensional illusion," *Information Display*, vol. 3, pp. 10-12, Dec. 1987.

[6] K. N. Ogle, "Spatial localization through binocular vision," in *The Eye*. London: Academic, 1962, pp. 289-290.

[7] A. Lit, "Presentation of experimental data," *J. Amer. Optom. Assoc.*, vol. 39, pp. 1098-1099. Dec. 1968.

[8] J. Ong and W. S. Burley, "Effect of induced anisometropia on depth perception," *Amer. J. Optom. Arch. Amer. Acad. Optom.*, vol. 49, pp. 333-335, Apr. 1972.

[9] W. L. Larson and A. Lachance, "Stereoscopic acuity with induced refractive errors," *Amer. J. Opto. Physiol. Opt.*, vol. 60, pp. 509-513, June 1983.

[10] G. Westheimer and S. P. McKee, "Stereoscopic acuity with defocused and spatially filtered retinal images," *J. Opt. Soc. Amer.*, vol. 70, pp. 772-778, July 1980.

[11] G. Stigmar, "Blurred visual stimuli II. The effect of blurred visual stimuli on vernier and stereo acuity," *Acta Opth.*, vol. 49, pp. 364-379, May 1971.

[12] B. Julesz, *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press, 1971.

[13] W. S. Kim, A. Liu, K. Matsunaga, and L. Stark, "A helmet mounted display for telerobotics," in *Proc. COMPCON'88* (San Francisco), Mar. 1988.

[14] L. Stark *et al.*, "Telerobotics: Display, control, and communication problems," *IEEE J. Robotics Automat.*, vol. RA-3, pp. 67-75, Feb. 1987.

[15] D. Regan, K. Beverly, and M. Cynader, "The visual perception of motion in depth," *Sci. Amer.*, vol. 241, p. 136-151, July 1979.

[16] C. W. Tyler, "Stereoscopic depth movement: Two eyes less sensitive than one," *Science*, vol. 174, pp. 958-961, Nov. 1971.

[17] W. J. Zinn and H. Solomon, "A comparison of static and dynamic stereoacuity," *J. Amer. Optom. Assoc.*, vol. 56, pp. 712-715, Sept. 1985.

[18] G. Westheimer and S. P. McKee, "Stereoscopic acuity for moving retinal images," *J. Opt. Soc. Amer.*, vol. 68, pp. 450-455, Apr. 1978.

[19] G. Perlman, *ANOVA(1CSL)-UNIX Programmer's Manual*, 1980.

Mobile Robot Localization by Tracking Geometric Beacons

John J. Leonard and Hugh F. Durrant-Whyte

Abstract—**This short paper presents the application of the extended Kalman filter (EKF) to the problem of mobile robot navigation in a known environment. We have developed an algorithm for model-based localization that relies on the concept of a *geometric beacon*—a naturally occurring environment feature that can be reliably observed in successive sensor measurements and can be accurately described in terms of a concise geometric parameterization. The algorithm is based on an EKF that utilizes matches between observed geometric beacons and an *a priori* map of beacon locations. We describe two implementations of this navigation algorithm, both of which use sonar. The first implementation uses a simple vehicle with point kinematics equipped with a single rotating sonar. The second implementation uses a "Robuter" mobile robot and employs six static sonar transducers to provide localization information while the vehicle moves at typical speeds of 30 cm/s.**

I. THE NAVIGATION PROBLEM

Stated most simply, the problem of navigation can be summarized into answering the following three questions: "where am I?", "where am I going?" and "how should I get there?". The first question is one of localization; how can I work out where I am in a given environment, based on what I can see and what I have previously been told? The second and third questions are essentially those of specifying a goal and being able to plan a path that results in achieving this goal. We are principally concerned with the first, localization, question, and maintain that finding a robust and reliable solution to this problem is an essential precursor to answering the remaining two questions.

The problem of position determination has been of considerable interest over the last 4000 years. The basic process of distance measurement, correlation, and triangulation was known to the Phoenicians,[1] who successfully managed to build and maintain quite accurate maps of the Mediterranean area. Today, navigation is a well-understood quantitative science, used routinely in maritime and aviation applications [22]. Given this, the question must be asked as to why robust and reliable autonomous mobile robot navigation remains such a difficult problem. In our view, the reason for this is clear; it is not the navigation process *per se* that is a problem, it is the reliable acquisition or extraction of information about navigation beacons, from sensor information, and the automatic correlation or correspondence of these with some navigation map that makes the *autonomous* navigation problem so difficult.

Implementing a navigation system that uses artificial beacons together with sensors that provide accurate and reliable measurements of beacon location is a straight forward procedure used by many commercial mobile robots today. For example, the GEC-Caterpillar AGV [3] uses a rotating laser to locate itself with respect to a set of bar codes that are fixed at known locations through the AGV's environment. More recently, the TRC Corporation has developed a system for localization that uses retro-reflective strips and ceiling lights as beacons that are observed by vision and active infrared sensors. Our goal for a *competence of localization* is to use the naturally occurring structure of typical indoor environments to achieve comparable performance to artificial beacon systems without modifying the environment.

We have developed a system in which the basic localization algorithm is formalized as a vehicle-tracking problem, employing an extended Kalman filter (EKF) to match beacon observations to a navigation map to maintain an estimate of mobile robot location. Kalman filtering techniques have been used extensively in location estimation problems such as missile tracking and ship navigation [21]. There have been many notable applications of the EKF in mobile robot systems. For example, Dickmanns uses an EKF in a real-time vision system that achieves autonomous road-following at speeds over 80 km/h [7]. Ayache and Faugeras [1], Matthies and Shafer [19], and Kriegman *et al.* [14] have used the EKF for visual map building and motion estimation. These systems address a much more complex task than that considered here, as they start without an *a priori* model. However, the *motion* estimation formulation does not by itself meet our requirements for long-term autonomous *position* estimation, for despite the high accuracy with which the relative motion between frames can be estimated, uncertainty in the globally referenced vehicle position estimate must accumulate with time. Hallam has developed an undersea navigation system that maintains an absolutely referenced estimate of vehicle position in an environment comprised of moving targets and clutter [13]. Our formulation to the problem deals with the much simpler case of a static environment, but has been demonstrated to be successful with real sonar data on several different robots.

II. THE LOCALIZATION ALGORITHM

In man-made indoor environments, we model the world in terms of geometry and consider each feature of the environment to be a geometric target. A *geometric beacon* is a special type of target that can be reliably observed in successive sensor measurements and that can be accurately described in terms of a concise geometric parameterization. Hence, geometric beacons are stable, naturally occurring environment features that are useful for navigation. The idea of a generalized geometric beacon arises as the result of our earlier work in describing sensors and processing algorithms as "geometry extractors" [9], [11], allowing many different types of information to be integrated easily in a common geometric framework.

With reference to Fig. 1, we denote the position and orientation of the vehicle at time step k by the state vector $x(k) =$

Manuscript received August 30, 1989; revised August 30, 1990. This work was supported in part by ESPRIT 1560 (SKIDS) and by the SERC-ACME under Grant GRE/42419. A portion of this work was presented at the IEEE/RSJ International Conference on Intelligent Robot Systems, 1989.

The authors are with the Department of Engineering Science, University of Oxford, Oxford, United Kingdom.

IEEE Log Number 9042230.

---

[1] The Phoenicians were pre-Greek seafarers from Syria whose main claim to fame was their hypothesis that the world was flat.

1042-296X/91/0600-0376$01.00 © 1991 IEEE