ornl

ORNL/TM-11733
CESAR-90/54

OAK RIDGE NATIONAL LABORATORY

MARTIN MARIETTA

Performance of Visual and Ultrasound Sensing by an Autonomous Robot

M. Beckerman
D. L. Barnett

MANAGED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

ORNL/TM-11733
CESAR-90/54

Engineering Physics and Mathematics Division

PERFORMANCE OF VISUAL AND ULTRASOUND SENSING BY AN AUTONOMOUS ROBOT

M. Beckerman and D. L. Barnett

DATE PUBLISHED — January 1991

Prepared for the
Engineering Research Program
Office of Basic Energy Sciences
and the
Office of Nuclear Energy
Office of Technology Support Programs
U.S. Department of Energy

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
managed by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE–AC05–84OR21400

CONTENTS

LIST OF FIGURES

INDEX TERMS

Autonomous Robots

Mobile Robots

Monte Carlo Methods

Region Analysis

Reliability

Robustness

Sensor Fusion

Ultrasound Sensing

Visual Sensing

ABSTRACT

This paper presents results of an experimental study of the reliability of an autonomous mobile robot operating in an unstructured environment. Examined in the study are the principal components of the visual and ultrasound sensor systems used to guide navigation and manipulation tasks of the robot. Performance criteria are established with respect to the requirements of the integrated robotic system. Repeated measurements are done of the geometric and spatial quantities used for docking the robot at a mock-up control panel, and for locating control panel devices to be manipulated. The systematic and random components of the errors in the measured quantities are exhibited, their origins are identified, and means for their reduction are developed. We focus on refinements of visual area data using ultrasound range data, and on extraction of yaw by visual and by ultrasound methods. Monte Carlo methods are used to study the sensor fusion, and angle-dependence considerations are used to characterize the precision of the yaw measurements. Issues relating to sensor models and sensor fusion, viewed as essential strategic components of intelligent systems, are then discussed.

1. INTRODUCTION

1.1 GENERAL

During the past several years a variety of autonomous mobile robots have been built and used as testbeds for studying basic issues in sensing, actuation and design. Some of these robotic testbeds, for example, those discussed in [13], [19], [29], [30], and [32] were developed for outdoor environments; others, such as the systems described in [7], [11], and [12] were built for indoor environments. Most recently, a number behavior-driven indoor mobile robots have appeared which differ conceptually from the primarily task-driven systems mentioned above. These systems have been discussed in [1], [2], [3], [10], and [23].

The intelligent machines described in the above-cited papers operate in environments which are unstructured, or ill-defined, to some appreciable extent. The machines are intelligent insofar as they sense their surroundings, make decisions based on the acquired data, and then take some action. A primary objective of both task- and behavior-driven systems is to achieve a performance in an unstructured environment which is reliable and robust. Methods for achieving this objective, and related performance issues have been discussed in the papers cited above and in [20], [24], and [26], and a performance assessment for near-perfect machines was presented in [18].

In the present work we examined experimentally some issues pertaining to the reliability and robustness of an autonomous robotic system. Our specific goal was a modest one. We studied the ultrasound and (monocular) vision subsystems as used to guide navigation and manipulation tasks of the HERMIES-IIB robot. As is the case in any experimental science, we wished to understand the errors beginning with their decomposition into random and systematic components.

We were motivated in part by the observation that systematic errors are of particular importance in automated, or robotic, systems. They arise in these systems whenever there is insufficient information to correctly interpret the data and whenever underlying assumptions concerning the ill-defined environment are violated. Systematic errors influence the accuracy of the measured physical and geometric quantities while the random errors affect the precision to which those quantities can be determined.

1.2 HERMIES-IIB EXPERIMENTS

The HERMIES-IIB mobile robot used as the testbed was an inexpensive indoor prototype developed at the Center for Engineering Systems Advanced Research at Oak Ridge National Laboratory. The HERMIES-IIB hardware includes a wheeled platform (two drive wheels, one passive caster), a pair of Heathkit manipulator arms, on-board computers and a sensor suite. The sensor suite contains an array of ultrasound range sensors and three CCD cameras. There is an additional ultrasound sensor facing forward in the center of the robot. This range sensor and two of the CCD cameras were used in the HERMIES-IIB experiments.

Upon receiving a CCD camera image, the visual data processing subsystem produced a list of geometric properties of the various connected regions in the thresholded binary image. These properties were used by the integrated system to identify a mockup control panel, navigate to a docking zone in front of the panel, and locate devices on the panel, which were then manipulated.

The specific tasks mentioned above established the performance requirements of the visual sensor subsystem. These requirements include the precision and accuracy to be achieved, and the type of information extracted from the sensor data. The measurements done in this study were of distance, angle and yaw of the robot with respect to the control panel, and the area and height-to-width ratio of the panel devices, as derived from the list of geometric properties. The first three of these spatial quantities were used to navigate up to and dock in front of the mockup control panel. The last two were used by HERMIES-IIB at the docking zone to locate the various devices on the control panel.

In previous experiments with the HERMIES-IIB robot, we examined ultrasound-guided navigation issues. In [7] we fused ultrasound data from different robot locations, and in [8] we fused visual and ultrasound data from one location. Measurements of ultrasound ranges were done in the present study to examine whether the addition of range data provides a cost-effective means of reducing errors in the visually derived areas, and to compare values for yaw deduced from the range data to those obtained from the visual images.

This paper is organized, as follows. The experimental setup is given in Sec. 2, and the experimental procedure and the (monocular) visual data processing are described in Sec. 3. The visually-derived area and height-to-width data are presented in Secs. 4.1 and 4.2. This is followed in Sec. 4.3 with a brief description of the Monte Carlo method, and its application to the refinement of visual area data using ultrasound ranges. Distance and yaw results obtained using visual and ultrasound methods are examined in Sec. 5. Some of the results and emerging issues are discussed further in Sec. 6, and the work is concluded in Sec. 7.

2. EXPERIMENTAL SETUP

2.1 ULTRASOUND ENVIRONMENT AND SENSING

The ultrasound sensor used in the experiments was mounted in the center of the front of the robot at a height of 47 cm above the floor. The sensor contained a Polaroid industrial grade transducer which functioned as both the transmitter and receiver, and produced 50 kHz bursts, 1 msec in duration. The signals from the sensor were processed to give the time of flight for the earliest echo, thereby determining the distance to the nearest object lying either partially or wholly within the beam cone. The intrinsic resolution of the sensor, after digitization, was 3.0 cm.

The ultrasound systematic errors depend upon the beam width, or resolution, the sensing frequency, the radiated power and sensor threshold, and the environmental geometry and surface properties of the objects scanned. A detailed treatment of some of the systematic errors encountered using this type of sensor in a laboratory environment has been presented in [7]. The sensing experiments reported herein involved scan distances from roughly 45 to 180 cm. At these near distances, and for the control panel being scanned, the systematic errors were minimal.

A photo of the control panel used in the experiments is shown in Fig. 1. Three features, or surfaces, of the control panel comprise the ultrasound sensing environment. These surfaces, the smooth metal front plating, the four raised plastic buttons, and a steel lip surrounding the front plate, are visible in Fig. 1. The buttons are 3.8 cm in diameter, and when not depressed stand out 4.2 cm from the main surface. The lip is raised by 1.5 cm from the main surface, and forms an interior corner. This corner allows for ultrasound returns from the control panel at yaw angles which otherwise would lead to specular reflections.

2.2 VISUAL ENVIRONMENT AND SENSING

The CCD cameras mounted on the HERMIES-IIB robot provided 256 $\times$ 256 pixels of 8-bit grey-level intensity values. These cameras were mounted on a pan/tilt table at a height of 94 cm above the floor, and were oriented facing forward in the horizontal plane. A 4.8 mm wide-angle lens, mounted on each of the two cameras, provided a 60 deg field of view. Ambient lighting was provided by the standard overhead hanging fluorescent lamps, and by a window located some 10 m from the experimental area.

The mockup control panel contains a pair of analog meters, a large rectangular "danger" light, a pair of cylindrical slides, and four round buttons and acompanying indicator lamps. The buttons and slides were made from white plastic. The danger light and indicator lamps were red. The analog meter background color was white, and the control panel was dark metallic grey.

The physical (geometric) characteristics of the devices on the control panel are listed in Table 1. As can be seen in the Table, there were considerable variations in device size and shape. At large distances only the analog meters were large enough to be discerned. At near distances the variations in device elevation were sufficient to require the use of tilt.

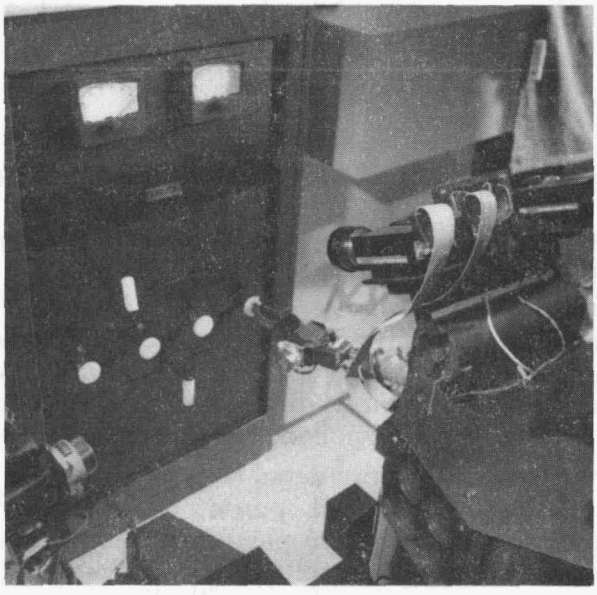

Fig. 1. Photo of the HERMIES-IIB robot performing a manipulation task from the docking zone. The CCD cameras are located on the topmost part of the robot. The central ultrasound sensor is housed in the box-like unit in the lower front of the robot. From top to bottom, the control panel devices visible are the pair of analog meters, "danger" light, upper slide, four buttons, and lower slide.

Table 1. Physical Properties of the Control Panel Devices

Device	Shape	Height/Width	Area (mm^2)	Elevation (cm)
Light	Rectangle	0.32	1720	73.0
Slide	Cylinder	2.79	1120[a]	59.2
Button	Circular	1.00	1200[b]	46.0

[a]Includes visible area of circular top surface.
[b]Includes visible area of rim of button.

3. EXPERIMENTAL PROCEDURE AND VISUAL DATA PROCESSING

3.1 CALIBRATION

The experiments were preceded by a series of calibration measurements. First, the zero (forward axis) of the sensor turret was determined to an accuracy of 1 deg. Second, the optic axis of each CCD camera was aligned. An adjustment of about 10 pixels was required in order to align the software crosshair with the physical centerline. Third, the plane of the 4.8 mm wide-angle lens was found to be misaligned with respect to that of the CCD array. This misalignment was not corrected, and has a discernable effect upon the data.

3.2 VISUAL IMAGE PROCESSING

The visual image processing consisted of converting the grey-scale image to a binary image, decomposing this binary image into connected regions, and then extracting the geometric properties of the connected regions. Techniques for doing region analyses of binary images, and extracting the quantities mentioned above, are well established in computer vision following, for example, the work described in [4], [9], [16], [17], [22] and [27]. In the present work the region analysis was done concurrently on-board the robot using a 16-node hypercube computer.

The five principal stages of low-level parallel processing are depicted in Fig. 2. The grey-scale morphology, or non-linear filtering, reduced noise while preserving edges and sizes (see, for example, [28]). The grey-scale image was then converted to a binary image. The binary morphology (dilation and erosion) ensured the stability of the grey-scale-to-binary transformation with respect to shifts or changes in size of major features.

A two-dimensional array was then produced from the binary image in which each pixel carried a label common to its connected neighbors. White regions were 4-connected, and black regions were 8-connected. A butterfly vector accumulator was used to produce a consistent set of labelled regions across nodes of the hypercube. The techniques for doing this have been described in detail elsewhere [21].

Geometric properties of moments of each region were then extracted and stored in a list. Two such lists were produced, one for black regions and one for white regions. Each entry in the list contains the quantities given in Table 2. Recall that the n^{th} moment of a discrete distribution of a variable, x, is defined as the mean of the n^{th} power of the variable, that is,

$$< x^n >= (1/N) \sum_i f_i x_i^n \tag{1}$$

In Eq. (1) f_i is the relative frequency of the i^{th} element of the distribution, x_i is the mean value of that element, and

$$N = \sum_i f_i \tag{2}$$

is the summed frequency of the distribution. These moments are related to the moments, μ_n, of the distribution about the mean by the expression

$$\mu_n = (1/N) \sum_i f_i (x_i - < x >)^n \qquad (3)$$

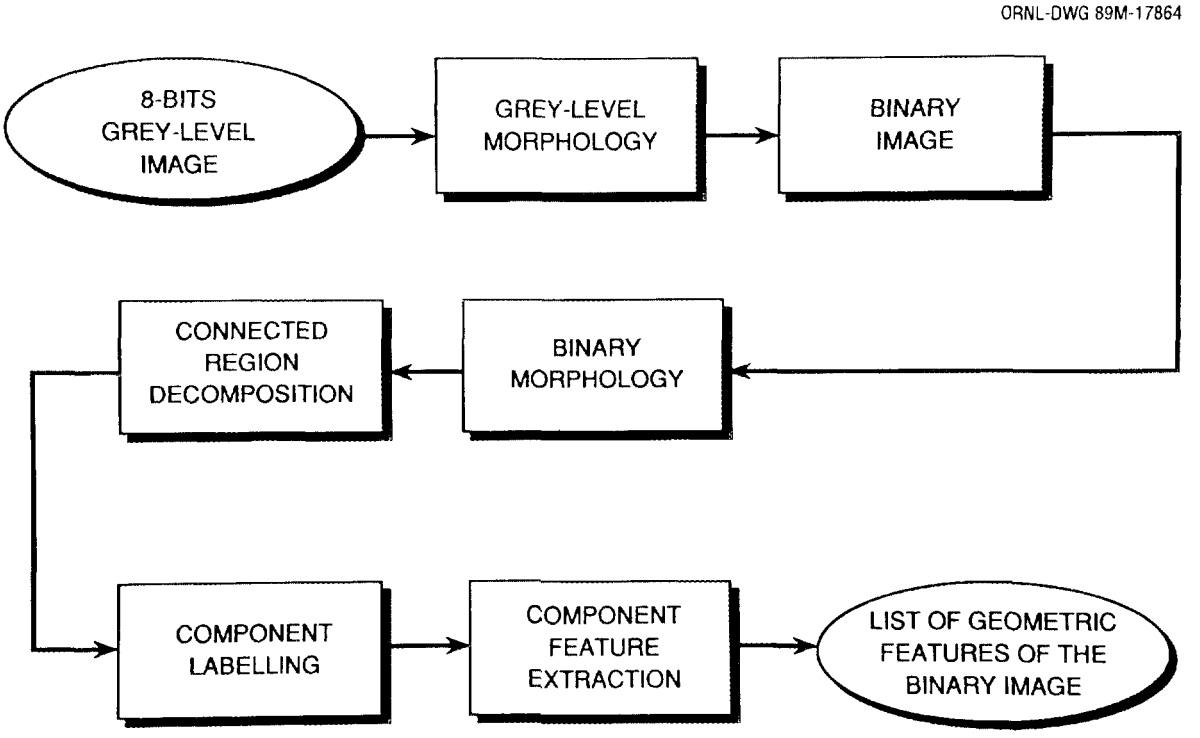

Fig. 2. Block diagram of the stages of the low level processing of the visual sensor data.

The first entry in Table 2 represents the total pixel area of the region. The second entries, the extrema, define a bounding box for the region. The third entries, the mean values, give the centers-of-gravity for the x and y coordinates of the region. The final set of entries in the Table are the root-mean-square deviations, defined as the square-root of the second moments, or variances, of the coordinate distributions. These quantities have been cast into the simple form given in Table 2. For pixel distributions there are, of course, two variables, and an additional summation over the second variable gives the frequencies, f_i.

Table 2. Geometric Properties Extracted for Each Region in the Binary Image

Property	***Definition***
Area	Cardinality
Extrema	Min x, Max x, Min y, Max y
Mean	$< x >, < y >$
RMS Deviation	$(< x^2 > - < x >^2)^{1/2}$, $(< y^2 > - < y >^2)^{1/2}$

3.3 INTEGRATED SYSTEM REQUIREMENTS

The final task performed by the HERMIES-IIB integrated system was to manipulate the slides and push the buttons on the mockup control panel. The CCD cameras were used in this task to locate both the end-effectors and the control panel devices by means of a stereo algorithm. The docking zone is defined as that region in front of the control panel from which this task was performed successfully. It was determined by the mechanical limitations of the manipulators and the sensing requirements for the particular placement of the cameras on the robot. The docking zone was found experimentally to be an area centered at a mean distance of 50 cm from the control panel, 5 cm × 8 cm in size, with a yaw, as defined in Fig. 3, of less than 2 deg.

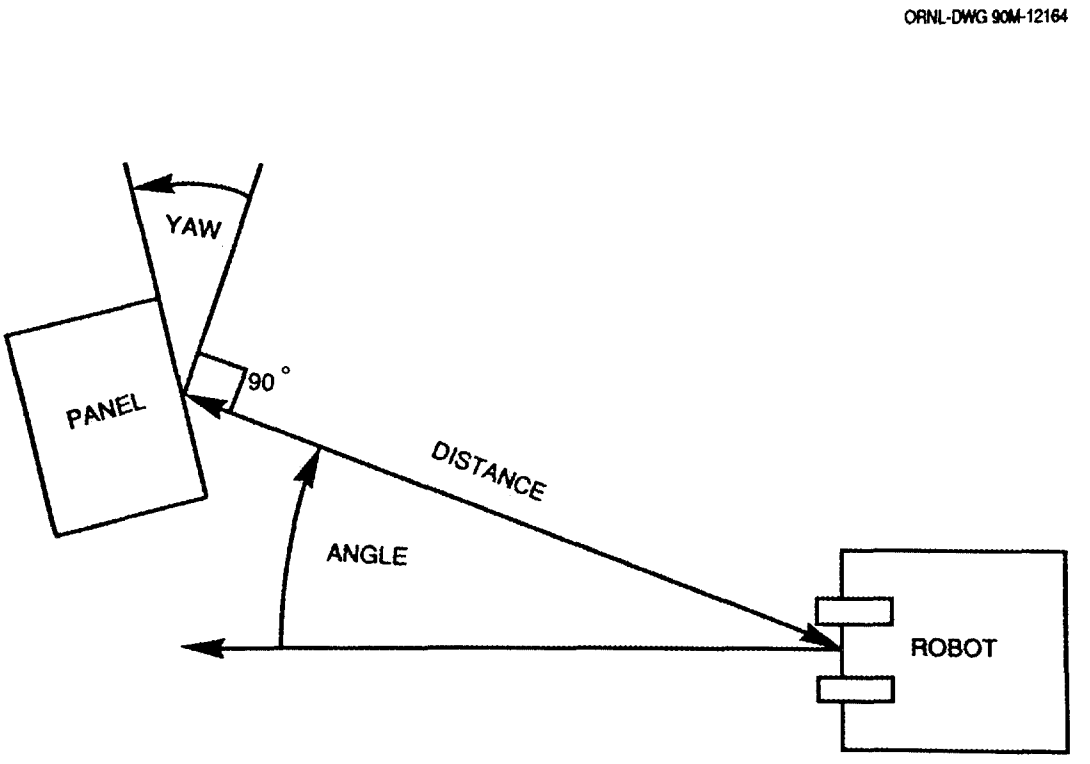

Fig. 3. Definition of distance, angle and yaw.

The manipulation tasks were proceded by the navigation tasks. The control panel was first identified in the binary image as a dark region containing a pair of light regions situated approximately side-by-side in its uper half. Distance, angle and yaw information was then used to plan a local path towards the docking zone, and eventually reaching this goal in three to five steps. The task-required operating range of these algorithms was from 45 cm to 4.5 m over the angular range for which the control panel was in the field-of-view of the cameras. The performance ranges and measurements done are summarized in Table 3.

Table 3. Summary of Visual Sensing Experiments

Sensing Task	Sensing Zone	Device	Measurement
Location	5 cm width 8 cm depth 2 deg yaw	Light Slide Buttons	Height/Width Area
Docking	45 cm min depth 4.5 m max depth 60 deg angular	Meters	Distance Angle Yaw

3.4 MEASUREMENTS

As shown in Table 3, areas and height-to-widths were measured for the danger light, four buttons and uppermost slide. These measurements were done from a uniform distribution of positions within the docking zone. At each position a CCD camera image was taken, and the desired quantities were extracted. Areas were given as the cardinality of the region, and the height-to-width ratios were determined using the values for the extrema.

Distance, angle and yaw (Fig. 3) were deduced from the observed geometric properties of the analog meters, from representative positions within the task-defined operating range. These three quantities were found using monocular vision and a standard pinhole camera model (see, for example, [6]). A priori knowledge was used of the physical dimensions of the analog meters, focal length of the lens of the CCD cameras, and the physical size of a pixel in the CCD array. Distance was calculated from the pixel area (cardinality) of the meters in the visual image. Angle was determined from the mean values of the horizontal offset of the meters from the center of the image. Yaw was the most difficult of the three quantities to determine, and its determination is essential for reliable docking. It was deduced visually from the reduction in the observed widths of the meters from the physical value. To remove the distance-dependence from the data, ratios of width to height were used. These were defined in terms of the ratios of the rms deviations.

Two sets of ultrasound range experiments were done. In the first series of measurements, ranges were found for a uniform distribution of locations within the docking zone located in front of the control panel at distances from 45 to 53 cm from the control panel. A simple representation of these data was then used to refine the visual sensor data. In the second series of measurements, range data was acquired at distances of 90 and 180 cm. The measurements were done from two positions separated from one another by a distance comparable to the width of the control panel. The difference between the ranges given by the two measurements was then used to calculate the control panel yaw.

4. AREA AND HEIGHT-TO-WIDTH DATA

In this section we describe the results of measurements of those geometric properties used to distinguish the control panel devices from one another. We start with the height-to-width ratios, then examine the area data. Finally, we explore refinements of the areas using ultrasound range data.

4.1 HEIGHT-TO-WIDTH RATIOS

Height-to-width ratios for the uppermost slide, the four buttons and the danger light are displayed in Fig. 4. In contructing this figure data for the two CCD cameras were added together. Also shown in the figure are the height-to-width ratios of the devices as listed in Table 2. Two observations can be made. First, the measured distributions are systematically shifted downward from the actual values. The shifts are pronounced for the buttons, less so for the slide and modest for the danger light. Second, the distributions are rather broad, indicative of the presence of large random variations in the measured quantities.

The systematic reductions in the measured height-to-width ratios from their actual values are due to perspective distortions. We recall from Sec 2.2 that the CCD cameras are situated on the sensor turret of the robot at a distance of 94 cm above the floor. The devices are located at various elevations on the control panel. The relative shifts in the device height-to-width ratios from their correct values are correlated with the device elevations listed in Table 2. These systematic effects are consistent with those expected of perspective distortions, and are a consequence of the less-than-optimal viewing conditions to be expected in any unstructured environment.

To understand the origins of the large widths of the distributions we examined individual values for the heights and widths. We found that the magnitudes of these quantities were small, ranging from 5 to 20 pixels. These small values together with the sensitivity of extrema to small differences in intensity produced variations in measured height and width for repeated measurements from a fixed position. Additional variations were produced when the camera positions were varied across the docking zone. Part of the spread in the frequency distributions presented in Fig. 4 can be attributed to these random effects. Other factors contributing to the spread in ratios will be discussed in the following sections, where their signatures are more easily observed.

4.2 AREAS

Area data for the three devices are presented in Fig. 5. In order to interpret these results we have included in the figure the cardinalities to be expected from the center of the docking zone, calculated using the pinhole camera model from the area data given in Table 2. Again, we observe that the frequency distributions are broad and are shifted from the expected values. The shifts in the area data are more complex than those for the height-to-width ratios. The button histogram is shifted downward while the danger light and slide distributions are shifted slightly upward.

Of the three devices types, the buttons were located furthest from the camera, and the reductions in area are consistent with those expected for perspective distortions. The increases in area for the slide and danger light are due to other, more dominant effects. If by spherical distortions we mean distortions which vanish at the center of the lens, and increase radially so that objects appear smaller

then they actually are, then this effect operates in a direction opposite to the observations. This leaves a general failure of the pinhole camera model at small distances as a possible explanation. In our case the distance to focal length ratio is 100, or less. A blurring, or defocussing, coupled to the subsequent digitization to a binary image, would serve to increase the observed areas of the devices. The magnitude of the effect needed to account for the data is roughly a single pixel in the extrema. Additional evidence for this effect will be presented in Sec. 5.

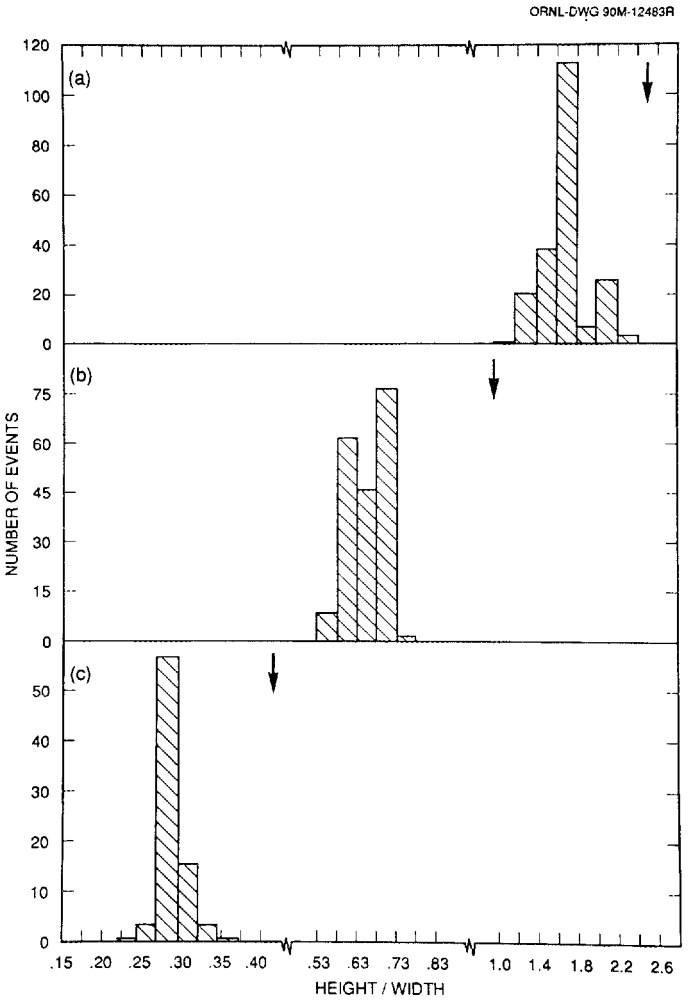

Fig. 4. Frequency distributions height/width ratios. (a) slide; (b) buttons, (c) light.

Unlike the hight-to-width data, the area results are influenced by small variations in distance. This dependence becomes visible when the histograms are rebinned into distributions corresponding to distance regimes within the docking zone. The results of a decomposition of area data for the danger light are diplayed in Fig. 6. In the figure we find that the histograms are centered at progressively larger values of the area as the distances are increased. The arrows included in the figure represent the expected inverse square dependence, and are in excellent agreement with the observed variations.

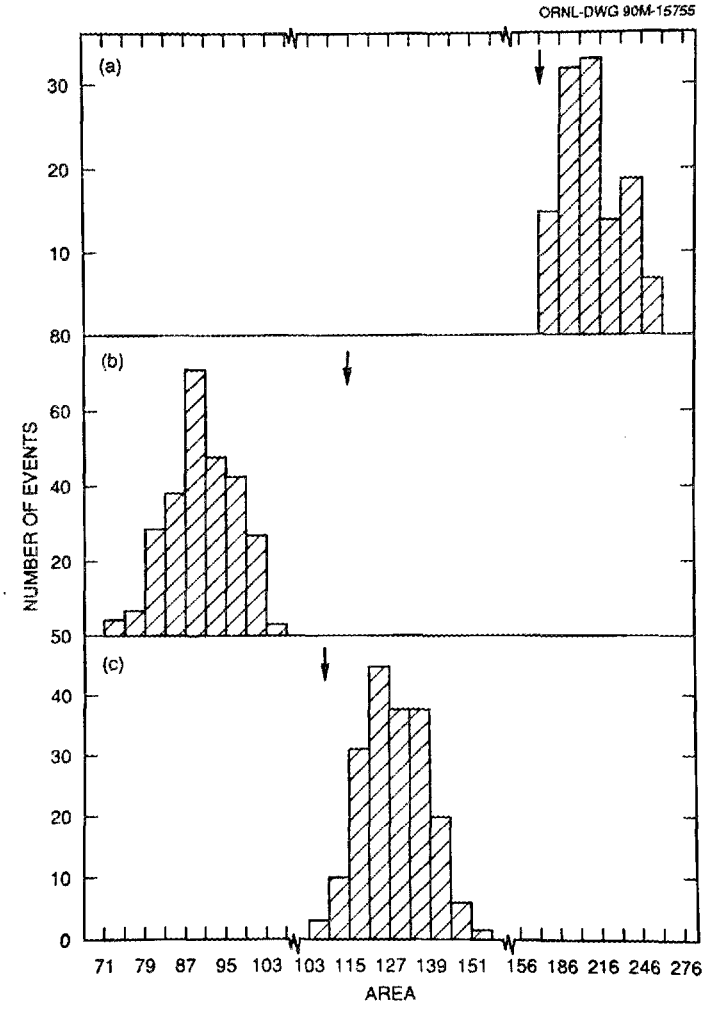

Fig. 5. Frequency distributions of areas (cardinalities). (a) light; (b) buttons; (c) slide.

4.3 MONTE CARLO METHOD AND SENSOR FUSION

It is clear that if we had independent information on the distance of the robot from the control panel, we could perhaps sharpen the area results. The ultrasound sensors provide such a measure of distance. In this section we present results of fusing visual area and ultrasound range data using the Monte Carlo method.

Monte Carlo integration methods, in their modern form, were introduced by Ulam and von Neumann [31], and by Fermi. These methods, first applied to calculating intranuclear cascades [15], have been applied since that time to a wide variety of physical and statistical problems. In the present work, we treat the (normalized) error histograms as measured probability distributions. We then use the Monte Carlo method to follow the propogation of errors as we fuse ultrasound ranges with the visual area data.

There are two techniques for selecting values for the random variables. In the rejection, or direct, Monte Carlo method of von Neumann [25], a pair of random numbers is generated. The first random number is used to select a candidate x_i. The

second random number is then compared to f_i to determine whether the candidate x_i is accepted or rejected. If it is rejected another pair of random numbers is generated and the procedure is repeated until an accepted pair is found.

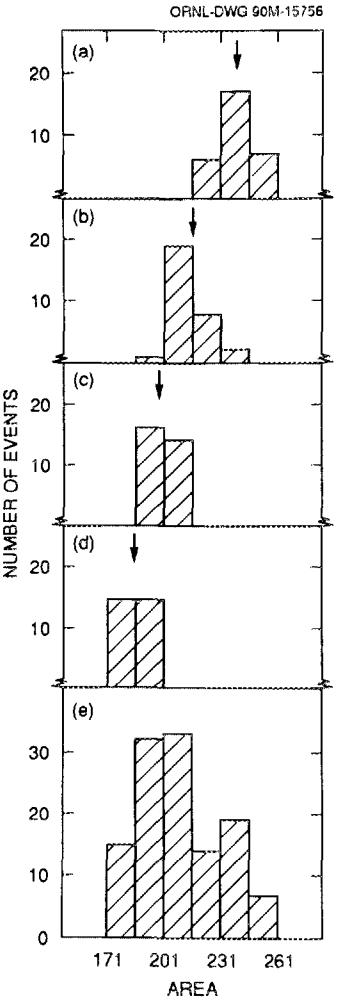

Fig. 6. Frequency distributions of areas (cardinalities) of the light at different docking zone (camera) distances. (a) 45.7 cm (57.2 cm); (b) 48.3 cm (59.2 cm); (c) 50.8 cm (62.2 cm); (d) 53.3 cm (64.8 cm); (e) summed over all distances.

In the summation, or inversion, Monte Carlo method only one random number is required in order to select an x_i. In this second approach the cumulative frequencies, g_k, are generated, as given by Eq. (4):

$$g_k = \sum_{i=0}^{k} f_i \tag{4}$$

The random number is used to select a value for g_k which, upon inversion, yields the corresponding x_i. The summation method produces rapid results when the distributions are simple and do not have to be recalculated for each trial.

The inversion method is illustrated in Fig. 7. In the first part, Fig. 7(a), we display a representative frequency distribution. The partial sums of the frequency distribution forms a staircase distribution. This distribution is plotted in Fig. 7(b), and the inversion of g_k to form h_l is displayed in Fig. 7(c). The frequency distribution generated from h_l is compared to the original distribution in Fig. 7(d). In making this comparison the regenerated distribution was normalized to the same total area as the original. As can be seen in Fig. 7(d), there are minor departures for the smallest bins; otherwise the original distribution has been reproduced exactly.

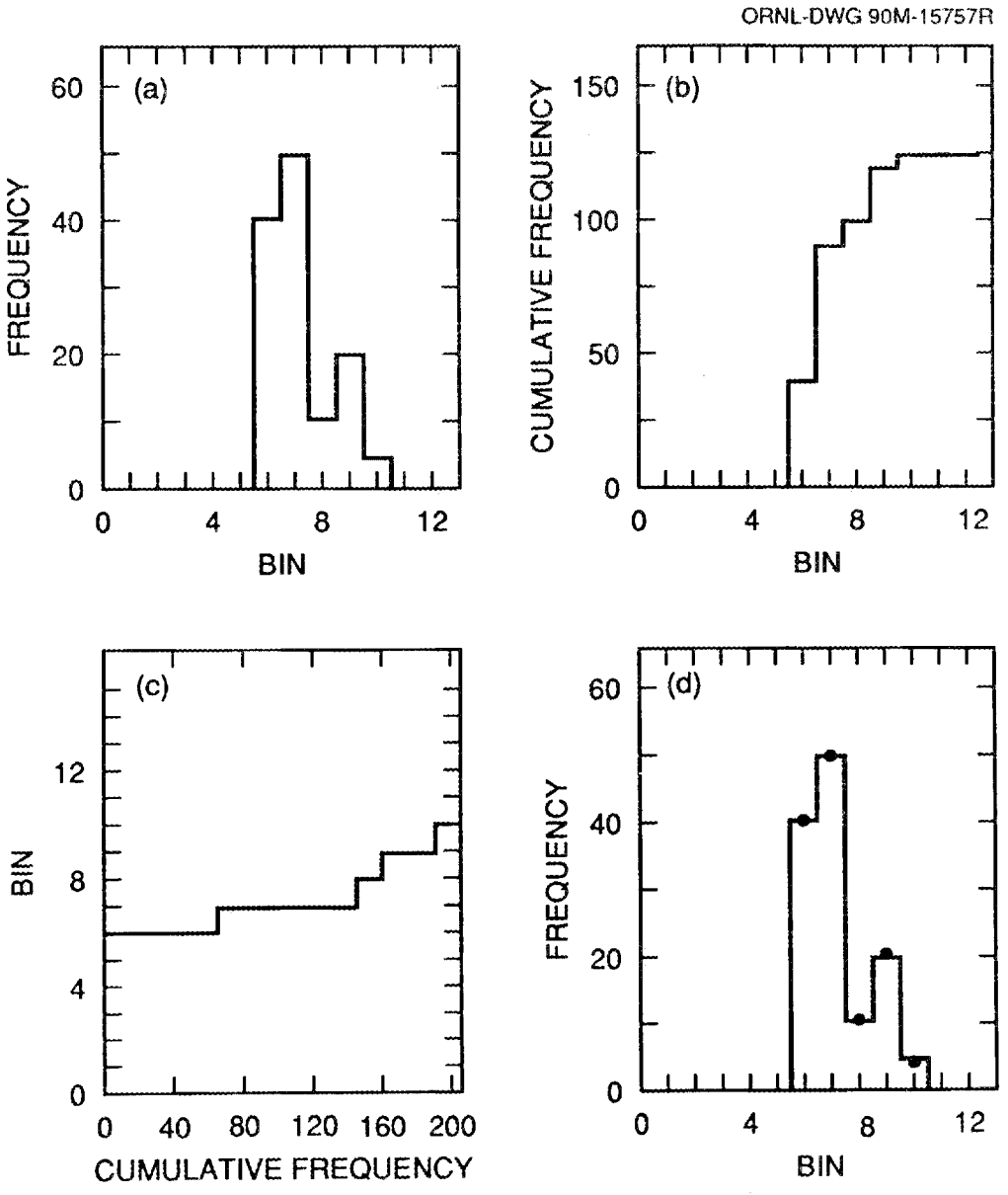

Fig. 7. Monte Carlo method. (a) Frequency distribution; (b) Partial-sum (staircase) distribution; (c) Inverted staircase distribution; (d) Comparison of original distribution (histogram) to regenerated distribution (filled circles).

In our calculations we scaled the random numbers to span the desired range, and then truncated to an integer representing the index of the array h_l. The magnitude of the range so chosen determined the reliability of this discrete implementation. In the calculations presented in Fig. 7(d) a range of 200 was selected. The number of Monte Carlo trials was 10,000.

Three types of distributions were used to integrate the area and range data. These were: (i) the distribution of positions within the docking zone; (ii) the distribution of ranges; and (iii) the distribution of areas. For (i) we used a uniform distribution of distances, neglecting the angular and yaw variations. In this approximation the distance, d, was given by the simple expression

$$d = d_{\min} + \Delta_d * \xi_1 \quad , \tag{5}$$

where

$$\Delta_d = d_{\max} - d_{\min} \quad , \tag{6}$$

$d_{\max}$ and $d_{\min}$ are the maximum and minimum distances within the docking zone, and ξ_1 is a random number.

For (ii) we adopted a simple two-step approximation to the results of repeated distance measurements. In these range measurements we delineated a small transition region, on the order of 10% of the binwidth, where the value returned by the sensor underwent a transition from one value to the next. The first step in the two-step approximation represented the single range value from the central 90% of the bin; the second step a 50-50 choice of range in the transition region. That is, the range, r, was determined as

$$\begin{aligned} r &= r_i, \quad |d - r_i| \leq \frac{\Delta_r}{2} \\ r &= \zeta_2 r_i + (1 - \zeta_2) r_{i\pm1}, \quad |d - r_i| > \frac{\Delta_r}{2} \quad , \end{aligned} \tag{7}$$

where r_i is the (midpoint) range for the i^{th} bin, corresponding to the distance d, Δ_r is the 90%-width, ζ_2 is either 0 or 1 depending on whether the random number ξ_2 is less than or greater than 0.5, and the plus/minus sign refers to upper or lower transition region.

Lastly, for (iii) we used the data presented in Fig. 6. Unit area distributions were generated from the four histograms. This was done using the total frequency for each bin as the unit area frequencies within that bin. The inverse distributions were then generated from the resulting cumulative distributions. To avoid rounding-off errors in the discrete approximation a range, $l_{\max}$, was calculated as

$$l_{\max} = \Delta_A * N \quad , \tag{8}$$

where Δ_A is the width of the area bins in Fig. 6, and N is the total frequency, given by Eq. (2). The Monte Carlo areas, A_r, were then generated using a third random number, ξ_3, as

$$A_r = h_{\xi_3 l_{\max}} * (r^2 / < d >^2) \quad , \tag{9}$$

where $< d >$ is the mean robot position.

The results of this calculation are presented in Fig. 8. We observe in this figure that distribution of visual areas has been sharpened through the fusion with the

ultrasound data using the model given by Eqs. (5) to (8). Specifically, the root-mean-square deviation of the refined distribution was 15.3 while that of the original distribution was 21.4.

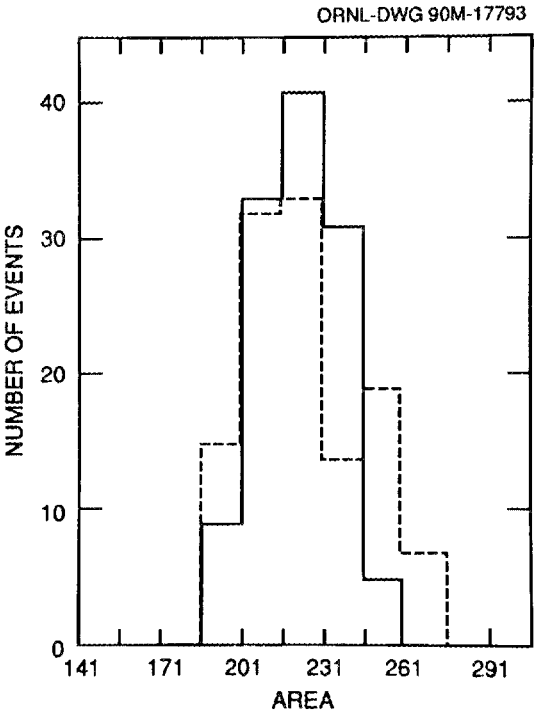

Fig. 8. Comparison of area (cardinality) distributions. Dashed histogram denotes the original frequency distribution shown in Fig. 6(e). Solid histogram gives the results after fusion. The Monte Carlo calculations were done for 10,000 events, and the results normalized to the total number of events in the original histogram.

5. DISTANCE AND YAW DATA

In this section we discuss the results of measurements of distance and yaw, neglecting the generally uninteresting angle data. We start with the monocular vision distance results, and then turn to the yaw data from both the CCD cameras and the ultrasound sensor.

5.1 DISTANCE

The first of two sets of distance results are presented in Fig. 9. Plotted in this figure as a function of distance are the errors in the distances extracted from the monocular vision data. In the experiments the control panel was centered on the robot with its surface orthogonal to the optic axes of the cameras, that is, the robot-control panel angle and yaw were zero. The two cameras used in the experiments were situated one on each side of the centerline, producing a small, distance-dependent angle with respect to the control panel.

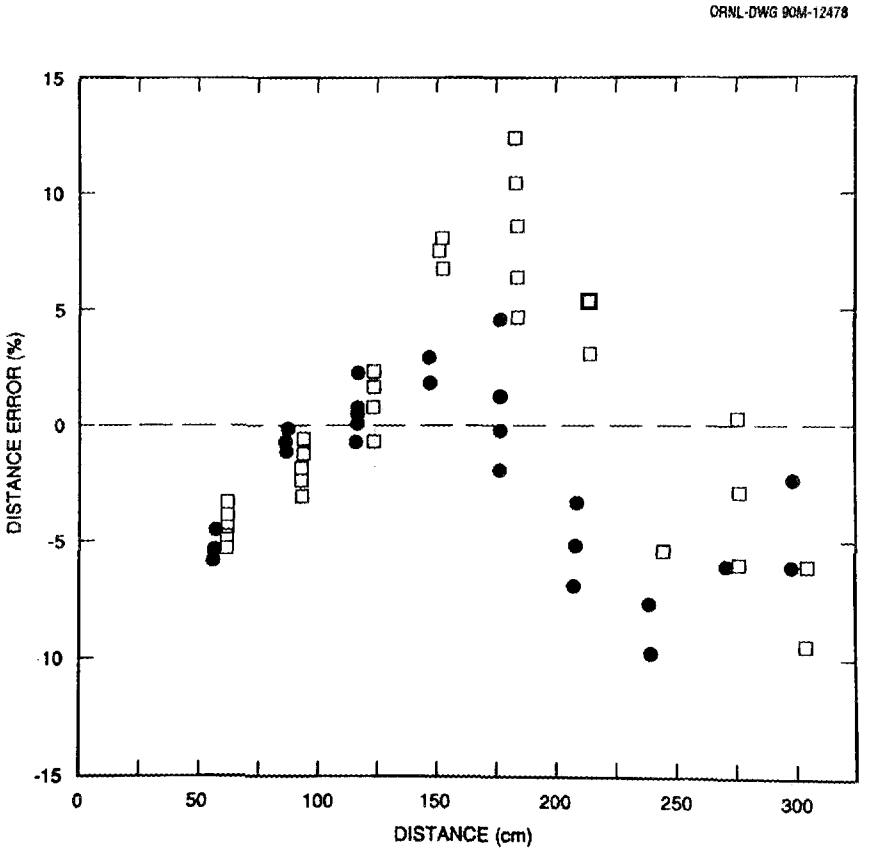

Fig. 9. Plot of distance errors as a function of distance. Open squares and filled circles denote the data from the left and right cameras, respectively. The dashed line is the zero error locus. Ten trials were done at each distance. The multiple entries for each distance give the corresponding values.

In the figure we find that the most accurate and precise results are for distances on the order of 1 m. At smaller distances the errors begin to increase. The errors at these small viewing distances correspond to extracted areas which are too large by about 10%. At distances greater than about 2 m the areas are again too large. More importantly, at large distances effects due to the small numbers of pixels in the regions of interest are present. As a result the precision of the estimates is comparable to the accuracy at these large distances.

The second set of data, showing distance errors as a function of angle, for zero yaw, are presented in Fig. 10. These data were acquired at distances up to 3 m. For completeness, the data from the previous figure have been included, and appear as dense groups of data points at angles less than 10 deg. In this figure we observe a variety of excursions away from the zero error line. The most pronounced excursions occur for large positive angles. These excursions and their apparent lack of symmetry about 0 deg may be ascribed to a misalignment of the plane of the lens of the camera and/or cylindrical distortions. A rotation, for example, of the lens had an observable effect upon the distortions.

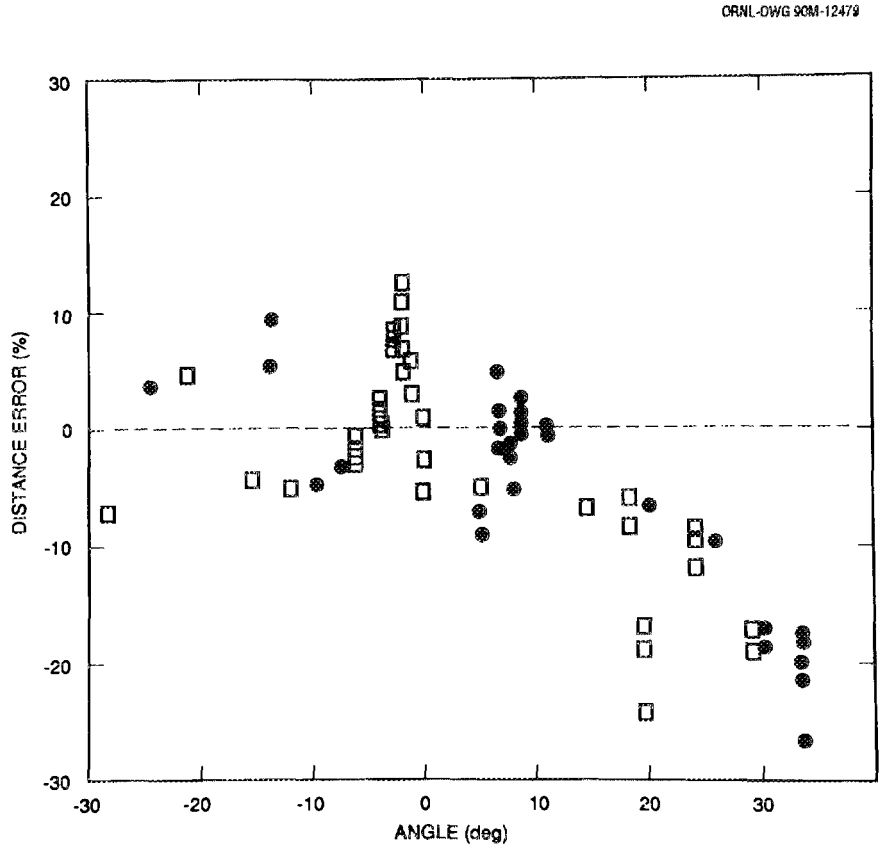

Fig. 10. Plot of distance errors as a function of angle. Definitions of the symbols are the same as in Fig. 9.

5.2 YAW

5.2.1 Visual Data

The distances over which yaw could be determined are smaller than those for which distance and angle can be deduced. This arises from the sensitivity, at large distances, of the measured cosine of the angle of rotation to small (unit pixel) variations in length. The measured quantities were stabilized as much as possible by using second moments to form the ratios of interest. Also, the control panel angle was set to zero, that is, the camera was pointed at the control panel, which was rotated through some angle to be measured.

Results of the measurements of yaw at two representative rotation angles are presented in Figs. 11 and 12. The data have been grouped by camera, and by the direction of rotation (positive or negative). The first observation to be made is that the overall agreement bewteen measured and physical values for the yaw is good. The error in the mean values for the yaw differ from the physical values by only a few percent. Two achieve these results an effective value for the ratio of meter width to height was used. This ratio differed from the physical ratio by about 10%. Otherwise, no parameter adjustments were made, and the same constants were used for both cameras.

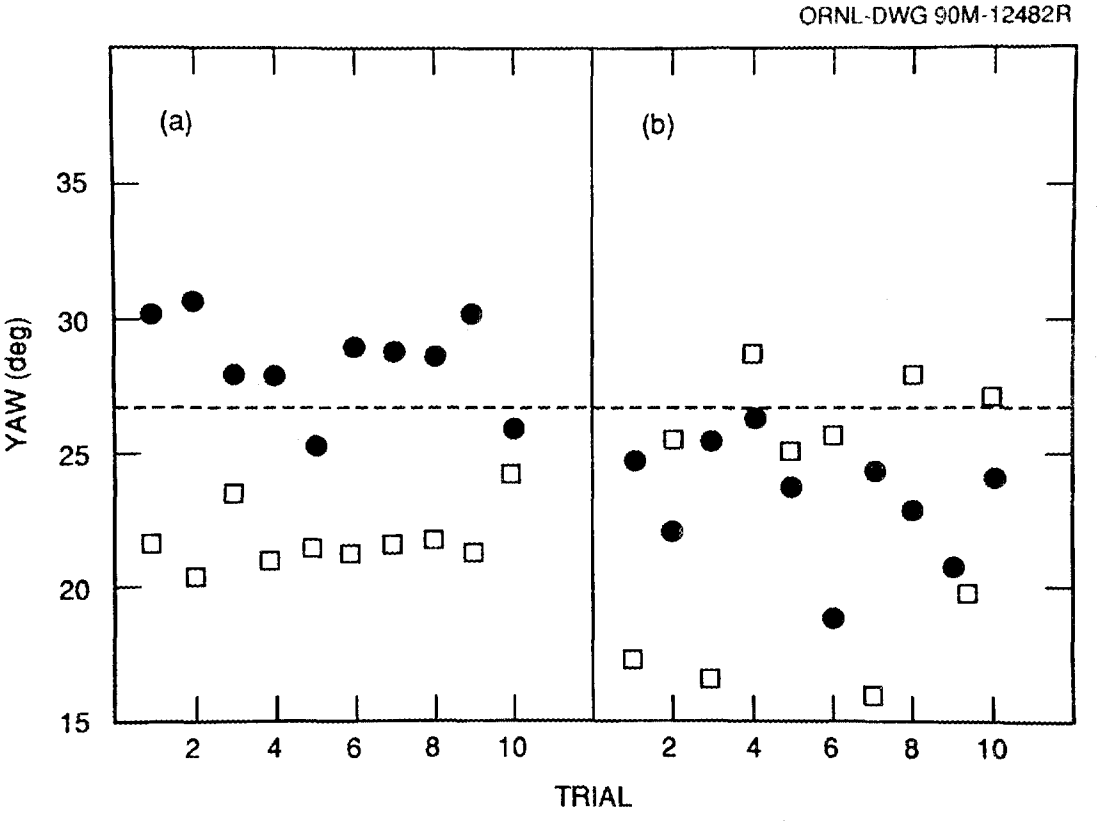

Fig. 11. Measurements of a 26.6 deg yaw. Part (a) negative yaw. Part (b) positive yaw. Open squares and filled circles as in Figs. 9 and 10. The dashed line is the zero error locus.

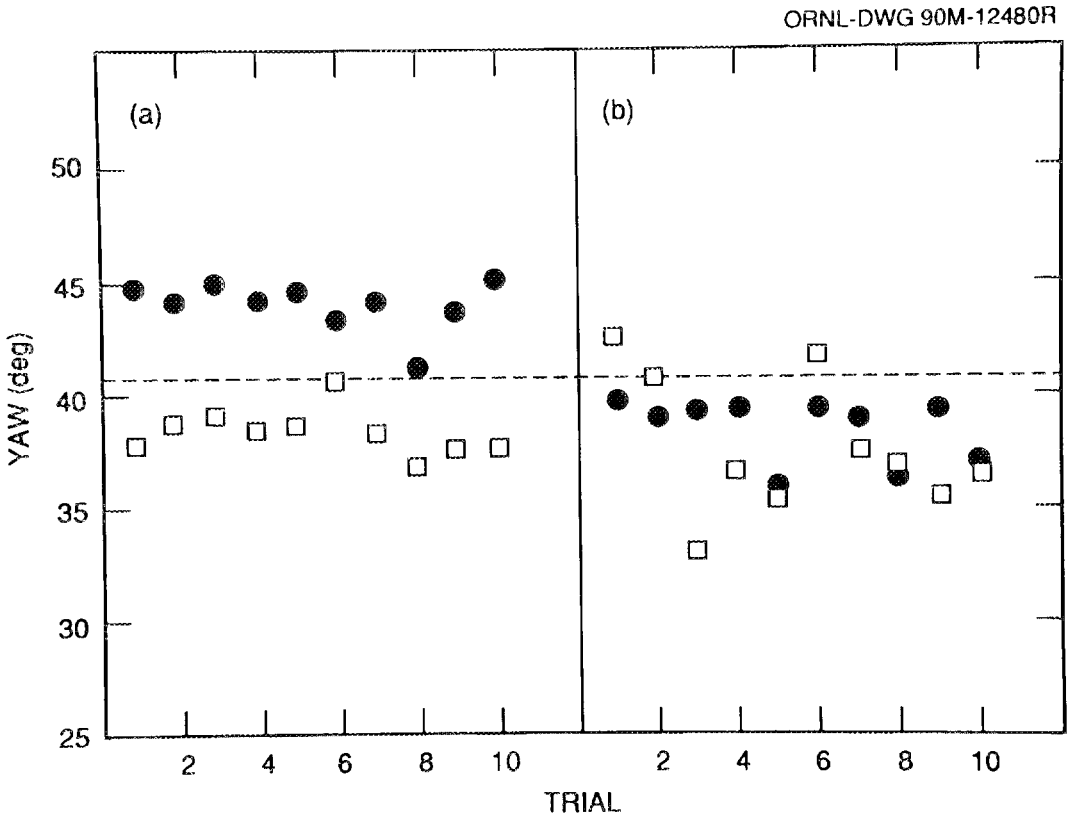

Fig. 12. Measurement of a 41.0 deg yaw. Part (a) negative yaw. Part (b) positive yaw. Symbols as before.

The mean values and root-mean-square deviations are listed in Table 4. We observe from the figures and table that the accuracy is higher at 41 deg than at 27 deg. Specifically, the errors range from 1.8 to 5.0 deg at 27 deg, and from 1.5 to 2.5 deg at 41 deg. The precision, represented by the root-mean-square deviations, is also superior at the larger angle. This is a general trend, and we find that the precision deteriorates rapidly at rotation angles less than 10 deg, due to the cosine dependence.

Table 4. Summary of Visual Yaw Measurements

Camera	Physical Angle (deg)	Measured Mean Angle (deg)	Measured RMS Deviation (deg)
Left	-26.6	-21.6	0.98
	26.6	23.2	4.26
	-41.0	-39.3	0.71
	41.0	38.9	2.06
Right	-26.6	-28.4	1.82
	26.6	23.5	1.87
	-41.0	-43.5	0.80
	41.0	39.5	0.97

Two other observations can be made. First, the spreads in the data for rotations in the positive direction are considerably greater than those for the negative direction. This change in spread can be ascribed to differences in lighting. In the case of positive rotations the front of the control panel is being illuminated, in part, from the window. Second, there are differences between the results from the two cameras. These differences are most likely calibration, or alignmnt, related.

5.2.2 Ultrasound Data

We noted in Sec. 5.1.1 that the $\cos \vartheta_y$ dependence of the yaw estimates extraced from the visual image severly limits its precision at small angles. The ultrasound method (Sec. 3.4), in contrast, provides an estimate of yaw which varies as $\sin \vartheta_y$. Thus, at small angles, the behavior of the quantities measured:

$$\text{measurement} \sim \sin \vartheta_y \approx \vartheta_y, \text{ for ultrasound data}$$

vs.

$$\text{measurement} \sim \cos \vartheta_y = 1 - \vartheta_y^2/2 \text{ for visual data}$$

The small-angle dependence of the ultrasound yaw estimate varies linearly with angle, and does not suffer from the insensitivity exhibited by the visual yaw data at angles near zero deg. (For example, the difference between the cosines of 0 deg and 10 deg is only 1.5%.) These simple considerations, when combined with a specification of the precision of the visual horizontal length and ultrasound range estimates, form elements of a model for the precision of the yaw estimates in the two sensor domains.

Ultrasound range data for 3.0, 9.1 and 45.0 deg are presented in Fig. 13. We observe in these plots that as yaw increases the differences in near and far side distances become larger. In more detail, the data at 9.1 deg displayed in Fig. 13(b) are typical of the small angle regime. We observe that the distribution of ranges from the near side, as well as the far side, is not sharp. Upon combining the root-mean-square deviations in quadrature we obtain a precision of 2.5 deg, while the error is 1.2 deg. Data at 3 deg is presented in Fig. 13(a) and Table 5. At this angle we have reached the limits of precision imposed by the bin size corresponding to the analog-to-digital conversion of the time-of-flight signals. This limit in precision at small angles can be overcome by acquiring more precise range data.

Table 5. Summary of Ultrasound Yaw Measurements

Distance (cm)	Physical Angle (deg)	Measured Mean Angle (deg)	Measured RMS Deviation (deg)
180	45.0	44.4	5.3
	26.6	23.8	6.5
	18.4	13.1	6.7
90	9.08	7.88	2.49
	3.02	3.33	3.96

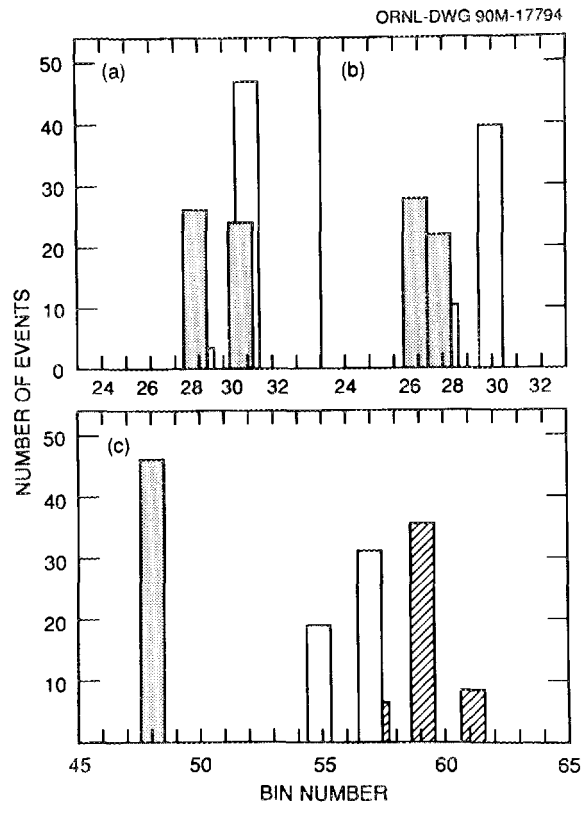

Fig. 13. Distribution of range values for rotation angles (yaw) of (a) 3.0 deg, (b) 9.1 deg, and (c) 45.0 deg. Fully shaded histograms denote the results from the near end of the control panel surface. Unshaded histograms show the frequencies from the far end of the surface. The partially shaded histogram displayed in part (c) gives the numbers of events for the "in" setting of the buttons; otherwise the buttons are "out." The ordinates are expressed in units of 3.0 cm. The skipping of ordinate values for sets of returns was produced by a hardware clock error.

Overall, useful data were obtained using the difference of ranges method at distances up to 2 m, and yaws from 0 deg to more than 45 deg. Within this broad zone the interplay between the beam and surface properties produced several observable effects. Range data for a yaw of 45 deg are displayed in Fig. 13(c). We observe that the distribution of ranges from the left-hand, or near, side of the panel is sharply peaked. This is contrast to the distribution of ranges from the right-hand, or far, side which is spread among several bin values. The broadening of the histogram is due to the presence of several scattering sources along the control panel surface. To illustrate this point two distributions are shown for the far side. In one histogram data were taken with the buttons in the "in" position; in the other case the buttons are "out." The differences between "in" and "out" configurations are apparent.

The data from 27 and 18 deg are similar. Values for the mean yaw and the root-mean-square deviation for the "in" configuration for the three rotation angles are

listed in Table 5. The yaw calculations were done assuming that the far returns were from the inner corner formed by the raised lip. As the rotation angle is decreased we note that the mean values fall progressively further below the that of the physical rotation angle. The reason for this is that the mean distances along the control panel from the left-side to the right-side returns are smaller than assumed in these calculations. This error manifests itself as a shortfall in the estimated yaw.

6. DISCUSSION

6.1 INTEGRATED SYSTEM PERFORMANCE

Docking trials were done in which distance and angle alone were measured and used together with prior information on the control panel yaw. We found in repeated measurements of the final docking position that the robot failed to successfully dock in 5 to 10% of the cases due either to wheel sticking (producing a 5 deg change in the yaw) or to human-induced initialization or calibration errors. It is clear that, even if prior information on yaw were available, without a reliable method for determining yaw the rate of successful dockings will be reduced in the vast majority of environments which are less gentle than that of the present study.

The 90 to 95% success rate in docking was achieved in an active approach where limitations of the sensor and the needs of the system were both taken into account. Returning to Figs. 9 and 10, we note that the most accurate and precise distance estimates were obtained at small angles and at distances on the order of 1 m. The path planning algorithm used by the integrated system was designed so that the final approach to docking zone was made through this "window." In doing so optimality of the path was sacrificed in favor of stability with respect to sensor errors.

Turning to the feature identification tasks, situations were encountered where there was insufficient information in the area and heigh-to-width data to disambiguate all of the devices on the control panel. In those instances binary filters were used to supplement the unary filters. (Unary filters are tests which depend exclusively on properties of the candidate region; binary filters are tests of relationships between two regions.) The small lamps located on the control panel just above the buttons were located in this manner. The objective was to discriminate bewteen these lamps and small regions of glare occurring primarily, but not always, along the edges of the lip around the control panel. The solution was to introduce a binary filter in which small regions were tested for an "above" association with the buttons.

In more detail, we recall from Sec. 3.2 that the visual data processing subsystem produced lists of the connected regions of the binary image and their geometric properties. In the panel recognition and device location tasks, several passes were made through these lists until only one candidate region remained. In each pass a specific unary or binary test was performed. The centroids and tolerances were converted to min-max bands sufficiently wide to accomodate the errors. The unary filtering was done first to produce a rapid decrease in the number of entries. Then, the more involved binary tests were performed.

6.2 STRATEGIC COMPONENTS

To prevent the performance of robotic sensor subsystems from deteriorating in unstructured, or ill-defined, environments, the interpretations of the output of the sensor data processing algorithms must be controlled and, if necessary, corrected. Useful sources of information for this purpose are sensor models and data from additional sensors and/or sensing locations. The first solution includes not only obvious models such as those for perspective and lens distortions, but also models which enable the system to choose between different data interpretations in a given situation. The yaw problem, where either ultrasound or visual methods can be used

with varying reliability at different distances and rotations, provides an example of this type of model.

Sensor fusion, the second solution, may be regarded as a strategy for reducing uncertainties and errors, and overcoming mechanical and sensor limitations, to achieve the desired performance goals. In a previous study [8], we examined sensor fusion done through a mutual exchange of information between sensor domains. For the cases examined in the present work, either the same quantity was measured in two different ways using different sensors, or data in one domain was refined, or conditioned, with data from a second sensor domain. Regardless of the type of fusion, a sensor model, in the second sense discussed in the previous papagraph, is needed to characterize the relative accuracy and precision of the sensor data and guide the fusion process.

To conclude this section, the above approaches, where sensor models and sensor fusion play key roles in the interpretation of sensor data, may provide the means for achieving reliable performance by automated systems operating in unstructured, and eventually, dynamic environments. These approaches are strategic in intent, and provide the means for guiding and controlling the interpretation of the output of data processing algorithms. Conclusions similar to ours have been reached by others in the field (see, for example, [5]).

7. SUMMARY AND CONCLUDING REMARKS

Our goal in this experimental study was to examine the performance of the visual and ultrasonic subsystems of a mobile robot operating in an unstructured environment. To do so we carried repeated measurements of the key geometric quantities used to guide navigation, docking, location and indentification tasks. A variety of errors were observed. We decomposed these errors into their systematic and random components, attempted to account for their origins, and devised means for their reduction.

This approach is the standard one in an experimental science. As noted by R. A. Fisher [14], "The experimenter interested in the causes which contribute to a certain effect is supposed, by a process of abstraction, to isolate these causes into a number of elementary ingredients, or factors, and it is often supposed, at least for purposes of exposition, that to establish controlled conditions in which all of these factors except one can be held constant, and then to study the effects of this single factor, is the essentially scientific approach to any experimental investigation." Fisher then observed that this prescription was a difficult one since "We are usually ignorant which, out of innumerable possible factors, may prove to be most important." In this latter regard we note that we have not specified in a precise quantitative form all environmental, hardware and software factors, but we have attempted to indicate which ones were perhaps most important.

Turning to the issues raised by the study, we note that some of errors reflect limitations of the hardware, and can be overcome by using more sophisticated (and costly) cameras, lens mounts, lenses, etc. More important are those errors which arise not because of failures of hardware, but rather as consequences of erroneous interpretations of the outputs of data processing algorithms. These systemtic errors, and the loss of precision through random errors, are likely to increase as the environments become progressive less pre-engineered.

In the present study, and in two previous ones [7], [8], we have tried to show how sensor models and sensor fusion can be used to minimize these errors to satisfy the performance requirements of a given task. These strategic components function in intelligent robotic systems in an active manner to control and guide the interpretations of the output of the data processing algorithms.

ACKNOWLEDGMENTS

The authors wish to acknowledge helpful discussions, especially during the preliminary stages of the study, with J. R. Einstein, J. P. Jones and S. M. Killough. This work was sponsored by the Engineering Research Program of the Office of Basic Energy Sciences, and the Office of Nuclear Energy, Office of Technology Support Programs of the U.S. Department of Energy under contract No. DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

REFERENCES

1. Anderson, T. L. and M. Donath, "Autonomous Robots and Emergent Behavior," Proc. IEEE International Workshop on Intelligent Robots and Systems IROS '90 (Tsuchiura, Japan), pp. 723–730, 1990.
2. Arkin, R. C., "Motor Schema-Based Mobile Robot Navigation," *Int. J. Robotics Res.*, Vol. 8, pp. 92–112, 1989.
3. Arkin, R. C. and R. R. Murphy, "Autonomous Navigation in a Manufacturing Environment," *IEEE Trans. Robotics Automat.*, Vol. RA-6, pp. 445–454, 1990.
4. Bajcsy, R., "Computer Identification of Visual Surfaces," *Comp. Graph. Image Process.*, Vol. 2, pp. 118–130, 1973.
5. Bajcsy, R., "Active Perception," Proc. IEEE, Vol. 76, pp. 996–1005, 1988.
6. Ballard, D. H., and C. M. Brown, *Computer Vision*, Englewood Cliffs: Prentice-Hall, 1982, Ch. 2.
7. Beckerman, M. and E. M. Oblow, "Treatment of Systematic Errors in the Processing of Wide Angle Sonar Sensor Data for Robotic Navigation," *IEEE Trans. Robotics Automat.*, Vol. RA-6, pp. 137–145, 1990.
8. Beckerman, M., L. A. Farkas and S. E. Johnston, "Treatment of Systematic Errors II: Fusion of Ultrasound and Visual Sensor Data," submitted to *IEEE Trans. Robotics Automat.*, 1990.
9. Brice, C. R. and C. L. Fennema, "Scene Analysis Using Regions," *Artificial Intell.*, Vol. 1, pp. 205–226, 1970.
10. Brooks, R. A., "A Robust Layered Control System for a Mobile Robot," *IEEE J. Robotics Automat.*, Vol. RA-2, pp. 14–23, 1986.
11. Chatila, R. and J.-P. Laumond, "Position Referencing and Consistent World Modelling for Mobile Robots," Proc. IEEE Int. Conf. Robotics Automat. (St. Louis, Missouri), pp. 138–145, 1985.
12. Crowley, J. L., "Navigation for an Intelligent Mobile Robot," *IEEE J. Robotics and Automat.*, Vol. RA-1, pp. 31–41, 1985.
13. Dunlay, R. T., "Obstacle Avoidance Perception Processing for the Autonomous Land Vehicle," Proc. IEEE Int. Conf. Robotics Automat. (Philadelphia, Pennsylvania), pp. 912–917, 1988.
14. Fisher, R. A., *The Design of Experiments*, Edinburgh: Oliver and Boyd, 1935, 3rd ed., 1942, pp. 90–91.
15. Goldberger, M., "The Interaction of High Energy Neutrons and Heavy Nuclei," *Phys. Rev.*, Vol. 74, pp. 1269–1277, 1948.
16. Haralick, R. M. and G. L. Kelly, "Pattern Recognition with Measurement Space and Spatial Clustering for Multiple Images," Proc. IEEE, Vol. 57, pp. 654–665, 1969.
17. Haralick, R. M., "Edge and Region Analysis for Digital Image Data," *Comp. Graph. Image Process.*, Vol. 12, pp. 60–73, 1980.
18. Haralick, R. M., "Performance Assessment of Near-Perfect Machines," *Mach. Vis. Appl.*, Vol. 2, pp. 1–16, 1989.

19. Harmon, S. Y., "The Ground Surveillance Robot (GSR): An Autonomous Vehicle Designed to Transit Unknown Terrain," *IEEE J. Robotics Automat.*, Vol. RA-3, pp. 266–279, 1987.

20. Harmon, S. Y. "A Report on the NATO Workshop on Mobile Robot Implementation," Proc. IEEE Int. Conf. Robotics Automat. (Philadelphia, Pennsylvania), pp. 604–610, 1988.

21. Jones, J. P., "A Concurrent On-Board Vision System for a Mobile Robot," in Proc. Third Conf. Hypercube Concurrent Computers and Applications, G. Fox, Ed., Vol. II, pp. 1022–1032, 1988.

22. Kanade, T., "Region Segmentation: Signal vs. Semantics," *Comp. Graph. Image Process.*, Vol. 13, pp. 279–297, 1980.

23. Materic, M. J. and R. A. Brooks, "Learning a Distributed Map Representation Based on Navigation Behaviors," Proc. 1990 Japan-U.S.A. Symposium on Flexible Automation (Kyoto, Japan), pp. 499–506, 1990.

24. Moravec, H. P., "The Stanford Cart and the CMU Rover," Proc. IEEE, Vol. 71, pp. 872–884, 1983.

25. Neumann, J. von, "Various Techniques Used in Connection with Random Digits," National Bureau of Standards Applied Mathematics Series, Vol. 12, pp. 36–38, 1951.

26. Nitzan, D., "Development of Intelligent Robots: Achievements and Issues," *IEEE J. Robotics Automat.*, Vol. RA-1, pp. 3–13, 1985.

27. Rosenfeld, A. and L. S. Davis, "Image Segmentation and Image Models," Proc. IEEE, Vol. 67, pp. 764–772, 1980.

28. Serra, J., *Image Analysis and Mathematical Morphology*, New York: Academic Press, 1982.

29. Thorpe, C., M. H. Herbert, T. Kanade and S. A. Shafer, "Vision and Navigation for the Carnegie-Mellon Navlab," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-10, pp. 362–373, 1988.

30. Turk, M. A., D. G. Morgenthaler, K. D. Gremban and M. Marra, "VITS - A Vision System of Autonomous Land Vehicle Navigation," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-10, pp. 342–361, 1988.

31. Ulam, S. M., and J. von Neumann, "On Combination of the Stochastic and Deterministic Processes," *Bull. Am. Math. Soc.*, Vol. 53, p. 1120, 1947.

32. Waxman, A. M., J. J. LeMoigne, L. S. Davis, B. Srinivasan, T. R. Kushner, E. Liang and T. Siddalingaiah, "A Visual Navigation System for Autonomous Land Vehicles," *IEEE J. Robotics Automat.*, Vol. RA-3, pp. 124–141, 1987.

INTERNAL DISTRIBUTION

1. B. R. Appleton
2. J. E. Baker
3–7. D. L. Barnett
8–12. M. Beckerman
13. B. L. Burks
14. P. L. Butler
15. G. de Saussure
16. J. R. Einstein
17. K. Fujimura
18. C. W. Glover
19. W. R. Hamel
20. J. N. Herndon
21. J. P. Jones
22. H. E. Knee
23. G. Liepins
24–28. R. C. Mann
29. S. A. Meacham
30. E. M. Oblow
31–35. F. G. Pin
36. D. B. Reister
37. P. F. Spelt
38. F. J. Sweeney
39. M. A. Unseren
40–44. R. C. Ward
45. EPMD Reports Office
46–47. Laboratory Records Department
48. Laboratory Records, ORNL-RC
49. Document Reference Section
50. Central Research Library
51. ORNL Patent Section

EXTERNAL DISTRIBUTION

52. Office of the Assistant Manager, Energy Research and Development, Department of Energy, Oak Ridge Operations, Oak Ridge, TN 37831
53. Dr. Peter Allen, Department of Computer Science, 450 Computer Science, Columbia University, New York, NY 10027
54. Dr. Harry Alter, Office of Technology Support Programs, Office of Nuclear Energy, U.S. Department of Energy, Washington, DC 20545
55. Dr. R. C. Arkin, School of Information and Computer Science, Georgia Institute of Technology, Atlanta, GA 30332
56. Dr. Ruzena Bajcsy, Department of Computer and Information Science, University of Pennsylvania, 3401 Walnut St., Philadelphia, PA 19104
57. Dr. Wayne Book, Department of Mechanical Engineering, J. S. Coon Building, Room 306, Georgia Institute of Technology, Atlanta, GA 30332
58. Dr. Rodney Brooks, MIT Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139
59. Dr. Max Donath, Department of Mechanical Engineering, University of Minnesota, 111 Church St., SE, Minneapolis, MN 55455
60. Prof. John J. Dorning, Department of Nuclear Engineering and Physics, Thornton Hall, McCormick Rd., University of Virginia, Charlottesville, VA 22901
61. Dr. Steven Dubowsky, Massachusetts Institute of Technology, Bldg. 3, Rm. 469A, 77 Massachusetts Ave., Cambridge, MA 02139
62. Dr. Scott Harmon, Hughes Research Laboratory, 3011 Malibu Canyon Rd., Malibu, CA 90265
63. Dr. Ewald Heer, Heer Associates, 5329 Crown Avenue, La Canada, CA 91011
64. Dr. Avi Kak, Robot Vision Lab, Department of Electrical Engineering, Purdue University, Northwestern Ave., Engineering Mall, West Lafayette, IN 47907

65. Dr. Takeo Kanade, Department of Computer Science and Robotics, Carnegie-Mellon University, Pittsburgh, PA 15213-3890
66. Dr. James E. Leiss, 13013 Chestnut Oak Drive, Gaithersburg, MD 20878
67. Dr. Oscar P. Manley, Division of Engineering, Mathematical, and Geosciences, Office of Basic Energy Sciences, ER-15, U.S. Department of Energy - Germantown, Washington, DC 20545
68. Dr. Max Mintz, Department of Computer and Information Science, University of Pennsylvania, 3401 Walnut St., Philadelphia, PA 19104
69. Dr. Tom Mitchell, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
70. Dr. Hans Moravec, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
71. Prof. Neville Moray, Department of Mechanical and Industrial Engineering, University of Illinois, 1206 West Green St., Urbana, IL 61801
72. Dr. Michael J. Rabins, Department of Mechanical Engineering, Texas A&M University, 100 Engineering Physics Building, College Station, TX 77843
73. Dr. Karl N. Reid, College of Engineering, Architecture and Technology, Oklahoma State University, Stillwater, OK 74074
74. Dr. Wes Snyder, Department of Radiology, Bowman Gray School of Medicine, 300 S. Hawthorne Dr., Winston-Salem, NC 27103
75. Prof. Mary F. Wheeler, Mathematics Department, University of Houston, 4800 Calhoun, Houston, TX 77204-3476

76–85. Technical Information Center, P.O. Box 62, Oak Ridge, TN 37831