Synergism of Binocular and Motion Stereo for Passive Ranging

BIR BHANU, Senior Member, IEEE
SUBHODEV DAS
University of California
PETER SYMOSEK
SCOTT SNYDER
BARRY ROBERTS, Member, IEEE
Honeywell Systems and Research Center

Range measurements to objects in the world relative to mobile platforms such as ground or air vehicles are critical for visually aided navigation and obstacle detection/avoidance. An approach is presented that consists of a synergistic combination of two types of passive ranging method: binocular stereo and motion stereo. We show a new way to model the errors in binocular and motion stereo in conjunction with an inertial navigation system (INS) and derive the appropriate Kalman filter to refine the estimates from these two stereo ranging techniques. We present results using laboratory images that show that refined estimates can be optimally combined to give range values which are more accurate than any one of the individual estimates from binocular and motion stereo. By incorporating a blending filter, the approach has the potential of providing accurate, dense range measurements for all the pixels in the field of view (FOV).

Manuscript received December 3, 1991; revised April 19, 1993.

IEEE Log No. T-AES/30/3/16636.

Authors' addresses: B. Bhanu and S. Das, College of Engineering, University of California, Riverside, CA 92521; P. Symosek, S. Snyder, and B. Roberts, Honeywell Systems and Research Center, 3660 Technology Dr., Minneapolis, MN 55418.

0018-9251/94/$4.00 © 1994 IEEE

I. INTRODUCTION

Range measurements to objects in the world relative to mobile platforms such as ground or air vehicles are critical for visually aided navigation and obstacle detection/avoidance. Active (laser) range sensors can be used to provide such range measurements although they have a limited field of view (FOV), suffer from slow data acquisition, and are expensive. Robust passive ranging techniques can be suitable alternatives. The passive visual cues of binocular and motion stereo have been the two most popular methods for range estimation. A plethora of algorithms have been proposed to estimate three-dimensional (3D) structure or motion or both, using these two cues individually. "Robustness" of the algorithms is sought by selecting stable features in the stereo images for matching, performing accurate camera calibration, removing lens distortion or employing less noise-sensitive computational methods. Applications of such algorithms are shown mostly for synthesized data, some for real scenes, and few for outdoor scenarios. However, it is to be realized that "robustness" cannot guarantee high precision of the estimates derived using any one of the cues which are inherently imprecise. Consequently, in any experiment involving these cues some 3D locations will consistently have better estimates than some others. Besides, most assumptions about "robustness" and the robust characteristics are unlikely to survive in major real-world applications, such as autonomous mobile robots operating outdoors.

The objective of this research is to develop a passive ranging system that utilizes the benefits of binocular and motion stereo. This system is based on the synergistic combination of the two stereo modalities which is achieved by the following sequence of operations: interest point matching, Kalman filtering, and range measurement blending. The important benefits of the proposed synergistic system are, 1) the system is cheap to build (compared with active sensors), 2) it is passive (i.e., nondetectable, covert), 3) a more dense and more accurate range map is generated than is possible by either passive technique alone (this is necessary for obstacle avoidance), and 4) negligible motion distortion is caused by the moving platform (i.e., fast data acquisition).

Previous efforts in the derivation of approaches for the synergistic combination of binocular and motion stereo ranging have placed restraints on the problem specification to reduce the complexity of the analysis. To date, no demonstration of a totally general, comprehensive characterization of the ranging problem for multiple binocular stereo frames has been derived.

The emphasis of this work is on modeling the errors in binocular and motion stereo in conjuction with an inertial navigation system (INS) for a real-world application of a passive ranging system, and deriving the appropriate Kalman filter to refine the estimates from these two stereo ranging techniques. Our particular approach is designed to allow empirical evaluation of the performance and robustness of the passive ranging system for various scenarios. The next section describes in greater detail the background and motivation behind the work reported here. Section III presents the technical approach adopted in designing the synergistic system. Section IV discusses results obtained during an empirical evaluation of the performance of this system with laboratory data and a simulated inertial reference unit (IRU) data. Section V discusses our plans for future research to further optimize this system.

II. BACKGROUND AND MOTIVATION

In this section we summarize the past research related to the work reported in this paper, and the motivations that lead to the development of the approach described in the following section.

A. Background

Features from stereo pairs of images can be matched over time to obtain better accuracy for disparity-based range calculations. Leung, et al. [8] derived an algorithm for finding point correspondences among stereo image pairs at two consecutive time instants (t_{i-1}, t_i). They demonstrated significantly improved feature matching accuracy for scenes that demonstrate large feature displacements due to object motion in the scene. Li and Duncan [9] estimated the platform motion from measures for the optical flow of the left and right cameras for a series of binocular stereo images, without point-to-point correspondences. In addition, stereo matching procedures based on the estimated translational velocity and the flow fields were derived. An empirical evaluation of the robustness of the approach to image noise (which degrades the accuracy of the flow field) for synthetic images was carried out. The approach was demonstrated to be robust for representative flow field magnitude and direction errors. Sridhar and Suorsa [13] describe recursive binocular and motion stereo algorithms and compare their performances. However, the confidence factors for each of the range measurements which form the basis of such comparison are obtained by considering only the errors in image locations of matched feature points. The uncertainty models of their passive ranging techniques are therefore inadequate for a real-world imaging system such as a mobile platform. Several researchers have used the Kalman filtering method to estimate range from binocular stereo images [5] and motion sequences [10].

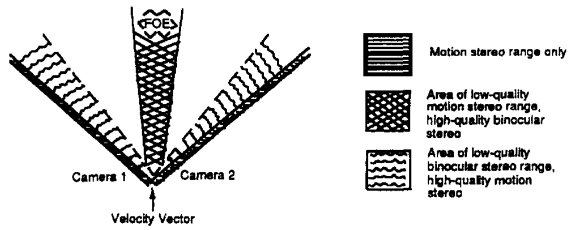


Fig. 1. Modalities for passive ranging.

B. Motivation

A synergistic combination of binocular and motion stereo is motivated by the following observations about their relative performance as illustrated in Fig. 1. Binocular stereo-based range computations suffer the greatest error at the edges of field of view (FOV) of the camera where motion stereo-based range is most accurate; the converse scenario holds true in the vicinity of the focus of expansion (FOE) where motion stereo-based range error is very large and binocular stereo-based range error is very small. Thus, a passive ranging system which employs only one of these two methods of range computation is likely to perform poorly even with the most robust algorithm. On the other hand, a passive ranging system which can successfully employ both methods, has the advantage of retaining only the best range estimate of a scene point. Which one of the methods provides the best range estimate is determined by the location of the point in the FOV. This may mean that the visual field can be appropriately segmented to be processed by either binocular stereo or motion stereo, thereby reducing the computational burden. Alternately, range values for distinct points in the visual field can be computed from both binocular and motion stereo and be refined using the statistics of their uncertainties. The refined range estimates for each point can be statistically combined to yield a more precise range value.

Most passive ranging techniques developed to date make idealistic assumptions about their operational conditions. On the contrary, in most real-world applications the conditions are far from being ideal. Some of these, such as vibration of the platform on which the cameras are mounted or the wind speed, may prove to be catastrophic to the ranging techniques, such as determination of motion parameters in order to compute range. Incorporation of hardware which can compute stable values of motion parameters under harsh operating conditions, will greatly improve the performance of any motion analysis technique for motion stereo-based ranging. An INS is one such item which is used in many types of land and air vehicles.

An INS includes an IRU and the necessary hardware and software to stabilize and process the

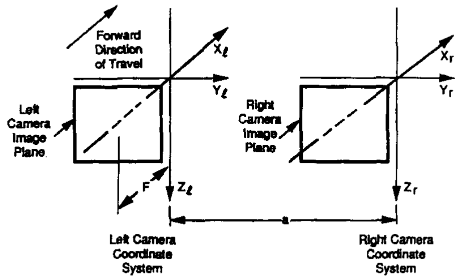


Fig. 2. Reference coordinate system.

IRU outputs to derive values for the position and velocity of the platform in a desired reference frame [4]. IRU measurements are made with gyroscopes, to provide an absolute measure of the rotation difference between the coordinate frame of the vehicle and a fixed, geographic, reference frame. Such measurements are also made with accelerometers, to provide the acceleration of the vehicle relative to the reference frame. The time integral of these accelerations gives the velocity and position of the vehicle. In addition, INS information can be used to select interest points in the visual field where range computations are to be made, and to determine sensor motion between frames [3, 12].

III. INTEGRATED APPROACH

With a two-camera system in motion, a stereo ranging system is formed which consists of *binocular stereo* and *motion stereo* range computations. In the case of binocular stereo two cameras are rigidly mounted on the same fixture such that their optical axes are parallel and yet horizontally displaced by a fixed, known distance. Whereas the cameras are laterally displaced for binocular stereo, the cameras are longitudinally displaced, due to forward vehicle motion, for motion stereo. On a moving platform, the same two cameras can provide the imagery required to perform one binocular and two motion stereo range calculations. We use a right-handed coordinate system for both cameras of the binocular stereo ranging system. As shown in Fig. 2, the x-axis is parallel to the forward direction of travel, the y-axis points rightward, and the z-axis points down. In subsequent analyses, the image coordinates are denoted by (u,v) while the pixel coordinates are denoted by (y,z).

Our integrated stereo system, shown in Fig. 3, uses the following two key elements which constitute the unique features of our approach: 1) matching of interesting points in binocular stereo and motion stereo imagery, and 2) modeling of range errors present in the motion and binocular stereo techniques. These errors are represented as the states of a Kalman filter applied to obtain improved estimates of range values.

The coincident points of interest, i.e., those points for which range is computed by both motion and binocular stereo techniques, are used as measurements to estimate errors in the ranging processes. The points in the ragne maps which are not coincident can be corrected with these error estimates, improving the overall quality of the composite range map. This can be achieved with the use of a blending filter as shown in Fig. 4. This filter derives a composite range map for each measurement location as the weighted average of the Kalman filter estimates for the range, where the averaging weights are the current estimates of the measurement noise obtained from each filter. The confidence in each range measurement is inversely proportional to the estimate of the measurement noise, so that when the measurement noise for the binocular stereo algorithm is large, the estimate obtained from the motion stereo Kalman filter is weighted more heavily. Conversely, when the measurement noise for the motion stereo algorithm is large, the estimate from the binocular stereo algorithm is given more weight. The filtered estimates of the measurement noise are used to smooth out the effects of isolated bad measurements.

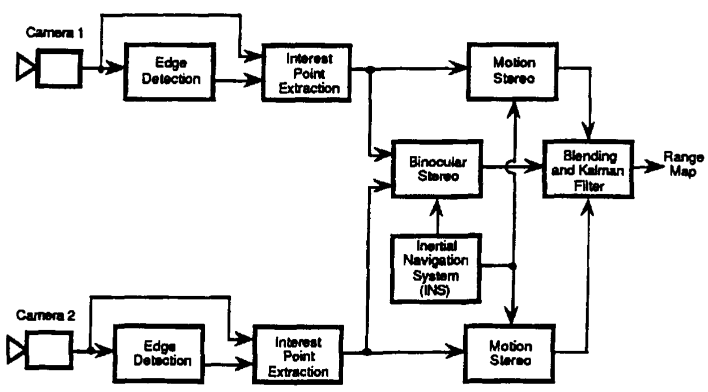


Fig. 3. System for integration of motion stereo and binocular stereo.

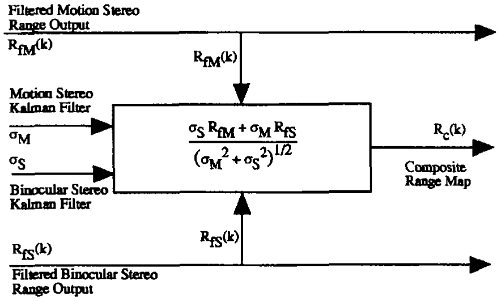


Fig. 4. Composite range map/blending filter.

A. Range Error Modeling

In this section, we discuss the approach for synergistic combination of motion and binocular stereo-based range estimates. The disagreement between the calculated ranges from the motion and binocular stereo algorithms for the coincident points of interest is attributed to the errors in inertial data and geometric alignment of the cameras. The computed discrepancies in the range values are used by a Kalman filter to refine the estimates for the errors in the inertial and system configuration parameters. New estimates could be obtained by adding the updated estimates for the errors to the expected system variable magnitudes. The refined estimates could then be used to calculate improved binocular and motion stereo ranges. Alternatively, the H-matrices for each coincident interesting point could be derived from the dependence of errors for binocular and motion range calculations on errors in the inertial and system configuration data, and the output ranges for these algorithms could be corrected using a linear combination of the error states of the Kalman filter. The latter is done in the current implementation of the passive ranging system.

The measurement for the binocular stereo component of the filter is the difference of the ranges from binocular and motion stereo. The motion stereo measurement of the filter is the negative of the binocular stereo measurement of the filter. Using the static Gauss–Markov discrete time model, the measurement process is described as follows:

$$y_{Mj}(k) = R_{Mj} - R_{Sj} = Hx_M + \nu_M \tag{1}$$

$$y_{Sj}(k) = R_{Sj} - R_{Mj} = Hx_S + \nu_S \tag{2}$$

where $y_{Mj}(k)$ is the measurement for the motion stereo component of the Kalman filter for the jth feature point location at time k, $y_{Sj}(k)$ is the measurement for the binocular stereo component of the Kalman filter for the jth feature point location at time k, R_{Mj} is the estimate of the range corresponding to the jth feature point from the motion stereo algorithm, R_{Sj} is the estimate of the range corresponding to the jth feature point from the binocular stereo algorithm, x_M is the error state vector for the motion stereo Kalman filter, x_S is the error state vector for the binocular stereo Kalman filter, ν_M is the measurement noise for the motion stereo Kalman filter where $E\{\nu_M^T \nu_M\} = \sigma_M^2$ is large near FOE and small near periphery, and ν_S is the measurement noise for the binocular stereo Kalman filter where $E\{\nu_S^T \nu_S\} = \sigma_S^2$ is small near FOE and large near periphery.

As stated previously, the binocular stereo and motion stereo range errors are linear combinations of the Kalman filter error states. The linear combination can be expressed as:

$$\delta R_S = H\hat{x}$$

$$\delta R_M = H\hat{x}$$

where H is the measurement matrix defined by the total differential of binocular stereo range and the total differential of motion stereo range, respectively, for the preceding pair of equations, and $\hat{x}$ is the estimated error state vector.

We derived the functional relationships of errors in range values. The total differential of motion stereo range is

$$\begin{aligned} dR_f = {} & \frac{\partial R_f}{\partial y'}dy' + \frac{\partial R_f}{\partial z'}dz' + \frac{\partial R_f}{\partial y}dy \\ & + \frac{\partial R_f}{\partial z}dz + \frac{\partial R_f}{\partial fov_h}dfov_h + \frac{\partial R_f}{\partial fov_v}dfov_v \\ & + \frac{\partial R_f}{\partial F}dF + \frac{\partial R_f}{\partial \Delta\psi'}d\Delta\psi' + \frac{\partial R_f}{\partial \Delta\theta'}d\Delta\theta' \\ & + \frac{\partial R_f}{\partial \Delta\phi'}d\Delta\phi' + \frac{\partial R_f}{\partial v_x}dv_x + \frac{\partial R_f}{\partial v_y}dv_y + \frac{\partial R_f}{\partial v_z}dv_z \end{aligned} \tag{3}$$

where (y',z') is the pixel location of an interest point in the left frame of a motion stereo pair of images that is acquired at time t_{i+1}, (y,z) is the pixel location of the interest point in the left frame of a motion stereo pair of images that is acquired at time t_i and matches (y',z'), fov_v is the camera vertical FOV, fov_h is the camera horizontal FOV, $\Delta\psi'$ is the change in yaw angle that occurred in the time interval $t_{i+1} - t_i$, $\Delta\theta'$ is the change in pitch angle that occurred in the time interval $t_{i+1} - t_i$, $\Delta\phi'$ is the change in roll angle that occurred in the time interval $t_{i+1} - t_i$, (v_x, v_y, v_z) is the velocity of the camera, and F is the focal plane to lens center distance.

The total differential of binocular stereo range is

$$\begin{aligned} dR_f = {} & \frac{\partial R_f}{\partial y_l}dy_l + \frac{\partial R_f}{dz_l}dz_l + \frac{\partial R_f}{\partial y_r}dy_r + \frac{\partial R_f}{\partial z_r}dz_r \\ & + \frac{\partial R_f}{\partial fov_h}dfov_h + \frac{\partial R_f}{\partial fov_v}dfov_v + \frac{\partial R_f}{\partial \Delta\psi}d\Delta\psi \\ & + \frac{\partial R_f}{\partial \Delta\theta}d\Delta\theta + \frac{\partial R_f}{\partial \Delta\phi}d\Delta\phi + \frac{\partial R_f}{\partial F}dF + \frac{\partial R_f}{\partial a}da \end{aligned} \tag{4}$$

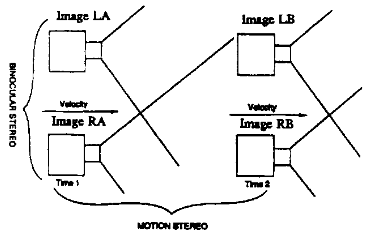


Fig. 5. Data acquisition for motion and binocular stereo.

where (y_l, z_l) is the pixel location of an interest point in the left frame of a binocular stereo pair of images acquired at time t_i, (y_r, z_r) is the pixel location of an interest point in the right frame of a binocular stereo pair of images acquired at time t_i and matches (y_l, z_l), $\Delta\psi$ is the boresight yaw angle, $\Delta\theta$ is the boresight pitch angle, $\Delta\phi$ is the boresight roll angle, and a is the camera separation distance.

In the above, we have given only the functional form of range errors. These equations are of great value in understanding the effect of various system and imaging parameters on the computed range. The complete equations for these partial derivatives are quite complicated and, for clarity, we have not presented them here. We have also derived the functional relationships between the variance of range error and the location of an interest point in the FOV. Further details of these steps may be found in [1, 2].

An approximation to the range calculation error for the case of motion stereo range computations is,

$$\sigma_M(u_A, v_A) = \sigma_{D_M} \frac{\Delta R_M(u_A, v_A)}{\sqrt{F^2 + u_L^2 + v_L^2}} \tag{5}$$

where σ_{D_M} is an initial estimate of the range calculation error due to the error in the motion stereo point matching algorithm, $\Delta R_M(u_A, v_A)$ is the computed error in range for the world point whose projection onto the image plane is described in three space by (F, u_A, v_A), and F is the distance between the lens center and the image plane.

Likewise, an approximation to the range calculation error for the case of binocular stereo range computations is,

$$\sigma_S(u_L, v_L) = \sigma_{D_S} \frac{\Delta R_S(u_L, v_L)}{\sqrt{F^2 + u_L^2 + v_L^2}} \tag{6}$$

where σ_{D_S} is an initial estimate of the range calculation error due to the error in the binocular stereo point matching algorithm, and $\Delta R_S(u_L, v_L)$ is the computed error in range for the world point whose projection onto the image plane is described in three space by (F, u_L, v_L). The variances of the measurement noises ν_M and ν_S of (1) and (2) are calculated using these approximations.

In computing range with either the motion stereo or binocular stereo techniques, all range measurements are made relative to the *first of a temporal pair of images* (i.e., A of A and B images) and the *left image of a stereo pair*, as shown in Fig. 5. Hence the subscripts A and L are used for the variables that describe points in three space on the image plane. In our implementation, the A and L images are the same image.

B. Kalman Filter Implementation

The navigation coordinate system for the INS is shown in Fig. 6. The true local level axes x_l, y_l, z_l are also known as north-east-down axes. The twenty-nine error states listed in Table I are mechanized in the Kalman filter. The first seven states [7] are based on the level axis "Psi-Angle" (ψ_1, ψ_2, ψ_3 of Fig. 6) IRU error model,

$$\dot{\psi} = -(\rho + \Omega) \times \psi - C\delta\omega \tag{7}$$

$$\delta\dot{\mathbf{V}} = C\delta\mathbf{A}^B - \psi \times \mathbf{A}^L(\delta\mathbf{R}\cdot\mathbf{R}/R)(\mathbf{R}/R) + \delta\mathbf{g}' \tag{8}$$

$$\delta\dot{\mathbf{R}} = \delta\mathbf{V} - \rho \times \delta\mathbf{R} \tag{9}$$

where ψ is the Psi-angle error (states 1, 2, and 3), $\delta\mathbf{V}$ is the Psi-angle horizontal velocity error (states 4 and 5), $\delta\mathbf{R}$ is the Psi-angle horizontal position error (states 6 and 7), ρ is the local level transport rotation rate (V/R), Ω is the Earth rate in local level coordinate frame $= [\omega_E \cos\lambda, 0, -\omega_E \sin\lambda]^T$, C is the body to local level direction cosine transformation matrix, $\delta\omega$ is the gyro error states (states 25, 26, 27), δA^B is the accelerometer error states (states 28 and 29), A^L is the local level acceleration, ω_S is the Shuler frequency ($\sim$ 0.00125 rps), $\mathbf{R}/R$ is the unit vector, and $\delta\mathbf{g}'$ is the gravity deflection and anomaly errors.

The vertical error states (8, 9, 10) assume an IRU vertical channel damped with altitude data from a radar altimeter. Fig. 7 shows a typical IRU vertical channel filter. The error model implemented in the Kalman filter can be expressed as

$$\dot{x}_8 = -x_9 \tag{10}$$

$$\dot{x}_9 = K_1 x_9 + x_{10} \tag{11}$$

$$\dot{x}_{10} = x_{24} + K_2 x_9 - K_3 x_8 \tag{12}$$

where K_1, K_2, K_3, are the vertical channel gains. In our system, these gains were selected as 0.6, 0.15, and 0.0156, respectively.

The remaining error states are modeled as Gauss–Markov processes with large time constants. A Gauss–Markov process can be represented as

$$\dot{x} = \frac{-1}{\tau} x + \eta \tag{13}$$

where η is a white noise process and τ is a time constant. For large time constants, the error sources are effectively modeled as constants.

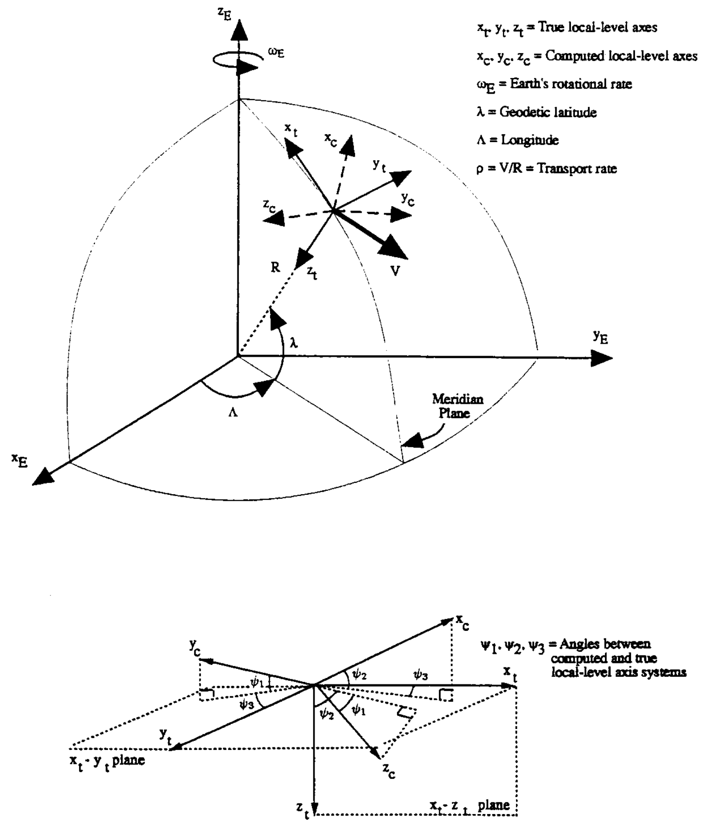


Fig. 6. Navigation coordinate system for INS.

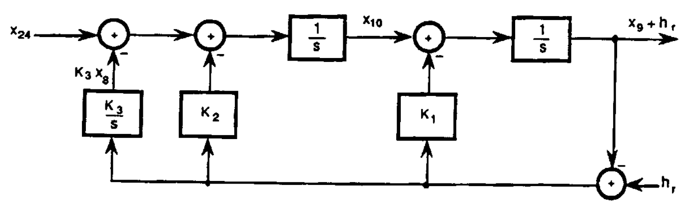


Fig. 7. Block diagram of vertical channel filter.

IV. IMPLEMENTATION AND RESULTS

In this section, we present details of implementing the synergistic combination of binocular and motion stereo.

A. Implementation Details

The following tasks were carried out to demonstrate the efficacy of our synergistic system.

1) laboratory collection of binocular and motion stereo data (a total of 5 pair of frames). For use in validating our integrated stereo technique, ground truth range to a selected number of image points was obtained.

2) development and software implementation of binocular and motion stereo algorithms and the integration of the two algorithm suites. This includes the extraction of "interest" points and the matching of interest points for subsequent binocular and motion stereo range computations.

3) Kalman filter implementation. We developed Kalman filter software to process range data measurements and derive estimates for the error states which contribute to range error.

4) evaluation of the integrated stereo system with real imagery.

For experimentation with the system, we wrote binocular and motion stereo algorithms in C and modified an existing Kalman filter software that was previously used in a real-time environment. The Kalman filter software is written in C and FORTRAN and was modified for a simulation environment.

TABLE I
Error States Used in Kalman Filtering

Error State	Description
1	IRU psi 1 angle error
2	IRU psi 2 angle error
3	IRU psi 3 angle error
4	IRU x velocity error
5	IRU y velocity error
6	IRU x position error
7	IRU y position error
8	Vertical channel acceleration error
9	Vertical channel velocity error
10	Vertical channel position error
11	Horizontal FOV error (fov_h)
12	Vertical FOV error (fov_v)
13	Camera focal plane to lens center distance (F)
14	y_l left camera Y optical axis offset error
15	z_l left camera Z optical axis offset error
16	y_r right camera Y optical axis offset error
17	z_r right camera Z optical axis offset error
18	Camera yaw angle boresight error
19	Camera pitch angle boresight error
20	Camera roll angle boresight error
21	Camera separation distance (a)
22	y_l' left camera Y optical axis offset error (past frame)
23	z_l' left camera Z optical axis offset error (past frame)
24	Z accelerometer bias
25	X gyro bias error
26	Y gyro bias error
27	Z gyro bias error
28	X accelerometer bias error
29	Y accelerometer bias error

Camera parameters used were: horizontal FOV, hfov=0.754160 rad; vertical FOV, vfov=0.313147 rad; focal length, F = 0.0410 ft; baseline, $a = 2.0$ ft.

For the purposes of efficiency, only one Kalman filter is used by the integrated system by stacking the binocular and motion stereo measurements into a single $2N \times 1$ column vector, where N is the number of feature points matched by both algorithms for a specific image. The H-matrix is obtained by stacking the total differential of binocular stereo range and the total differential of motion stereo range into a single $2N \times 29$ matrix, where 29 is the number of states of the Kalman filter for our integrated system.

IRU errors were simulated by running an off-line IRU error simulation and adding the resulting errors to our nominal motion. The simulation used was a Monte-Carlo simulation of the IRU error equations. For this research, we simulated a GG1328 gyro-based IRU which is a low cost, 0.1 mrad angular orientation accuracy, integrated gyroscope and INS. The trajectory chosen was a northern cruise at 15 ft/s. Fig. 8 shows simulation results for the first 10 s. For a cruise scenario such as the trajectory above, IRU errors are essentially a function of time. Therefore, to formulate IRU errors for our two cases, the true trajectory was subtracted from the Fig. 8 data. The resultant error data was then added to our integrated system trajectory to simulate corrupted IRU data.

For the initial evaluation phase, our stereo system was not fully integrated, but left in modular components. These components consist of laboratory collected video files, an IRU error model simulation, binocular stereo range algorithm, motion stereo range algorithm, and Kalman filter algorithm. Each of these components are run separately with communication between the components through input/output files.

B. Experimental Results

Five frames (each 512 × 512 pixels) of video data were collected in the laboratory at 2 ft intervals. An example of the experimental data is shown in Fig. 9. To simulate motion for the motion stereo algorithm we chose two velocities, 2 ft/s and 20 ft/s. These two velocities correspond to processing the four frames at 1 s time intervals or 0.1 s intervals. The attitude of both experiments was chosen to be level and in a northerly direction.

From these five frames, the interest points which have the highest promise of repeated extraction throughout multiple frames are extracted using a combination of the Hessian and Laplacian operators [11]. The binocular stereo ranges are calculated to various points using the well-known Marr–Pogio–Grimson algorithm [6].

To aid the process of interest point matching, each vector, (F, y_j, z_j) corresponding to the jth interest point in the frame $m+1$, is derotated so that the image plane $m+1$ appears to be parallel to image plane m. The matching of interest points is performed in two passes. The goal of the first pass is to identify and store the top three candidate matches for each interest point in frame $m+1$. The second pass looks for multiple interest points being matched to a single point in frame m. The range computations are further improved (for three or more sequential frames) by predicting and smoothing the range to each interest point that can be tracked through multiple frames.

The output binocular stereo and motion stereo range files, and simulated IRU data files are read into the Kalman filter software. "Frame i" (sequence i) processing consists of the following steps: the left and right binocular stereo images (L_i and R_i) are matched; the left image frames L_i and L_{i+1} (of sequence $i+1$) are matched by motion analysis. The filter software runs a range matching algorithm to detect coincident range points. For each coincident point the corresponding H-matrix and filter measurements are calculated and processed by the filter.

Results from processing "Frame 1" of the sequence with the Kalman filter is shown in Table II. The results of Table II are simulated with IRU noise and a 1 Hz video frame iteration rate, which equates to

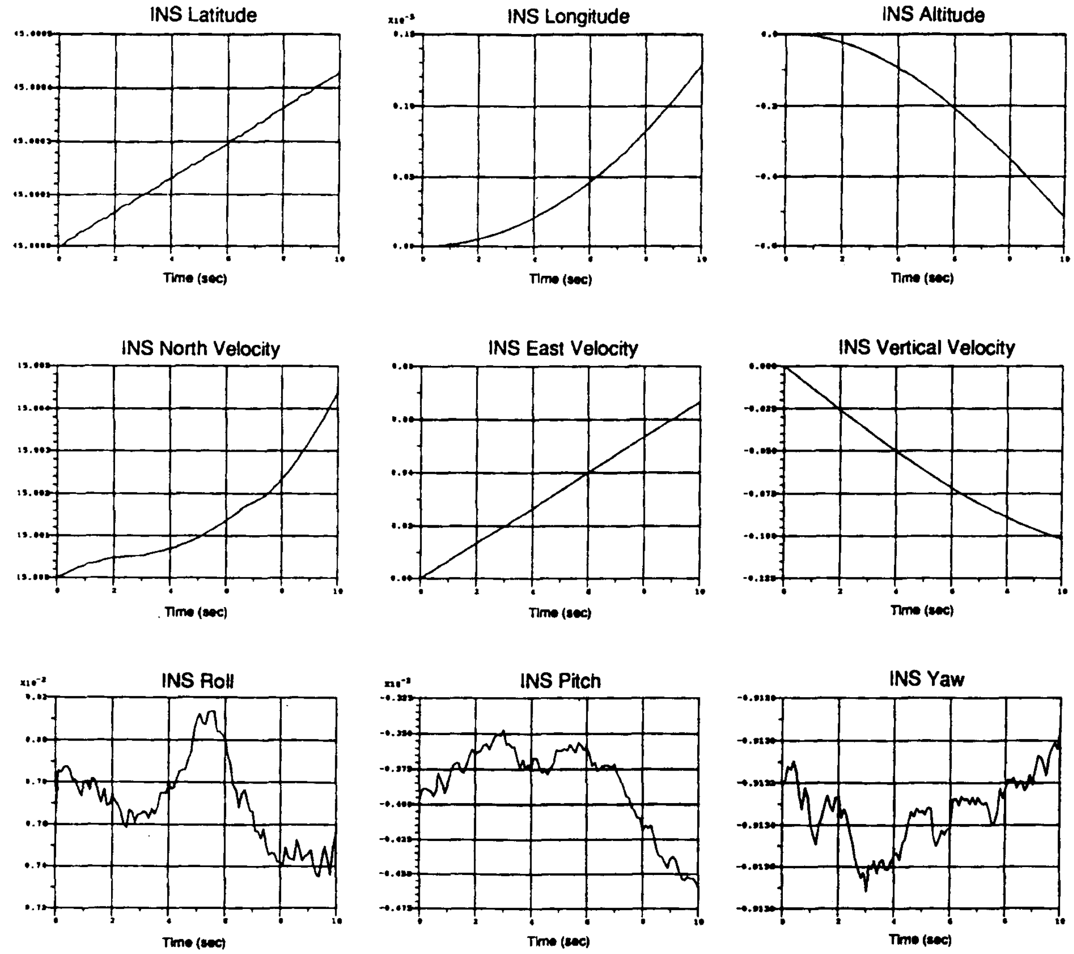


Fig. 8. Simulation of GG1328 gyro-based IRU.

TABLE II
Kalman-Filter-Computed Range Errors for 1 Hz Processing Rate (2 ft/s velocity)

Time = 1.0 sec (Frame 1)						
Measurement	**Raw Binocular Range**	**Raw Motion Range**	**Kalman Filter Binocular Error**	**Kalman Filter Motion Error**	**Corrected Binocular Range**	**Corrected Motion Range**
1	23.470947	13.229049	9.257078	−1.906973	14.213869	15.136022
2	15.250125	19.987539	−2.715582	−1.591671	17.965708	21.579210
3	23.286278	11.497955	5.687111	−4.775131	17.599167	16.273087
4	17.710770	22.075123	−1.643602	1.591987	19.354372	20.483137
5	13.850588	20.908190	−3.369545	7.472088	17.220133	13.436102
6	15.973729	21.540310	−3.376971	2.922868	19.350700	18.617441
7	16.092087	17.760782	1.714687	2.406360	17.806774	15.354422
8	16.151932	14.471107	−2.965745	2.350610	19.117678	12.120497
9	21.183895	22.302511	2.563853	2.367056	18.620041	19.935455
10	15.358275	14.538172	−3.013412	1.480840	18.371687	13.057332
11	18.167021	20.215263	−0.081987	1.600948	18.085033	18.614315
12	15.797955	21.457747	−1.679191	1.594449	17.477146	19.863297

TABLE III
Ground Truth Measurements for 1 Hz Processing Rate

Time = 1.0 sec (Frame 1)							
Measurement	y_l' (pixels)	z_l' (pixels)	y_r (pixels)	z_r (pixels)	y_l (pixels)	z_l (pixels)	R_f (ft)
1	-152	-59	-182	-53	-128	-51	13.9764
2	-52	122	-131	114	-46	111	18.8474
3	-47	-50	-94	-38	-38	-43	12.2229
4	-12	113	-83	106	-9	104	19.1300
5	41	65	-58	57	37	60	19.2826
6	49	202	-39	189	44	184	20.2861
7	92	-61	0	-49	82	-53	18.0941
8	95	167	0	149	82	144	14.0358
9	124	-15	51	-8	113	-12	22.3442
10	157	152	48	143	134	138	14.2818
11	170	-8	81	-1	153	-5	20.1257
12	170	-16	71	-9	154	-12	21.4648

TABLE IV
Kalman-Filter-Computed Range Errors for 10 Hz Processing Rate (20 ft/s velocity)

Time = 0.2 sec (Frame 1)						
Measurement	Raw Binocular Range	Raw Motion Range	Kalman Filter Binocular Error	Kalman Filter Motion Error	Corrected Binocular Range	Corrected Motion Range
1	23.470947	12.980159	9.617279	-1.099180	13.853668	14.079339
2	15.250125	19.459459	-2.733061	0.434156	17.983187	19.025303
3	23.286278	10.937357	5.861856	-2.925393	17.424423	13.862750
4	15.922381	14.836158	-1.956218	-4.532531	17.878599	19.368689
5	17.710770	22.263815	-1.671963	2.354116	19.382732	19.909698
6	13.850588	21.918085	-3.384527	9.700621	17.235115	12.217464
7	15.973729	21.828583	-3.465710	6.287485	19.439438	15.541098
8	16.092087	18.314566	-1.659888	0.328105	17.751974	17.986462
9	16.151932	14.771038	-3.026224	2.168971	19.178156	12.602067
10	21.183895	22.844698	2.680561	1.348993	18.503334	21.495705
11	15.358275	14.761926	-3.051116	0.234672	18.409391	14.527253
12	18.167021	20.570257	0.180036	0.792287	17.986984	19.777969
13	15.797955	21.834696	-1.613573	0.830073	17.411528	21.004623
Time = 0.3 sec (Frame 2)						
1	15.395482	15.375104	-2.823521	-1.199930	18.219004	16.575033
2	15.055490	18.953318	-3.251050	-3.732748	18.306541	22.686066
3	15.658957	19.364252	-2.859818	-3.794225	18.518774	23.158478
4	14.771465	16.280893	-3.150485	-4.380288	17.921949	20.661182
5	16.188807	18.075811	-3.643293	6.040530	19.832100	12.035282
6	16.690779	18.752043	-2.562827	4.724038	19.253607	14.028005
7	15.174622	18.209389	-3.081228	3.214266	18.255850	14.995123
8	19.333355	24.704124	-1.267558	6.274734	20.600912	18.429390
9	14.181705	17.482733	-3.954892	1.894178	18.136597	15.588555
10	16.496565	16.344479	-3.117268	0.850061	19.613832	15.494417
11	14.511797	16.740181	-4.223449	1.269790	18.735247	15.470390
12	15.358172	18.337030	-2.874749	2.921894	18.232922	15.415136
13	13.100904	19.790014	-4.306006	2.210015	17.406910	17.580000

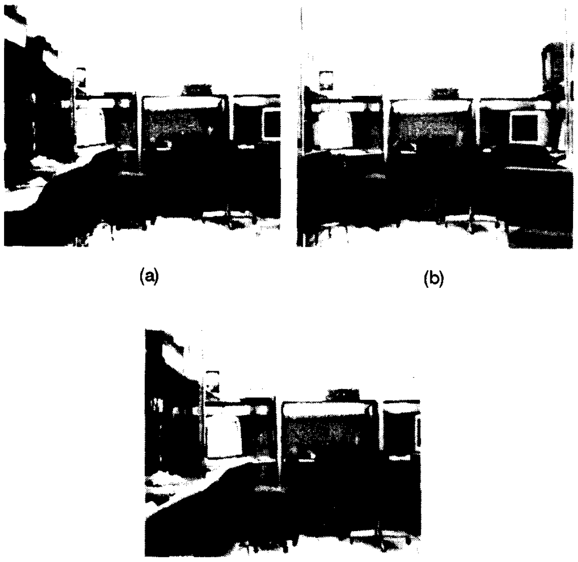


Fig. 9. (a) Frame 1 image obtained from left camera of stereo pair. (b) Frame 1 image obtained from right camera of stereo pair. (c) Frame 2 image obtained from left camera of stereo pair.

a velocity of 2 ft/s. Ground truth measurements for the 12 matched feature point locations of Frame 1 are presented in Table III. The center of the image plane is the origin of the pixel coordinates; (y_l, z_l) in the left image matches (y_r, z_r) in the right image of Frame i, while (y_l', z_l') in the left image of "Frame $i+1$" matches (y_l, z_l).

As shown in Table II, the corrections added to the binocular stereo range and motion stereo range tend to converge the solutions to a common point as expected, i.e., the corrected range values are in the direction (increasing or decreasing) as that of the ground truth values with respect to the raw range values. In general this behavior can be observed in the results for measurements 1 through 12. There are some exceptions (measurement 8 and 10) which could possibly be due to the measurement weighting. Since the results are for only a pair of frames, the actual convergence of the corrected range values cannot be seen.

Table IV contains results from processing the first and second "frames" for the 20 ft/s velocity case (10 Hz video frame iteration rate). Results for Frame 1 processing are good; the revised range estimates for the binocular and motion stereo ranging algorithms are converging to a unique value with the exception of measurements 3 and 9. Ground truth measurements for the 13 matched feature point locations of Frame 1 and the 13 matched feature point locations of "Frame 2" are presented in Table V. It is to be noted that the same scene points matched in Frame 1 will not necessarily appear in the results of processing Frame 2. Therefore, the ground truth range values of the corresponding measurements between Frame 1 and Frame 2 in Table V are not for the same scene point.

V. CONCLUSIONS

The basic concept and results of our binocular and motion stereo synergistic system have been presented. These results demonstrate that it is possible to effectively combine binocular range and stereo range measurements by incorporating a blending filter. This approach has the potential of providing dense range measurements. We plan to do this in the future.

REFERENCES

[1] Bhanu, B., and Roberts, B. (1989)
Obstacle detection during rotorcraft low-altitude flight and landing.
Honeywell Systems and Research Center Report to NASA, Apr. 1989.

[2] Bhanu, B., and Roberts, B. (1990)
Obstacle detection during rotorcraft low-altitude flight and landing.
Honeywell Systems and Research Center Report to NASA, July 1990.

[3] Bhanu, B., Roberts, B., and Ming, J. (1989)
Inertial navigation sensor integrated motion analysis.
In *Proceedings of the DARPA Image Understanding Workshop*, May 1989.

[4] Farrell, J. A. (1976)
Integrated Aircraft Navigation.
New York: Academic Press, 1976.

[5] Faugeras, O. D., Ayache, N., and Faverjon, B. (1986)
Building visual maps by combining noisy stereo measurements.
In *Proceedings of the International Conference on Robotics and Automation*, Apr. 1986, 1433–1438.

[6] Grimson, W. E. L. (1981)
From Images to Surfaces: A Computational Study of the Human Early Visual System.
Cambridge, MA: MIT Press, 1981.

[7] Ignagni, M. (1989)
Error analysis of a strapdown terrestrial navigation system.
Technical report System and Control Sciences, Honeywell Systems and Research Center, Minneapolis, MN, 1989.

[8] Leung, M. K., Choudhary, A. N., Patel, J. H., and Huang, T. S. (1989)
Point matching in a time sequence of stereo image pairs and its parallel implementation on a multiprocessor.
In *Proceedings of the Workshop on Visual Motion*, Irvine, CA, Mar. 1989.

[9] Li, L., and Duncan, J. H. (1989)
Recovering 3-D translational motion and establishing stereo correspondence from binocular image flows.
In *Proceedings of the Workshop on Visual Motion*, Irvine, CA, Mar. 1989.

[10] Matthies, L., Kanade, T., and Szeliski, R. (1989)
Kalman filter-based algorithms for estimating depth from image sequences.
International Journal of Computer Vision, **3** (1989), 209–236.

TABLE V
Ground Truth Measurements for 10 Hz Processing Rate

Measurement	y_l' (pixels)	z_l' (pixels)	y_r (pixels)	z_r (pixels)	y_l (pixels)	z_l (pixels)	R_f (ft)
Time = 0.2 sec (Frame 1)							
1	−152	−59	−182	−53	−128	−51	13.9764
2	−52	122	−131	114	−46	111	18.8474
3	−47	−50	−94	−38	−38	−43	12.2229
4	−13	0	−93	13	−11	7	12.2720
5	−12	113	−83	106	−9	104	19.1300
6	41	65	−58	57	37	60	19.2826
7	49	202	−39	189	44	184	20.2861
8	92	−61	0	−49	82	−53	18.0941
9	95	167	0	149	82	144	14.0358
10	124	−15	51	−8	113	−12	22.3442
11	157	152	48	143	134	138	14.2818
12	170	−8	81	−1	153	−5	20.1257
13	170	−16	71	−9	154	−12	21.4648
Time = 0.3 sec (Frame 2)							
1	−158	−26	−218	−22	−137	−21	15.9601
2	−85	−7	−161	3	−76	−3	20.0576
3	−34	−71	−115	−53	−32	−59	22.9181
4	−33	−95	−118	−86	−30	−81	19.5996
5	79	205	−11	183	71	180	16.4573
6	85	−19	−3	−8	76	−13	17.7225
7	103	−73	5	−56	92	−61	17.7855
8	125	−43	47	−30	115	−34	23.7157
9	143	56	33	60	126	54	16.6384
10	145	138	46	125	126	126	15.4135
11	172	168	59	158	150	152	15.9483
12	204	−50	96	−35	181	−40	17.9549
13	207	50	85	52	185	49	19.1903

[11] Nagel, H. H. (1987)
Displacement vectors derived from second-order intensity variations in image sequences.
Computer Vision, Graphics, and Image Processing, **21** (1987), 85–117.

[12] Roberts, B., and Bhanu, B. (1990)
INS-integrated motion analysis for autonomous vehicle navigation.
In *Proceedings of the DARPA Image Understanding Workshop*, Sept. 1990.

[13] Sridhar, B., and Suorsa, R. (1991)
Comparison of motion and stereo sensors in passive ranging systems.
IEEE Transactions on Aerospace and Electronic Systems, **27** (July 1991), 741–746.

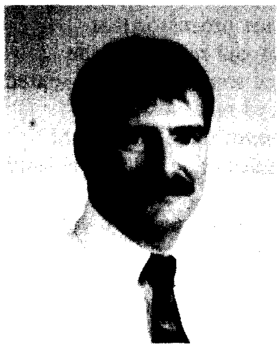
Bir Bhanu (S'72—M'82—SM'87) received the B.S. degree (with Honors) in electronics engineering from Institute of Technology, BHU, Varanasi, India; the M.E. degree (with Distinction) in electronics engineering from Birla Institute of Technology and Science, Pilani, India; the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge; the Ph.D. degree in electrical engineering from the Image Processing Institute, University of Southern California, Los Angeles; and the M.B.A. degree from the University of California, Irvine. He also received a Diploma in German from B.H.U., Varanasi, India.

Since 1991, Dr. Bhanu has been a Professor of Electrical Engineering and Computer Science and Director of Visualization and Intelligent Systems Laboratory at the University of California, Riverside. Prior to his current position, Dr. Bhanu was a Senior Honeywell Fellow at Honeywell Systems and Research Center. He has also worked with the Department of Computer Science, University of Utah, IBM San Jose Research Laboratory, INRIA-France, and Ford Aerospace and Communications Corporation. His current interests are computer vision, robotics, graphics and visualization, machine learning, artificial intelligence, and signal and image processing.

Dr. Bhanu has over 125 reviewed technical publications and 5 patents in the areas of his interest. He has given national short courses on intelligent automatic target recognition. He is a reviewer to over two dozen technical publications and government agencies. He is the guest editor of a special issue of IEEE Computer on "CAD-Based Robot Vision" published in August 1987, and guest co-editor of a special issue of *Journal of Robotic Systems* on "Passive Ranging for Robotic Systems" published in September 1992. He is the co-author of books on "Qualitative Motion Understanding" and "Genetic Learning for Adaptive Image Segmentation" published by Kluwer Academic Publishers. He is a member of ACM, AAAI, Sigma Xi, Pattern Recognition Society, and SPIE.

Peter F. Symosek was born in Lawrence, MA in 1953. He received the B.S. degree in electrical engineering from Merrimack College, North Andover, MA in 1978. He received the Sc.M. and Ph.D. degrees in electrical engineering from Brown University, Providence, RI in 1980 and 1985, respectively.

From 1978 to 1985 he was a Research Assistant at the Division of Engineering of Brown University. Since 1985, he has been working at Honeywell Systems and Research Center, Minneapolis, MN. He is Principal Investigator for all IR&D work for the development of model-based vision technology and multisensor algorithms for surveillance applications and for the mapping of state-of-the-art algorithms to VME-based, commercially available demonstration hardware. Currently, his research interests are in the fields of pattern recognition and scene analysis, the design and development of hybrid architectures for machine vision, and the design of robot vision systems which rely on the integration of multiple sources of visual information.



Scott I. Snyder received a B.S. degree in aerospace engineering from Iowa State University, Ames, in 1978.

In 1979, he joined Honeywell's Space and Aviation Division in Clearwater, FL, where he worked on the Space Shuttle flight control system. In 1983 Scott transferred to Honeywell's Systems and Research Center in Minneapolis where he has been active in inertial navigation and integrated navigation. He has designed, implemented, and tested integrated land navigation systems, GPS/inertial systems, and differential GPS systems.

Barry A. Roberts (S'85—M'88) received the B.S., M.S., and Ph.D. degrees in electrical engineering, all from Purdue University, Lafayette, IN, in 1982, 1984, and 1987 respectively.

During the 1987–1988 school year, he was a Visiting Assistant Professor in the School of Electrical Engineering at Purdue, teaching the graduate level course in computer vision and performing autonomous robot navigation research in the Robot Vision Lab. He joined the Systems and Research Center of Honeywell, Inc., in July of 1988 where he is currently a Principal Research Scientist in the Signal Processing Systems section. His current research interests include signal and image processing, computer vision (e.g., passive ranging using motion stereo and binocular stereo processing), and the application of machine learning techniques to computer vision problems.

Subhodev Das received his B.Tech. (Hons.) degree in electronics and electrical engineering from the Indian Institute of Technology, Kharagpur, in 1984, his M.S. degree in electrical engineering from the University of Hawaii, Honolulu, in 1986, and his Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 1991.

Currently, he is with the College of Engineering at the University of California, Riverside. His research interests include intelligent systems, computer vision, machine learning, human-computer interaction, parallel and distributed processing, and applications of artificial intelligence. He is a member of the IEEE Computer Society and the IEEE Robotics and Automation Society.